

ECO 434 : Economie Internationale

Gregory Corcos et Isabelle Méjean
Ecole Polytechnique, 2e année

2 septembre 2015

Table des matières

1	Introduction	7
1.1	Qu'est-ce que l'économie internationale ?	7
1.1.1	Les questions de l'économie internationale	7
1.1.2	Transactions réelles et transactions financières	9
1.2	Les flux internationaux de biens et services	14
1.2.1	Evolution du commerce mondial	14
1.2.2	Qui commerce avec qui ?	17
1.2.3	Nature des échanges	22
1.3	Les flux internationaux de capitaux	24
1.3.1	Evolution des échanges d'actifs financiers	24
1.3.2	Qui échange avec qui ?	30
1.3.3	Nature des échanges	32
1.4	Le marché des changes	33
2	Modèles néoclassiques de commerce	41
2.1	Le concept d'avantage comparatif	41
2.2	Le modèle ricardien	47
2.2.1	Hypothèses	47
2.2.2	L'équilibre autarcique	48
2.2.3	L'équilibre en économie ouverte	50
2.2.4	La détermination des salaires	54
2.2.5	Extensions et tests empiriques	55
2.3	Le modèle HOS	61
2.3.1	Hypothèses du modèle	62
2.3.2	L'équilibre : Autarcie versus économie ouverte	69
2.3.3	Ouverture et inégalités de revenus : le théorème de Stolper-Samuelson	73
2.3.4	Le rôle des dotations factorielles : le théorème de Rybczynski	77
2.3.5	Tests empiriques : Le paradoxe de Leontief	81
2.4	Conclusion	84

3	Concurrence imparfaite	87
3.1	Le modèle de Krugman	90
3.1.1	Ingrédients du modèle	90
3.1.2	Hypothèses du modèle	92
3.1.3	L'équilibre du modèle	96
3.2	Le modèle de Helpman-Krugman	102
3.2.1	Hypothèses du modèle	102
3.2.2	L'équilibre en économie ouverte	103
3.3	Evidences empiriques : L'équation de gravité	106
3.3.1	Microfondation de l'équation de gravité	106
3.3.2	Estimation	108
3.3.3	Les enseignements de l'équation de gravité	113
3.3.4	Limites	120
3.4	Conclusion	121
4	Solde Courant	123
4.1	Définitions	124
4.1.1	La Balance des Paiements	124
4.1.2	Solde courant et épargne nette	126
4.1.3	La position extérieure nette	128
4.1.4	Les taux de change nominal et réel	129
4.2	L'approche intertemporelle du solde courant	130
4.2.1	Le modèle de Fisher : un bien, un actif, deux périodes .	131
4.2.2	Deux biens, un actif, deux périodes : solde courant et change réel	134
4.3	Conclusion	138
5	Déséquilibres structurels	139
5.1	Croissance et flux financiers internationaux à long terme . . .	139
5.1.1	Equilibre en économie fermée	140
5.1.2	Equilibre dans une petite économie ouverte	143
5.2	Déséquilibres structurels mondiaux	147
5.2.1	Un aperçu des déséquilibres mondiaux dans les données	147
5.2.2	Explications possibles des déséquilibres structurels . . .	149
5.2.3	Un modèle avec marchés financiers incomplets	154
5.3	Conclusion	160
6	Taux de change	161
6.1	Change réel	161
6.1.1	L'hypothèse de Parité de Pouvoir d'Achat	161
6.1.2	TCR, prix des non-échangeables et revenu par habitant	165

6.1.3	TCR et compétitivité	167
6.1.4	Discrimination géographique par les prix (Pricing to Market)	169
6.2	Change nominal	170
6.2.1	Les parités de taux d'intérêt	170
6.2.2	Modèles monétaires du change nominal	174
6.3	Conclusions	184
7	Change fixe et flexible	187
7.1	Typologie des régimes de change	187
7.2	Ajustement en change fixe et flexible	189
7.2.1	Ajustement en change flexible	190
7.2.2	Ajustement en change fixe	193
7.2.3	Une comparaison de l'ajustement en change fixe et flexible	195
7.3	Crises de change	197
7.3.1	Crises de change de première génération	198
7.3.2	Crises de change de deuxième génération	203
7.3.3	Crises de change de troisième génération	208
7.3.4	Conclusion sur les crises de change	209
7.4	Conclusion	212
8	Politique économique dans la zone euro	215
	Bibliographie	216
	Exercices en Petites Classes	222

Chapitre 1

Introduction générale

1.1 Qu'est-ce que l'économie internationale ?

1.1.1 Les questions de l'économie internationale

L'économie internationale est l'étude des interdépendances entre économies nationales. Ces interdépendances prennent la forme de flux internationaux de biens et services, d'échanges d'actifs financiers, de migrations ou encore d'investissements directs à l'étranger. L'étude de l'économie internationale utilise les mêmes méthodes analytiques que les autres champs de l'économie. En particulier, les comportements des agents participant aux marchés internationaux sont dictés par les mêmes motifs que pour des transactions sur les marchés domestiques.

Certaines spécificités expliquent cependant le statut à part de ce champ. Ces spécificités concernent à la fois les déterminants microéconomiques de l'échange international et ses déterminants macroéconomiques. Au niveau macroéconomique, les échanges se font entre états indépendants, qui décident de leur politique économique sur la base d'intérêts purement nationaux. Cette spécificité implique que les politiques économiques choisies à l'optimal d'un pays peuvent se révéler néfastes pour le pays avec qui il échange. En outre, l'existence même de ces échanges implique que les nations souveraines ont à décider de la conduite de nouvelles politiques économiques. Ces nouveaux instruments de politique économique concernent la régulation des flux internationaux de biens et services (politique commerciale), la réglementation des flux de capitaux (régulation financière) ou encore la politique de change.

Au niveau microéconomique, la participation par des entreprises ou des institutions financières à l'échange international implique de prendre en compte les spécificités de ces marchés. Les coûts à l'échange y sont plus importants, du fait de l'existence de barrières tarifaires, de différences de régulation ou

encore d'écarts culturels. Ces coûts peuvent limiter la propension des agents économiques à participer à l'échange international. A l'inverse, l'échange international peut offrir de nouvelles opportunités de contournement des régulations nationales (optimisation fiscale par exemple). Pour estimer le volume des échanges et l'impact des régulations nationales sur ce volume, il est crucial de bien comprendre ce qui guide les stratégies microéconomiques. Tous ces éléments justifient d'étudier les flux internationaux de manière spécifique.

Les questions que se pose l'économie internationale sont étroitement liées à la nature des enjeux de politique économique spécifiques aux transactions internationales. Par exemple, la question de l'optimalité d'une politique commerciale régulant les flux de biens et services est étroitement liée à celle des gains au commerce. S'il est intuitif que certains flux de biens et services sont efficaces du point de vue économique (par exemple, les importations d'agrumes par des pays nordiques), la question de l'optimalité globale du libre-échange est beaucoup plus controversée, historiquement et aujourd'hui encore. Les modèles que nous étudierons dans ce cours montrent comment l'échange international de biens et services peut être source de gains en bien-être grâce à une utilisation plus efficace des ressources productives.

La question de l'échange international est également liée à celle des flux internationaux de capitaux. Ici aussi, la question de l'optimalité est largement débattue. Il n'est pas évident a priori de comprendre pourquoi il est bénéfique à un pays d'être un prêteur net sur les marchés internationaux ou, au contraire, de vivre au-dessus de ses moyens en empruntant à l'étranger (comme c'est le cas par exemple des Etats-Unis depuis le début des années 90). Pourtant, l'analyse macroéconomique montre comment de telles positions extérieures nettes permettent aux pays ouverts aux capitaux internationaux d'échanger des biens contre la promesse de biens futurs, ce qui améliore le lissage des flux intertemporels de revenus. En outre, l'échange d'actifs risqués entre les économies nationales peut conduire à une meilleure diversification des portefeuilles nationaux et donc, au niveau agrégé, réduire la variance des revenus.

Si les gains à l'échange international, de biens et services comme d'actifs financiers, sont positifs dans la plupart des cadres analytiques, cela ne signifie pas qu'il améliore la situation de tous les agents qui constituent l'économie nationale. Même avec des gains globaux positifs, certains groupes peuvent se retrouver négativement affectés par l'ouverture aux échanges internationaux. L'analyse économique s'inquiète évidemment de cette possibilité qui peut conduire à des politiques de redistribution visant à compenser les "perdants", voire à des politiques de régulation limitant les effets de l'ouverture sur la répartition des ressources. Dans ce dernier cas, les outils de l'économie internationale sont utilisés pour évaluer les pertes et les gains induits par de

telles politiques.

Historiquement, la question de la politique de change a été un thème central de l'économie internationale. Jusqu'au début des années 70, la plupart des économies contrôlent la valeur de leur monnaie en régulant les flux de capitaux et/ou en jouant le rôle de contrepartie de la demande nette de monnaie nationale pour maintenir une parité - en or ou en dollars. Pourtant, depuis le milieu des années 70, de nombreux pays ont mis en place des politiques de change flottant, laissant fluctuer la parité de leur monnaie nationale en fonction de l'offre et de la demande sur le marché des changes. Les politiques de change fixe se sont en effet révélées contraignantes pour les économies nationales, pouvant même mener à des crises financières lorsque la parité se révélait incohérente avec les fondamentaux de l'économie nationale.

Alors même que la tendance est au flottement des monnaies, l'économie mondiale a récemment vécu une expérience importante d'intégration monétaire. En 1999, 12 pays européens ont décidé d'abandonner leur souveraineté monétaire pour former une union monétaire. Une telle politique conduit *de facto* à fixer irrévocablement la parité des monnaies nationales à l'intérieur de la zone. Dans le même temps, le régime de change choisi pour la nouvelle Union Monétaire Européenne implique une flexibilité complète vis-à-vis des autres monnaies nationales. Les fluctuations de l'euro face aux grandes monnaies ne sont pas neutres pour les économies européennes. En se répercutant sur les prix relatifs des biens échangés, elles affectent la compétitivité de ces économies sur les marchés internationaux. Cet effet est hétérogène entre économies nationales, certains pays étant plus isolés des fluctuations de change, du fait de la structure de leur commerce. L'intégration monétaire conduit donc à des nouvelles questions de politique économique que l'économie internationale est amenée à traiter.

1.1.2 Transactions réelles et transactions financières

L'économie internationale se subdivise en deux champs. L'étude du commerce international est centrée sur les transactions internationales de biens et services, i.e. sur des transactions réelles. La macroéconomie internationale s'intéresse quant à elle à la contrepartie monétaire du commerce international, aux flux financiers internationaux. Evidemment, les deux types de transaction ne sont pas orthogonaux. Cette interdépendance des flux réels et financiers est illustrée dans la structure de la balance des paiements.

Pour comprendre l'interdépendance entre les flux réels et financiers, repartons de l'équation comptable de base en macroéconomie¹. Pour un pays,

1. Pour une description détaillée de cette équation comptable, voir le polycopié du

le produit intérieur brut (Y) mesure la valeur de l'ensemble des biens et services produits avec les facteurs productifs d'un pays. C'est aussi la somme des revenus perçus par les agents économiques nationaux, au cours d'une période donnée². Par définition, ces biens et services peuvent être utilisés pour différents types de dépenses : la consommation (C), l'investissement (I), les dépenses publiques (G) et, en économie ouverte, l'exportation nette (différence entre la valeur des exportations et la valeur des importations, $X - M$) :

$$Y = C + I + G + (X - M)$$

En notant S le montant épargné de la production nationale et en utilisant la définition de l'épargne comme la différence entre la production et la somme de la consommation privée et publique ($S = Y - C - G$), on en déduit :

$$S - I = X - M \tag{1.1}$$

La partie de gauche de l'équation (1.1) est l'épargne nationale (publique et privée) qui n'est pas utilisée pour investir dans des facteurs productifs nationaux. Quand $S - I$ est positif, ce supplément d'épargne est disponible pour être investi à l'étranger, il correspond aux flux monétaires internationaux entre l'économie nationale et le reste du monde. C'est ce qu'on appelle le solde financier. La partie de droite est l'excédent de la balance commerciale ou solde courant³. Il correspond à la demande nette de biens produits domestiquement par le reste du monde. Les fluctuations de cette demande nette conduisent à des fluctuations de la production nationale (et donc de l'em-

cours de Macroéconomie, Eco 432, Chapitre 3.

2. On ignore ici les revenus primaires perçus du reste du monde. En comptabilité nationale, on distingue le produit national brut (PNB) et le produit intérieur brut (PIB). Le produit intérieur brut mesure la production de biens et services produits à l'intérieur d'un pays, avec les facteurs de production localisés dans ce pays. Le produit national brut mesure la production de biens et services produits avec des facteurs appartenant aux résidents, que cette production se fasse sur le sol national ou à l'étranger. Le produit national brut est donc égal au PIB plus les revenus primaires nets perçus du reste du monde.

3. Dans ce cours, nous ne distinguerons pas, la plupart du temps, le solde de la balance commerciale et le solde de la balance courante (solde courant dans ce qui suit). En comptabilité nationale, le solde courant est le solde des flux monétaires résultant d'échanges internationaux de biens et services (solde de la balance commerciale) et d'échanges internationaux des revenus et transferts courants. Les flux internationaux de revenus incluent les salaires, dividendes et intérêts payés par un pays aux résidents d'un autre pays. Les transferts courants internationaux incluent les transferts de fonds de travailleurs étrangers ou les contributions aux organisations internationales. Ces deux dernières catégories ne sont généralement pas modélisées de sorte que le solde courant correspond, dans les modèles, au solde de la balance commerciale. La description de la balance des paiements dans le Chapitre 4 donne plus de détails sur cette distinction.

ploi). C'est pour cette raison que le solde courant est un élément important des discussions de politique économique.

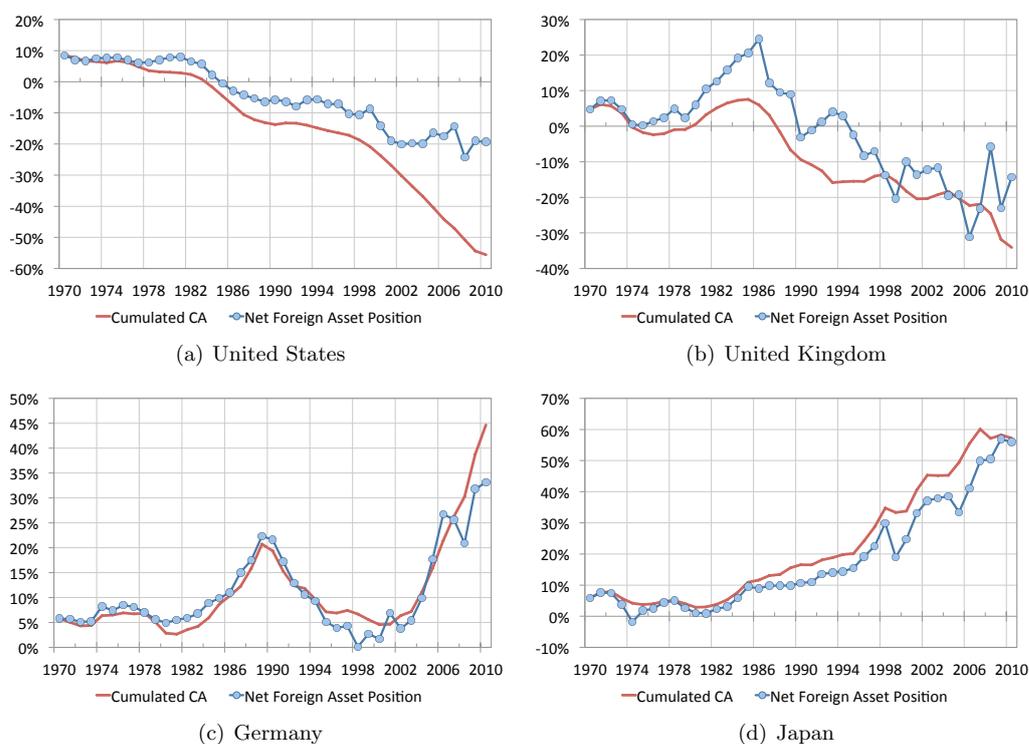
Le fait que la relation (1.1) entre solde financier et solde courant soit dérivée d'une équation comptable montre bien que, par définition, ces soldes sont intimement liés, ce sont les deux faces d'une même pièce. Un pays qui exporte plus que ce qu'il importe (la Chine ou l'Allemagne par exemple) est dit en excédent courant. Par définition, c'est aussi un pays qui épargne plus qu'il investit. Il n'a pas d'autre choix que d'exporter une partie de son épargne excédentaire pour l'investir à l'étranger. Il devient donc créancier du reste du monde. A l'inverse, un pays comme les Etats-Unis ou la France, en déficit courant récurrent, est un pays qui consomme plus que ce qu'il produit / qui investit plus que ce qu'il épargne. Pour financer cette consommation excédentaire, il a besoin d'emprunter à l'étranger, il devient donc emprunteur net vis-à-vis du reste du monde.

Le solde courant se répercute donc sur la position extérieure nette du pays. Un pays comme les Etats-Unis qui accumule des déficits courants en important plus qu'il exporte s'endette auprès du reste du monde et voit donc sa position extérieure nette se détériorer. C'est ce qu'on observe sur la Figure 1.1. Pour chacun des quatre pays considérés sont représentés le solde courant cumulé (somme des soldes courants accumulés au cours du temps) (ligne rouge) et la position extérieure nette du pays (ligne bleue). Les deux lignes sont très fortement corrélées. Par exemple, l'accumulation de déficits courants au Royaume-Uni depuis le début des années 90 a conduit à une détérioration de sa position extérieure nette, à un endettement du pays vis-à-vis du reste du monde. C'est aussi le cas des Etats-Unis mais, pour ce pays, les deux lignes sont plus dissemblables, la dégradation de la position extérieure nette étant moins forte que ce que ses déséquilibres courants cumulés ne le suggère. Cette spécificité américaine est liée à des effets de valorisation des actifs financiers (voir Gourinchas & Rey (2014) pour une discussion de ces effets).

L'interdépendance entre les flux réels et financiers est prise en compte dans la *balance des paiements*, un document comptable qui décrit l'ensemble des transactions du pays avec le reste du monde, au cours d'une période donnée. Les statistiques de balance des paiements sont une source utile de données en économie internationale. Comme toute table comptable, la balance des paiements inscrit les transactions au débit ou au crédit du pays. Les transactions au débit correspondent à des transactions conduisant à un paiement de l'économie nationale à l'étranger. Les transactions au crédit impliquent un paiement du reste du monde au pays considéré.

La structure et le fonctionnement de la balance des paiements sont décrits en détails dans le Chapitre 4. A ce stade, il est utile cependant de mentionner le principe de comptabilité en partie double utilisé pour établir

FIGURE 1.1 – Soldes courants cumulés et position extérieure nette, 1970-2010



Source: Lane and Milesi-Ferretti (2007a) updated to 2010.

Source : Gourinchas & Rey (2014).

les statistiques de balance des paiements : chaque transaction est enregistrée deux fois, au débit et au crédit. Pour comprendre ce principe, prenons un exemple. Supposons qu'une entreprise française achète un bien à une société américaine. Le paiement du bien par l'entreprise française est inscrit au débit de la balance courante de la France. Lorsque l'entreprise française paie le fournisseur, par exemple par virement bancaire sur son compte américain, une seconde transaction est enregistrée au crédit du compte financier. Cette transaction correspond au transfert d'un actif financier de la France vers les Etats-Unis, un titre de dette de la banque américaine sur la banque française. On retrouve bien le principe de l'égalité comptable décrit plus haut. L'achat par la France d'un produit américain participe au déficit courant français

(i.e. entre négativement dans le solde $X - M$ de l'équation (1.1)). Il conduit en outre à un endettement financier de la France vis-à-vis du pays partenaire (et entre donc négativement dans le solde $S - I$). Les deux transactions se compensent, l'équilibre (1.1) est maintenu et le solde de la balance des paiements est nul.

On peut résumer la littérature en économie internationale en considérant le type de flux internationaux étudiés. Les travaux en commerce international étudient principalement l'échange croisé de biens et services. Les questions que se pose cette littérature concernent la nature des biens qui sont échangés, le *volume* des échanges et les déterminants de la spécialisation des pays dans tel ou tel type de biens. Les économistes en commerce international s'intéressent donc à la partie de droite de l'équation (1.1), au solde courant. La plupart du temps et ça sera le cas dans les chapitres 2 et 3 de ce cours, on suppose pour simplifier que le solde courant est nul. Cette hypothèse éminemment contrefactuelle permet de concentrer l'analyse sur le volume et la composition des flux d'importation et d'exportation d'un pays.

Dans cette littérature, l'analyse est la plupart du temps *statique*. Au contraire, le caractère *dynamique* de l'échange est au coeur de la macroéconomie internationale, qui s'intéresse à la raison pour laquelle certains pays s'endettent sur les marchés internationaux, en important des capitaux étrangers. Cette littérature s'intéresse donc à un second type d'échanges internationaux, l'échange de biens et services contre des actifs financiers. Ces échanges déterminent le *niveau* du solde courant (et du solde financier) dans l'équation comptable (1.1). Comme nous le montrerons dans le chapitre 4 de ce cours, on peut interpréter un tel endettement international comme un achat de consommation présente en échange d'une promesse de consommation future, ie un commerce *intertemporel* de biens et services. Le pays prêteur net accepte de céder une partie de sa production courante de biens et services au pays importateur net en échange de la promesse d'être remboursé dans le futur, donc de pouvoir consommer plus.

Finalement, la finance internationale s'intéresse à un troisième type de transactions internationales, l'échange croisé d'actifs financiers. Cette littérature se concentre donc sur la partie de gauche de l'équation (1.1), sur le solde financier. Elle étudie le *volume* des échanges bruts d'actifs et la nature des actifs échangés. Le modèle le plus simple permettant de justifier cet échange croisé d'actifs financiers est un modèle de choix de portefeuille. Dans ce modèle, les marchés internationaux de capitaux offrent aux investisseurs des possibilités supplémentaires de diversification du risque. L'investisseur averse au risque choisit d'acheter des actifs financiers étrangers pour diversifier son exposition. C'est ce qui explique des échanges croisés d'actifs financiers à l'équilibre du modèle.

TABLE 1.1 – Exportations mondiales en pourcentage du PIB, sur longue période (en %)

1850	1880	1913	1950	1973	1985	1993
5.1	9.8 ^(a)	11.9 ^(a)	7.1	11.7	14.5	17.1

Source : Krugman (1995). Exportations de marchandises. ^(a) Calculé uniquement sur les pays de l'OCDE.

Dans ce qui suit, nous allons brièvement décrire les différents éléments entrant dans l'équation comptable (1.1). Nous étudierons d'abord les transactions réelles entrant dans le solde courant. Puis nous décrirons l'évolution et la structure des échanges d'actifs financiers internationaux. Nous concluons le chapitre par une analyse du marché des changes.

1.2 Les flux internationaux de biens et services

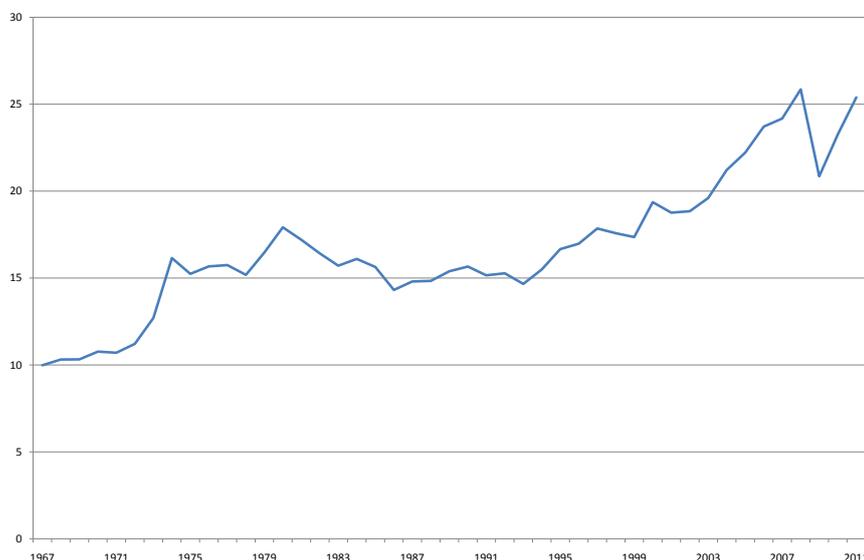
1.2.1 Evolution du commerce mondial

Au début du 21^{ème} siècle, les nations sont plus interdépendantes qu'elles ne l'ont jamais été. Au niveau mondial, la croissance du commerce est plus rapide que la croissance du PIB sur longue période, ce qui conduit à une augmentation tendancielle du ratio du commerce au PIB (Table 1.1 et graphique 1.2). Alors que le ratio des exportations au PIB mondial n'est que de 5% à la fin du 19^{ème} siècle, il atteint 25% en 2010. Cette évolution est évidemment hétérogène au cours du temps et entre pays.

Au cours du temps, on observe différentes "vagues" de globalisation, conduisant à une augmentation des flux internationaux de biens et services, suivies de périodes de protectionnisme et de contraction du commerce. Sur la période relativement récente, l'économie mondiale a connu une période forte de globalisation de la fin du 19^{ème} siècle à la première guerre mondiale. L'expansion des échanges a été la conséquence de la chute des coûts de transport (développement du chemin de fer et du bateau à vapeur) et des coûts de télécommunication (développement des réseaux télégraphes). L'expansion coloniale des économies européennes a également contribué au développement d'un commerce Nord-Sud, échange de matières premières contre des produits manufacturés.

Comme le montrent les chiffres du tableau 1.1, cette première vague de globalisation a conduit à un doublement du ratio du commerce à la produc-

FIGURE 1.2 – Exportations mondiales en pourcentage du PIB, 1967-2011 (en %)



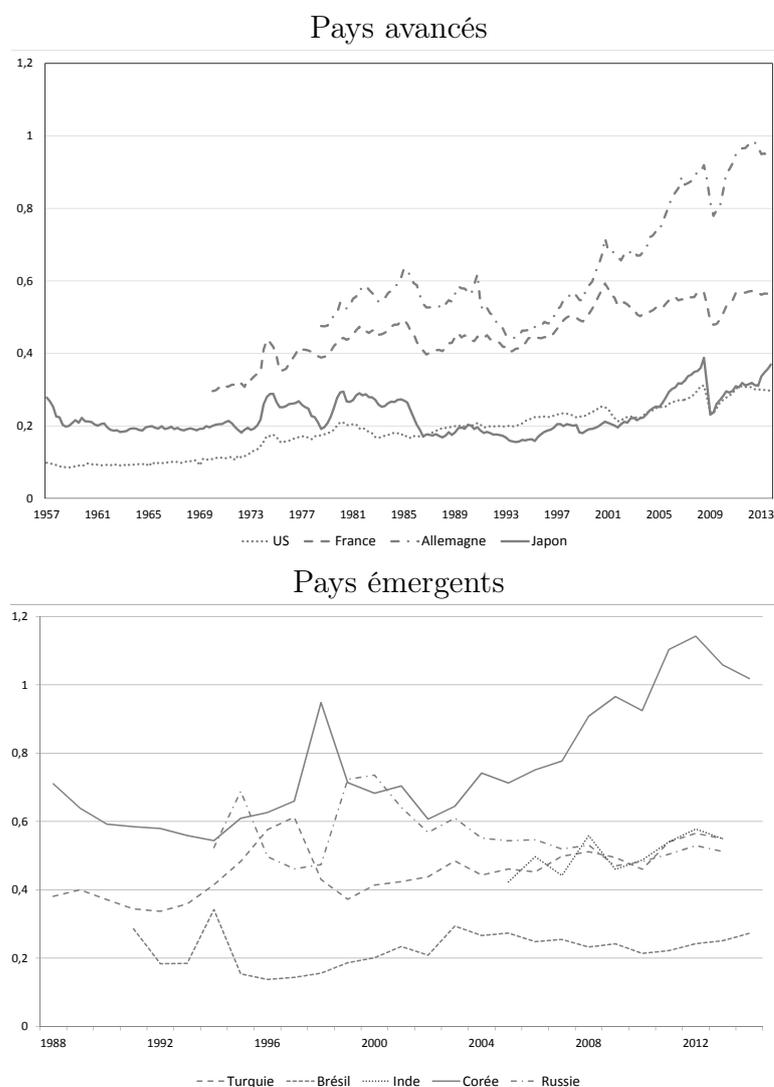
Source : CEPII-Chelem. Exportations de biens et services.

tion mondiale, qui atteint 11,9% à la veille de la première guerre mondiale. La guerre et, surtout, la dépression des années 30 ont mis fin à cet “âge d’or” du commerce mondial. Les pressions protectionnistes qui émergent en réponse à la hausse du chômage aux Etats-Unis et en Europe augmentent considérablement les coûts à l’échange international, qui se contracte dramatiquement en conséquence⁴. A la sortie de la deuxième guerre mondiale, le ratio des exportations à la production mondiale est revenu à 7.1%.

A partir des années 70, on observe une deuxième vague de globalisation, encore plus marquée que la première. De 1950 à 2011, le ratio des exportations à la production mondiale augmente de plus de 250%, pour atteindre 25% en 2011. Cette seconde vague de globalisation est elle aussi tirée par une chute des coûts de transport (apparition des containers et, plus récemment, transport aérien de marchandises) et de télécommunications (internet). La baisse des coûts à l’échange est amplifiée par les négociations internatio-

4. Aux Etats-Unis, la loi Smoot-Hawley de 1930 augmente les droits de douane à l’importation de plus de 20 000 biens. Cette loi conduit à des mesures de rétorsion des pays européens, qui augmentent également leurs taxes à l’importation. En conséquence de ces pressions protectionnistes, les importations américaines depuis l’Europe déclinent de 70% entre 1929 et 1932 alors que les exportations américaines vers l’Europe sont réduites de deux tiers. La loi Smoot-Hawley est considérée par les historiens comme un facteur aggravant important de la Grande Dépression.

FIGURE 1.3 – Degré d'ouverture au commerce, pays avancés et pays émergents



Source : IMF-IFS. Le degré d'ouverture est mesuré comme le ratio du commerce (importations + exportations) au PIB du pays. Le commerce couvre l'ensemble des flux de biens et services.

nales de libéralisation au sein du GATT (*General Agreement on Tariffs and Trade*, ancien nom de l'Organisation Mondiale du Commerce, OMC). Ces négociations conduisent à une réduction substantielle des barrières tarifaires (droits de douane, quotas, etc.). De 1950 à 1980, la croissance des échanges

se fait principalement entre économies développées, notamment au sein des zones de libre-échange que sont la Communauté Européenne et la zone de libre-échange nord-américaine.

A partir des années 80, on observe une émergence d'un certain nombre de pays du "Sud", notamment asiatiques. L'industrialisation rapide de ces pays induit une modification de la structure du commerce Nord-Sud, les exportations des pays émergents étant de plus en plus tournées vers les biens manufacturés. Si cette phase de globalisation des économies mondiales s'est accélérée à partir des années 90, sous l'effet de l'ouverture au commerce de la Chine, l'épisode récent de la crise de 2008 a montré que la croissance du commerce n'est pas irréversible. En 2007-2008, on a en effet observé une réduction massive des échanges internationaux, plus de trois fois supérieure à la contraction du PIB mondial. C'est ce que les économistes ont appelé le "Trade Collapse" (effondrement du commerce), un phénomène dont l'ampleur reste difficile à expliquer.

La croissance du commerce ne se fait pas de manière homogène entre les pays. Le graphique 1.3 montre l'évolution du taux d'ouverture au commerce pour un certain nombre de pays avancés et émergents. Dans cet échantillon, le taux d'ouverture, mesuré par le ratio du commerce au PIB, tend à augmenter sur la période d'observation. En 2014, ce taux d'ouverture s'établit entre 20% du PIB pour le Brésil et plus de 100% en Corée. Dans les données en coupe, on observe une forte hétérogénéité des taux d'ouverture au commerce, les grands pays comme les Etats-Unis ayant tendance à être moins ouverts au commerce que des plus petits pays comme l'Allemagne.

1.2.2 Qui commerce avec qui ?

Pour comprendre l'hétérogénéité des taux d'ouverture de différents pays, il est utile de regarder plus en détails la *structure* du commerce de ces pays. C'est ce que fait la littérature en commerce international. Il s'agit ici d'identifier et de comprendre pourquoi certains pays échangent plus que d'autres et la nature des biens qu'ils échangent.

Même au niveau agrégé, la matrice du commerce mondial est un objet qu'il est difficile de résumer car elle fait intervenir des milliers de relations bilatérales entre pays. Empiriquement, une manière efficace de prédire ces flux bilatéraux consiste à utiliser l'estimation économétrique d'une équation dite de gravité⁵. Si on note X_{ij} la valeur du commerce bilatéral entre deux

5. L'application de l'équation de gravité aux données de commerce international a été proposée par Tinbergen (1962). D'après Leamer & Levinsohn (1995), l'équation de gravité est une des relations les plus stables en économie.

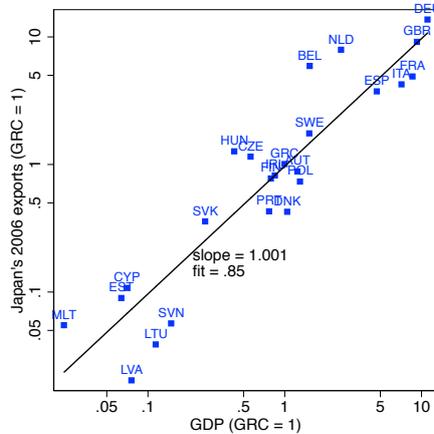
pays i et j , on peut montrer qu’une bonne estimation des flux bilatéraux peut être obtenue en utilisant les prédictions de l’équation suivante :

$$X_{ij} = \frac{AY_i^\alpha Y_j^\beta}{D_{ij}^\gamma}$$

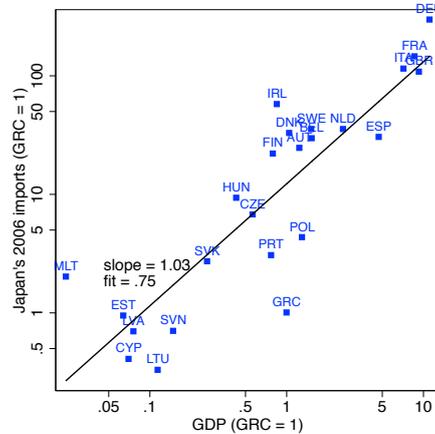
où A est une constante, Y_i (Y_j) le PIB du pays i (j) et D_{ij} la distance entre ces deux pays. α , β et γ sont des paramètres à estimer, qui se révèlent systématiquement positifs dans les exercices empiriques. L’analogie avec la théorie de Newton est donc immédiate : la force d’attraction entre deux pays, mesurée par les flux de commerce bilatéraux, est croissante de la “masse” de ces pays, mesurée par leur PIB, et décroissante de la distance les séparant. Les grands pays sont des acteurs importants du commerce international, à la fois côté importation et côté exportation. En outre, les flux internationaux de biens et services se concentrent au sein de zones géographiques relativement limitées.

FIGURE 1.4 – Volume du commerce et taille des pays

Exportations du Japon en UE,
2006



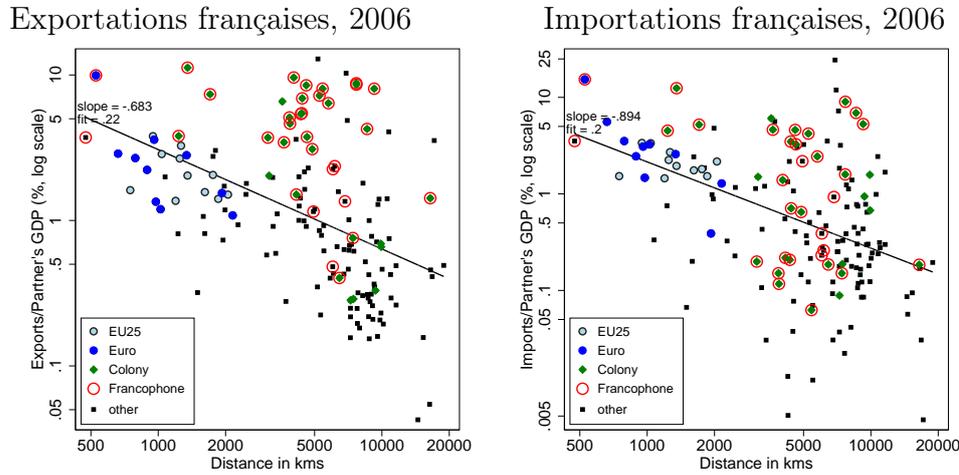
Importations du Japon en provenance de l’UE,
2006



Ces graphiques étudient la corrélation entre le commerce du Japon avec l’Union Européenne et la taille des pays membres. L’axe des abscisses représente le PIB des pays européens, en relatif par rapport au PIB grec. L’axe des ordonnées mesure la taille des exportations japonaises (Figure de gauche) et la taille des importations japonaises (Figure de droite), exprimées en termes relatifs par rapport au commerce avec la Grèce. Source : Head & Mayer (2014).

Ces relations sont illustrées dans les graphiques 1.4 et 1.5, extraits d’un chapitre du *Handbook of International Economics* écrit par Head & Mayer

FIGURE 1.5 – Volume du commerce et distance



Ces graphiques étudient la corrélation entre le commerce de la France et la distance qui la sépare de ses partenaires. L'axe des abscisses représente la distance en kilomètres à la France. L'axe des ordonnées mesure la taille des exportations françaises (Figure de gauche) et la taille des importations françaises (Figure de droite), exprimées en pourcentage du PIB du pays partenaire. Source : Head & Mayer (2014).

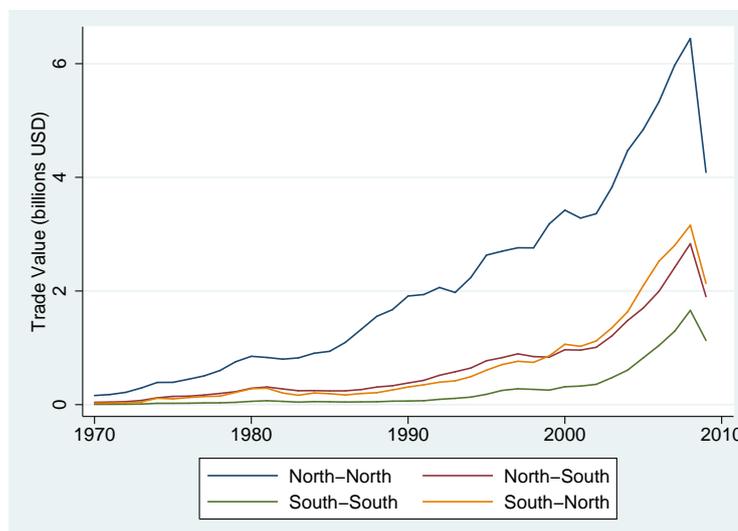
(2014) et consacré à l'estimation de l'équation de gravité. La figure 1.4 illustre la manière dont le commerce du Japon avec les différents pays membres de l'Union Européenne croît avec la taille du pays partenaire, du côté des exportations comme du côté des importations. La corrélation est forte, autour de 1, et le pouvoir prédictif élevé (plus de 70% de la variance expliquée par la variable de taille).

La figure 1.5 utilise quant-à-elle des données sur les flux bilatéraux de commerce entre la France et ses partenaires étrangers. Elle montre une relation significative et négative forte avec la distance au pays de destination : la France importe plus en provenance de et exporte plus à destination de pays proches géographiquement. En outre, conditionnellement à la distance, la France commerce relativement plus avec les pays francophones et/ou qui faisaient anciennement partie de l'empire colonial français⁶.

Nous reviendrons dans le chapitre 3 sur l'estimation de l'équation de gravité. A ce stade, il est important de noter qu'elle implique une *géographie* du

6. En effet, on remarque que les anciennes colonies de la France et les pays francophones se situent en moyenne au-dessus de la droite de corrélation entre la taille du commerce et la distance. Cela signifie que le commerce observé pour ces groupes de pays est en moyenne plus élevé que ce qui est prédit par une simple relation univariée basée sur la distance à la France.

FIGURE 1.6 – Géographie du commerce mondial



Source : UN ComTrade

commerce mondial, résumée dans la figure 1.6. Le graphique montre l'évolution du commerce depuis le début des années 70, en distinguant quatre catégories d'échange : le commerce entre pays développés ("North-North"), les flux de commerce entre pays développés et pays en développement ("North-South" et "South-North") et le commerce entre pays émergents ("South-South")⁷. Les échanges entre pays développés dominent le commerce mondial en valeur. En termes de contribution à la valeur du commerce mondial, le commerce Nord-Sud (dans un sens comme dans l'autre) est également important tandis que la part des échanges entre économies en développement reste minoritaire. Cette hiérarchie des flux d'échanges est cohérente avec l'équation de gravité. Les pays riches sont des acteurs importants du commerce international, à la fois côté offre et côté demande.

Une autre caractéristique de la géographie du commerce agrégé est liée à l'effet de la distance mis en évidence par l'équation de gravité. Le commerce bilatéral se concentre essentiellement au sein d'espaces géographiques limités. C'est ce que montrent les chiffres du tableau 1.2 qui décompose le commerce mondial par zones géographiques. En moyenne, les nombres situés sur la diagonale de la matrice sont élevés. La part du commerce intra-zone

7. Les pays sont classés dans les catégories "North" (pays avancés) et "South" (pays en développement) en fonction de leur niveau de PIB par tête au début de la période d'observation.

TABLE 1.2 – Concentration géographique du commerce mondial

	North Amer.	Latin Amer.	West. Eur.	CEECs ^(a)	Africa	Mid. East	Asia	Total
North Amer.	36.9	14.1	18.9	0.8	1.5	2.6	25.2	100
Latin Amer.	48.4	20.2	17.8	0.8	1.2	1.0	9.4	100
West. Eur.	8.2	2.5	68.1	4.2	2.8	3.0	9.5	100
CEECs	5.2	1.7	59.5	15.9	1.6	1.8	13.7	100
Africa	14.6	2.2	52.7	1.1	9.7	1.4	12.1	100
Mid. East	13.0	2.8	23.2	1.5	1.9	9.1	45.6	100
Asia	25.9	2.5	16.3	1.0	1.3	2.5	48.5	100

^(a) CEECs: Central and Eastern European Countries and former USSR

Source: Table 4, Chapter 1, Appleyard and Field (1998) *International Economics*, Third Edition

dans les échanges totaux est importante. Ainsi, pour l'Europe de l'Ouest, 68% des exportations sont destinées à d'autres pays européens. En dehors de la diagonale, on voit apparaître des liens importants entre zones proches géographiquement (Europe de l'Est et Europe de l'Ouest, Amérique latine et Amérique du Nord) et entre zones partageant des liens historiques (Afrique et Europe de l'Ouest).

Une des composantes de ce commerce intra-zone est liée au développement d'accords d'intégration régionale à partir de la deuxième guerre mondiale. En parallèle des négociations multilatérales au sein du GATT, un certain nombre de pays ont commencé à négocier des accords de libre-échange au niveau régional⁸. Ces accords de libre-échange facilitent les mouvements de biens et services au sein de la zone, tout en conservant un certain niveau de protectionnisme vis-à-vis du reste du monde. Une telle stratégie d'intégration régionale s'est révélée extrêmement efficace pour stimuler les échanges intra-zone. Dans le cas de l'intégration européenne par exemple, les données

8. Ce type d'accords d'intégration régionale déroge au principe de la "clause de la nation la plus favorisée" du GATT/OMC. Cette clause indique que les baisses de tarifs négociées au niveau bilatéral doivent automatiquement être étendues à tous les partenaires du pays concerné. Cette clause est un élément central du système de négociations multilatérales. Dans l'immédiat après-guerre, les pays signataires du GATT ont cependant accepté d'exclure de cette clause les accords de libre-échange régionaux, notamment ceux négociés au sein de la Communauté Economique Européenne.

récentes montrent que 70% des biens exportés par les pays membres sont vendus à d'autres partenaires européens. Ces flux de commerce intra-européen représentent un quart du commerce mondial. Pour les zones de libre-échange plus récentes, la proportion est généralement plus faible (autour de 50% pour les membres de l'Accord de Libre-Echange Nord-Américain, autour de 20% pour les zones de libre-échange d'Asie du Sud-Est et d'Amérique Latine). Cependant, la proportion reste importante. Alors même que les Etats-Unis et l'Europe négocient les bases d'un accord de libre-échange, cette dimension de la géographie du commerce est importante à prendre en compte.

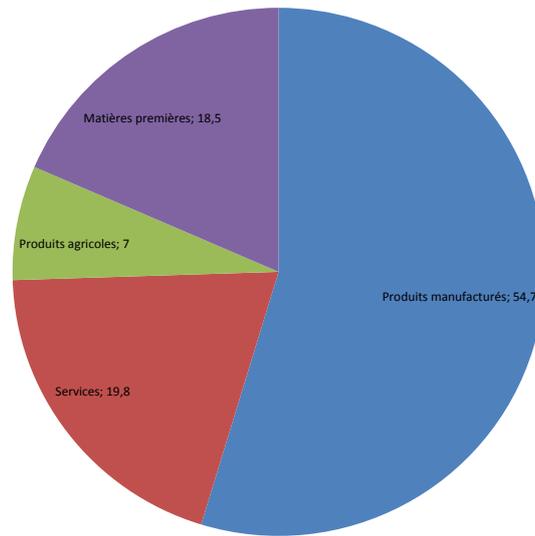
1.2.3 Nature des échanges

Le graphique 1.7 illustre la manière dont le commerce mondial se répartit entre différentes catégories de produits. En 2008, plus de la moitié de la valeur du commerce porte sur des biens manufacturés (voitures, vêtements, machines-outil, etc.). Les catégories de services et de matières premières représentent environ 20% chacune tandis que les produits agricoles représentent moins de 10% du commerce mondial. Dans la catégorie des matières premières, le pétrole occupe évidemment une place prédominante. Parmi les services facilement échangés, on peut enfin citer les services de transport, les dépenses de touristes à l'étranger ainsi qu'un certain nombre de services aux entreprises (activités des *call centers* à l'étranger par exemple).

Cette décomposition sectorielle du commerce international varie au cours du temps. Dans le passé, les biens manufacturés représentaient une part beaucoup plus faible des échanges (et de la consommation). En particulier, les Etats-Unis ont été longtemps un exportateur net de biens agricoles, produits à faible coût grâce à un accès à de vastes terres exploitables. Aujourd'hui, plus de trois quarts de la valeur des exportations américaines sont des produits manufacturés. Sur la période plus récente, les pays émergents ont également fortement contribué à la croissance du commerce de biens manufacturés. L'industrialisation d'un certain nombre d'économies émergentes a modifié la composition de leurs exportations. Alors qu'elles exportaient principalement des produits agricoles et des matières premières au début des années 60, la part des produits manufacturés est aujourd'hui proche de 70% (Figure 1.8). La Chine exporte par exemple plus de 90% de biens manufacturés.

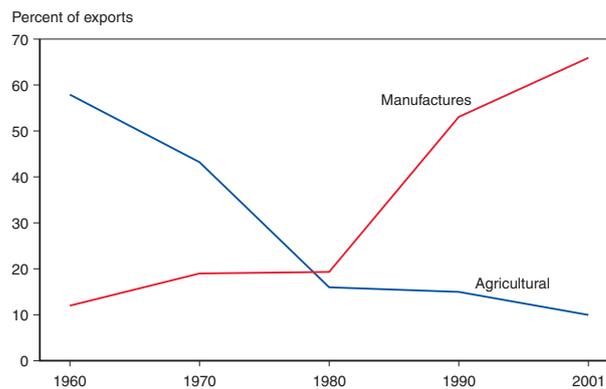
Si les services représentent encore une part relativement faible des échanges internationaux, de nombreux économistes s'attendent à observer une forte augmentation de la part des services dans le commerce mondial. Les progrès technologiques notamment liés aux technologies de l'information et de la télécommunication ont fait de certains services qui étaient jusqu'ici non échangeables des activités susceptibles d'être échangées au niveau interna-

FIGURE 1.7 – Décomposition sectorielle du commerce mondial



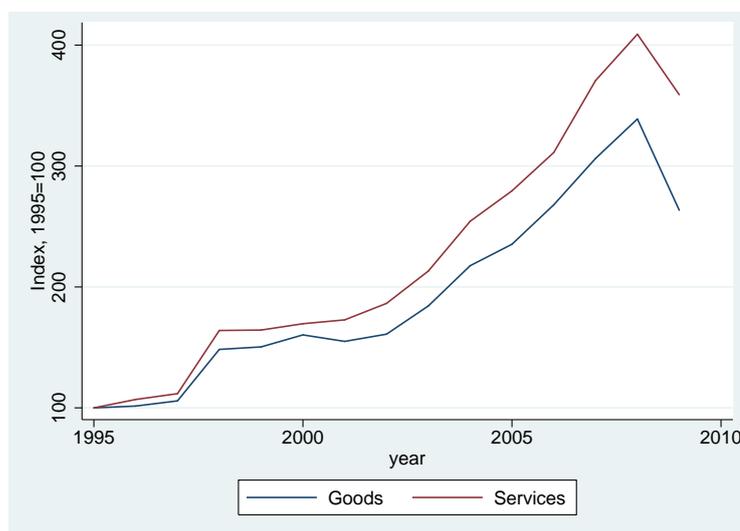
Source : OMC, cité par Krugman, Obstfeld & Melitz (2012). Les données concernent l'année 2008.

FIGURE 1.8 – Part des biens manufacturés et des produits agricoles dans les exportations des pays en développement



Source : United Nations Council on Trade and Development, cité par Krugman et al. (2012).

FIGURE 1.9 – Valeur des échanges mondiaux de biens et de services (base 100 en 1995)



Source : IMF-IFS.

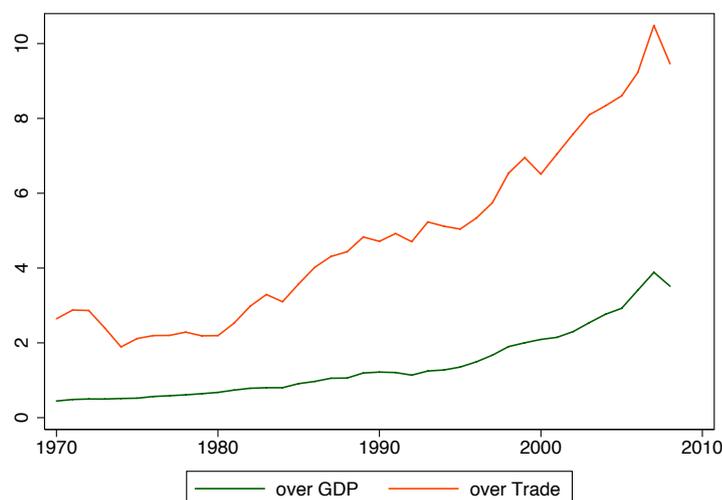
tional. C'est déjà le cas par exemple des activités de SAV opérées par des *call centers* à l'étranger. Dans le futur, cela pourrait être le cas d'autres activités (e.g. services de comptabilité aux entreprises). Depuis le milieu des années 90, la croissance des échanges de services est plus rapide que celle des échanges de biens (voir Figure 1.9). Cet écart de croissance explique que le commerce de services soit devenu un enjeu important dans la littérature récente en commerce international.

1.3 Les flux internationaux de capitaux

1.3.1 Evolution des échanges d'actifs financiers

On l'a vu l'échange international de biens et services va de pair avec un échange d'actifs financiers. La forte croissance des échanges internationaux de biens et services depuis le début des années 70 s'accompagne ainsi d'une croissance des flux internationaux d'actifs financiers, ce qu'on appelle parfois le phénomène de globalisation financière. Ce phénomène est illustré sur le graphique 1.10, pour les économies développées. De la même manière que l'ouverture au commerce était mesurée par le ratio du commerce total (expor-

FIGURE 1.10 – Intégration financière dans les économies développées



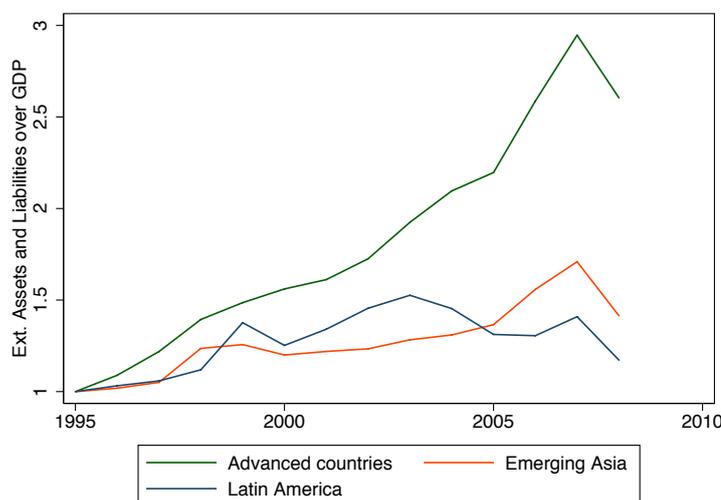
Source : Lane & Milesi-Ferretti (2007). L'intégration financière est définie comme la somme des actifs financiers étrangers et des dettes externes rapportée au PIB (ligne verte) et à la valeur du commerce (ligne rouge).

tations plus importations) au PIB, l'ouverture financière se mesure comme la somme des échanges d'actifs financiers entre une économie nationale et le reste du monde, en pourcentage de la production nationale. Entre 1970 et 2010, ce ratio a été multiplié par 7, passant de 5 à 35% dans l'ensemble des économies avancées. Ce phénomène de globalisation financière est particulièrement prononcé à partir de la deuxième moitié des années 90.

Si la valeur nette des flux commerciaux est nécessairement égale à la valeur nette des échanges d'actifs financiers, la somme des transactions financières entrantes et sortantes peut être bien supérieure à la valeur des exportations et des importations, du fait d'échanges croisés d'actifs financiers. C'est effectivement ce qu'on observe dans les données (ligne rouge sur le graphique 1.10). La valeur des transactions financières a cru beaucoup plus rapidement que la valeur des transactions réelles sur la période d'observation. Le taux d'ouverture financière des économies avancées est ainsi 9 à 10 fois supérieur à leur taux d'ouverture commerciale, en 2010.

Comme pour les transactions réelles, le phénomène de globalisation financière n'est pas homogène entre les pays. L'explosion des échanges d'actifs financiers à partir du milieu des années 90 est d'abord le fait des économies développées (graphique 1.11). Cette hétérogénéité tient notamment à

FIGURE 1.11 – Intégration financière, par régions



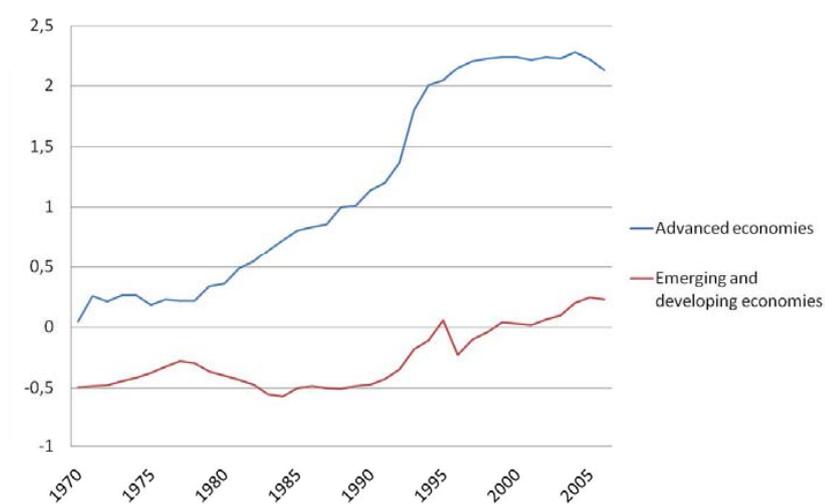
Source : Lane & Milesi-Ferretti (2007). L'intégration financière est définie comme la somme des actifs financiers étrangers et des dettes externes rapportée au PIB.

des différences entre pays en termes de régulation financière. Si les économies développées ont très largement dérégularisé les mouvements internationaux d'actifs financiers à partir des années 80, les régulations sont encore fortes dans un certain nombre de pays en développement.

La figure 1.12 illustre cette hétérogénéité à travers des mesures d'ouverture financière pour les deux types de pays. De la même manière que les échanges de biens et services sont régulés par l'existence de nombreux types de barrières protectionnistes (tarifs, quotas mais aussi barrières non tarifaires et réglementaires), la mesure de la régulation financière est un problème compliqué car elle résulte de l'existence de nombreuses règles freinant les mouvements internationaux de capitaux. Chinn & Ito (2006) proposent un indice basé sur des régulations observées sur les transactions réelles et financières. La séparation des pays en deux groupes, économies avancées et économies en développement, permet de mettre en évidence l'hétérogénéité entre pays dans le degré de dérégulation financière. L'augmentation de l'ouverture financière "officielle" des pays développés est beaucoup plus rapide que dans les pays en développement, notamment à partir des années 90.

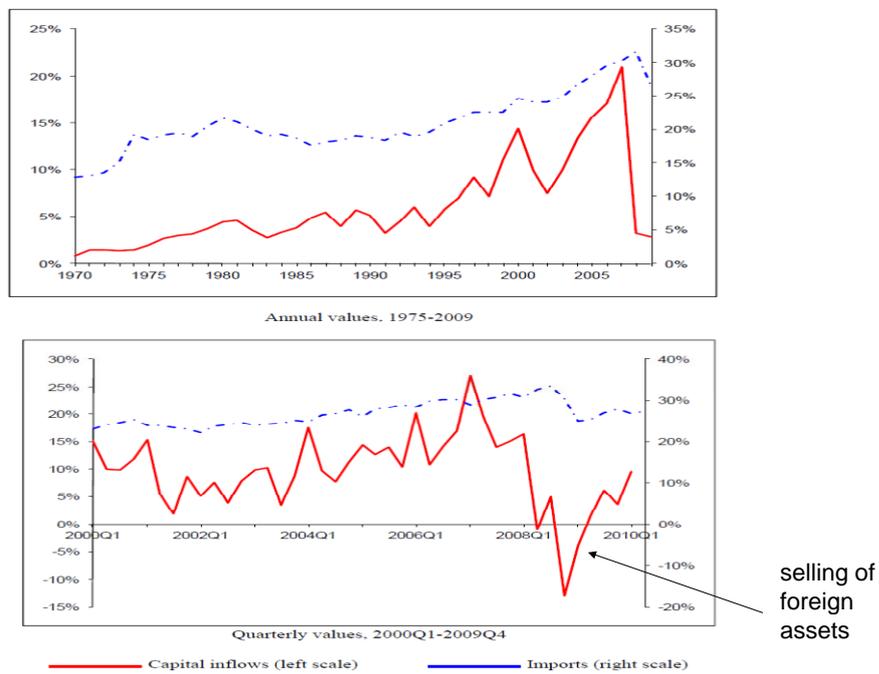
Une des raisons pour lesquelles les pays en développement sont encore réticents à libéraliser complètement les mouvements de capitaux tient au risque auquel cette libéralisation peut les exposer. Les flux internationaux

FIGURE 1.12 – Ouverture financière *de jure*, Pays avancés et pays en développement



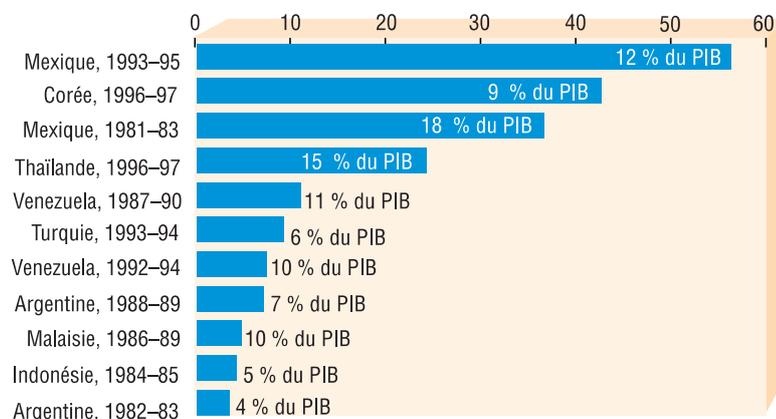
Source : Chinn & Ito (2006). L'indice d'ouverture financière est mesuré à partir d'une analyse en composantes principales sur quatre critères : i) l'existence de multiples taux de change, ii) l'existence de restrictions sur les transactions courantes, iii) l'existence de restrictions sur les transactions financières, iv) l'existence de règles de reddition sur les exportations.

FIGURE 1.13 – Volatilité des flux internationaux de capitaux



Source : Lane & Milesi-Ferretti (2007) et IMF-Balance of Payments. Les graphiques comparent le volume des flux internationaux d'actifs financiers à celui des transactions réelles. La ligne en rouge mesure le ratio des entrées de capitaux au PIB, au niveau mondial. La ligne en bleu est le ratio des importations au PIB, toujours mesuré au niveau mondial.

FIGURE 1.14 – Sorties massives de capitaux dans des épisodes de crise



Source : López-Mejía (1999). Le graphique illustre la taille des sorties de capitaux dans un certain nombre d'épisodes de crises dans les pays en développement. Les sorties sont mesurées en milliards de dollars et en pourcentage du PIB du pays.

de capitaux sont en effet extrêmement volatiles, du fait du fonctionnement même des différents marchés de capitaux. Cette volatilité est illustrée sur la Figure 1.13, en comparaison de la volatilité des flux internationaux de biens et services. Le graphique du haut montre l'évolution sur la période 1970-2008 du degré d'ouverture financière mondiale (mesuré par le ratio des actifs financiers entrants au PIB) et du degré d'ouverture commerciale (importations sur PIB). L'augmentation de l'ouverture au commerce est relativement régulière. Au contraire, le ratio des actifs au PIB est très volatile, notamment à partir du début des années 90, lorsque l'ouverture des économies aux flux internationaux de capitaux s'accélère. Cette volatilité atteint son apogée au moment de la crise financière de 2007-2008, pendant laquelle on observe une forte contraction des échanges internationaux d'actifs financiers. Si la volatilité est importante dans les données annuelles (graphique du haut), elle l'est encore plus lorsqu'on considère les données trimestrielles sur le graphique du bas.

Les pays en développement sont particulièrement exposés à la volatilité des flux de capitaux, notamment dans des épisodes de crises (voir chapitre 7 de ce cours). Cette exposition au risque est illustrée sur la Figure 1.14 qui montre le volume des sorties de capitaux (en niveau et en pourcentage du PIB) pour un échantillon de pays en développement lors d'épisodes de crises. Pour certains pays, les sorties de capitaux peuvent représenter une part massive du PIB (jusqu'à 15% du PIB en Thaïlande au moment de

la crise asiatique). En général, ces sorties se font sur une période courte. Un tel reflux de capitaux se révèle extrêmement coûteux pour le pays ce qui peut expliquer la réticence des économies en développement à libéraliser complètement leur compte financier.

Si les pays en développement sont encore à la traîne, le développement des échanges financiers internationaux reste rapide au niveau mondial, ce qui pose de nouvelles questions de politique économique. En particulier, la libéralisation financière crée des problèmes de “risque systémique” que les Etats ont beaucoup de mal à contrôler. De tout temps, les marchés financiers ont été fortement régulés du fait du rôle central des établissements bancaires dans le fonctionnement de l’économie. Le développement des transactions financières internationales rend la régulation beaucoup plus difficile. En effet, le marché international des capitaux n’est en fait qu’une collection de marchés nationaux, plus ou moins ouverts aux capitaux étrangers. Dans ce cadre, des régulations nationales trop exigeantes peuvent conduire les investisseurs nationaux à se détourner vers des marchés plus souples.

De telles possibilités d’arbitrage limitent *de facto* la capacité des Etats à réguler les innovations financières. Dans le même temps, l’existence même de transactions internationales rend l’ensemble du système plus fragile car il peut conduire à des phénomènes de contagion entre marchés. Conscientes de ce problème, les économies développées et, plus récemment, les économies émergentes ont créé un comité international chargé de coordonner la surveillance du système bancaire international. C’est le comité de Bâle qui définit un certain nombre de “bonnes pratiques” que les organes de surveillance nationaux sont supposés mettre en place. La crise financière de 2007-2008 a douloureusement prouvé que les règles imposées par le Comité de Bâle étaient nettement insuffisantes, ce qui a conduit à l’instauration de nouvelles règles prudentielles⁹.

1.3.2 Qui échange avec qui ?

Si l’équation de gravité a été utilisée de longue date pour prédire les flux bilatéraux de commerce, la finance internationale utilise depuis peu ce cadre empirique pour étudier la géographie des échanges internationaux d’actifs financiers. Si on note T_{ij} la valeur du flux d’actifs vendus par le pays i à des résidents du pays j , l’équation de gravité sur les échanges d’actifs financiers

9. Pour une description plus complète de ces questions, notamment dans le contexte de la crise financière de 2007, voir Krugman et al. (2012), Chapitre 21.

s'écrit :

$$T_{ij} = \frac{AM_i^\alpha M_j^\beta}{D_{ij}^\gamma}$$

où A est une constante, M_i (M_j) une mesure de la taille du marché financier dans le pays i (j) et D_{ij} la distance entre les pays, qui est corrélée à la taille des coûts de transaction financière.

Portes & Rey (2005) sont les premiers à appliquer ce cadre empirique aux échanges internationaux de titres d'actions. La taille du marché financier est mesurée par la capitalisation boursière du pays ainsi que par une mesure de la sophistication de ce marché. L'utilisation de la capitalisation boursière permet de rendre compte de la forte concentration des transactions sur quelques grosses places financières (Londres, New York, Hong-Kong, Singapour, Zurich, Séoul, Tokyo, etc.). Le coût des transactions financières est approximé par la distance ainsi que par un certain nombre de variables *proxy* des coûts de transaction (volume du trafic téléphonique entre les pays, nombre de branches de banque du pays i dans le pays j , nombre d'heures pendant lesquels les marchés des deux pays sont ouverts simultanément, etc.). Ils montrent que ce cadre empirique a une bonne qualité prédictive des flux bilatéraux d'actifs. 55% de la variance s'explique par la taille des marchés financiers et leur sophistication. Ce nombre augmente à près de 70% lorsque la distance est utilisée comme contrôle. L'élasticité des flux d'échange d'actions à la distance est égale à -0.9, presque aussi élevée que pour les flux de transactions réelles. Elle diminue mais reste très significative lorsque d'autres variables *proxy* des coûts de transaction sont également prises en compte.

La prise en compte de ces frictions à l'échange est importante car elle interfère avec d'autres déterminants de l'échange d'actifs. Portes & Rey (2005) montrent ainsi que, sans contrôle pour les coûts de transaction sur les marchés financiers, l'effet de la covariance des indices boursiers est positif : Les flux d'échanges d'actifs tendent à être plus importants entre pays dont les indices boursiers sont positivement corrélés. Ce résultat est contraire aux modèles de portefeuille qui expliquent l'échange de titres financiers par une volonté des investisseurs de diversifier leur portefeuille d'actifs en échangeant avec des pays aux rendements peu corrélés à ceux des actifs nationaux. Ce coefficient positif est cependant un artefact de l'existence de frictions à l'échange. Une fois que la distance entre les pays est prise en compte, l'effet de la covariance des indices est bien négatif, comme prédit par un modèle de portefeuille.

Si les résultats initiaux de Portes & Rey (2005) portent sur des flux d'échange de parts dans des sociétés privées, plusieurs applications ont montré que l'équation de gravité est également un bon cadre empirique pour expliquer des stocks de créance bancaire (Aviat & Coeurdacier, 2007), des

flux d'investissement direct à l'étranger (Mody *et al.*, 2002) ou encore des flux bancaires internationaux (Buch, 2005). En première approximation, on peut donc dire que les échanges financiers se font principalement entre places boursières importantes et suffisamment proches.

1.3.3 Nature des échanges

Comme évoqué brièvement dans la section précédente, les échanges internationaux d'actifs financiers peuvent être de nature variée. Parmi ces actifs, on peut distinguer les titres de créances et les actifs risqués. Les titres de créances impliquent une dette de l'émetteur sur l'acheteur, remboursée à échéance fixe et avec un taux d'intérêt certain. Ces titres sont donc non-risqués puisque le rendement est invariant aux circonstances économiques. Cette catégorie inclut les bons du Trésor émis par un Etat et les obligations de sociétés privées. Les actifs risqués sont au contraire payés avec un rendement aléatoire. Dans cette catégorie, on trouve par exemple les parts du capital d'une société privée, qui représentent un titre de propriété sur les profits (par définition aléatoires) de l'entreprise. Il faut noter que l'échange international rend cette distinction un peu plus floue. En effet, l'acquisition de titres de créances ne permet pas à l'investisseur de s'assurer un rendement sans risque si le rendement est payé dans une devise différente de celle de l'acheteur. Le risque de change pèse ainsi sur tous les actifs financiers, ce qui justifie l'existence d'instruments de couverture contre ce risque de change.

La structure des marchés internationaux de capitaux peut se résumer par les différents acteurs qui y participent.

- Les banques commerciales sont probablement l'acteur le plus important. Elles organisent le système international de paiements en effectuant les transferts monétaires internationaux. Elles échangent des titres de créances sous la forme de prêts internationaux aux entreprises, aux gouvernements ou à d'autres banques (prêts interbancaires) ou encore d'achats obligataires. Les banques conduisent également sur les marchés internationaux des activités de courtage en aidant les sociétés privées à vendre leurs titres d'action et/ou d'obligation sur les marchés étrangers.

Aujourd'hui, la plupart des grandes banques commerciales ont des filiales dans plusieurs centres financiers dans lesquelles elles conduisent ces activités. Ces filiales peuvent aussi participer à des échanges *off-shore* de monnaie, en acceptant des dépôts en monnaie étrangère (Eurodollars par exemple)¹⁰.

10. En finance internationale, on appelle "Euro-monnaie" (*Eurocurrencies*) les dépôts

- Les sociétés privées, lorsqu'elles sont suffisamment grandes, participent elles-mêmes à l'échange international d'actifs financiers. Sur ces marchés, elles ont accès à un marché plus important pour financer leurs investissements sous forme d'emprunts auprès de banques internationales ou d'émissions d'obligations. Pour les entreprises privées, notamment celles exerçant leurs activités dans plusieurs pays, un avantage du financement sur les marchés internationaux réside dans les possibilités qu'ils offrent d'emprunter en monnaie étrangère.
- Les institutions financières non-bancaires (compagnies d'assurance, *hedge funds*, etc.) représentent une part croissante des transactions internationales. Comme les banques commerciales, elles peuvent s'occuper d'activités de courtage pour les entreprises privées ou les gouvernements. Elles peuvent également investir leurs capitaux dans des actifs plus ou moins risqués.
- Les banques centrales et autres agences gouvernementales participent aux marchés internationaux pour deux raisons. Elles interviennent sur les marchés pour des motifs de régulation. Elles peuvent également être amenées à emprunter sur les marchés internationaux dans des conditions meilleures que ce qu'elles trouvent sur leur marché domestique.

1.4 Le marché des changes

Avant de conclure cette introduction, il est important de discuter le fonctionnement d'un marché financier particulier qu'est le marché des changes. La politique de change occupera en effet une place importante dans ce cours, notamment dans les Chapitres 6 et 7. Avant de considérer cette dimension de l'économie internationale sous l'angle, abstrait, de la modélisation, il est utile de décrire brièvement le mode de fonctionnement de ce marché.

Sur le marché des changes, les transactions financières portent sur l'échange d'une devise contre une autre, à condition bien sûr que ces monnaies soient convertibles. En principe, n'importe quelle monnaie peut être échangée contre

d'une monnaie dans des établissements localisés en dehors de la zone monétaire. Par exemple, les dépôts de dollars américains dans des banques en Europe ou au Japon sont appelés "Eurodollars". De tels dépôts en monnaie étrangère ont vu le jour à la fin des années 50, principalement pour des motifs liés au développement du commerce international. Le marché des Euro-monnaies s'est cependant fortement développés en lien avec des problèmes politiques. Lors de périodes de tensions internationales, il peut arriver par exemple que les pays détenteurs de dollars ne souhaitent pas les déposer dans des banques américaines, ce qui favorise le développement des Euro-banques. Historiquement, ça a été le cas par exemple pendant la guerre froide ou au moment des chocs pétroliers.

n'importe quelle autre. En pratique, la plupart des transactions se font en utilisant une monnaie véhicule, très souvent le dollar. Par exemple, un agent qui souhaiterait échanger des roubles russes contre des roupies pakistanaises va d'abord vendre ses roubles contre du dollar, qu'il échangera ensuite contre des roupies. La raison pour laquelle les monnaies véhicules sont en général utilisées réside dans le fait que le marché est beaucoup plus *liquide* pour de telles transactions (il y a plus d'individus qui veulent échanger du dollar contre des roubles ou des roupies contre du dollar que des roubles contre des roupies).

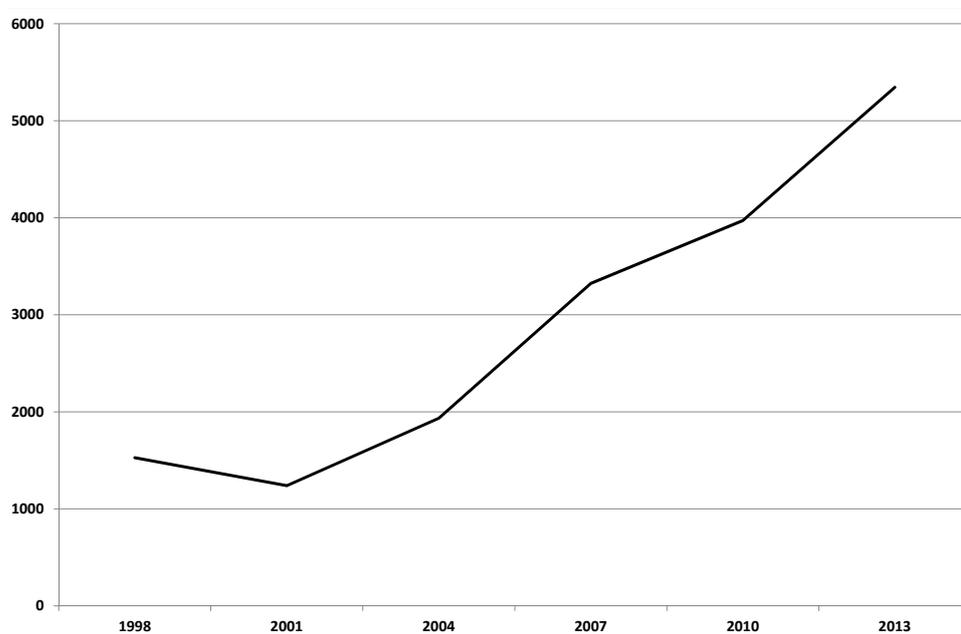
Le marché des changes est un marché en pleine expansion, comme le montre la Figure 1.15 tirée d'un rapport de la BIS (*Bank for International Settlements*). En avril 2013, le volume moyen des échanges journaliers atteint 5 345 milliards de dollars. Ce volume n'était "que" de 1 527 milliards de dollars en 1998. Malgré la taille du marché, la plupart des transactions s'y font de gré à gré, moins de 1% du trafic passant par des marchés organisés. Cependant, ce marché étant très liquide, la convergence des taux de change bilatéraux est quasi-instantanée.

Une caractéristique essentielle de ce marché est son extrême concentration. La plupart des transactions font intervenir un très petit nombre de monnaies différentes. C'est ce qui est illustré sur la Figure 1.16 qui donne la part de différentes monnaies dans le volume total des transactions. En 2013, 87% des transactions font intervenir le dollar. Les deux monnaies les plus utilisées si on exclut le dollar sont ensuite l'euro et le yen japonais. Alors que la monnaie unique européenne était supposée concurrencer le dollar dans son rôle de monnaie internationale, elle n'intervient que dans 33% des transactions, très souvent face au dollar. Le marché est également concentré dans l'espace (Londres, New York) et par le nombre d'intervenants qui y participent (surtout quelques grosses banques).

Sur ce marché interviennent différents types d'acteurs.

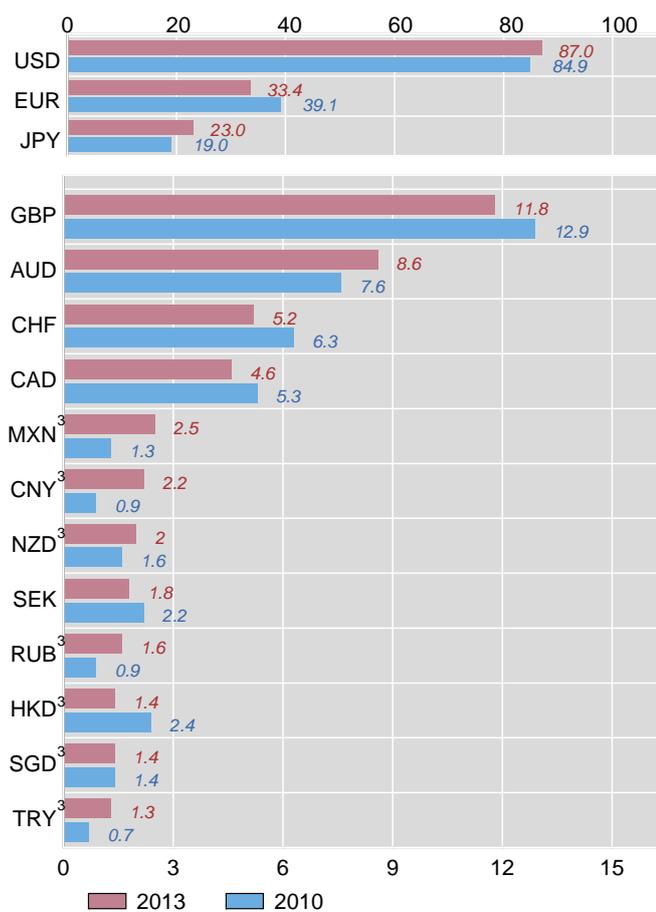
- En premier lieu, les banques commerciales qui sont en charge du système international de paiement. Presque toutes les transactions internationales de grande ampleur conduisent à une écriture au crédit et une écriture au débit de banques situées dans des centres financiers différents. Ces dépôts bancaires dénommés en différentes monnaies font ensuite l'objet d'une transaction de change, la plupart du temps interbancaire. L'avantage comparatif des banques à effectuer les opérations de change tient au grand nombre de transactions qu'elles gèrent simultanément, qui participe à la liquidité du marché malgré son fonctionnement de gré à gré. Si un individu ayant besoin d'échanger de l'euro contre de la couronne suédoise se rendait sur le marché des changes, il aurait probablement des difficultés à localiser l'individu qui accep-

FIGURE 1.15 – Volume des échanges total sur le marché des changes (en milliards de dollars par jour)



Source : BIS, Triennial Central Bank Survey (<http://www.bis.org/publ/rpfx13fx.pdf>). Le turnover est calculé en moyenne sur le volume journalier des échanges en avril.

FIGURE 1.16 – Part des différentes monnaies dans le volume total des échanges (en %)



Source : BIS, Triennial Central Bank Survey (<http://www.bis.org/publ/rpfx13fx.pdf>). Les chiffres sont basés sur la moyenne des transactions au cours du mois d'avril de l'année considérée. Les transactions faisant intervenir deux monnaies, par définition, la somme totale des parts est égale à 200%. L'indice ³ indique une surestimation potentielle du pourcentage pour 2010 du fait de problèmes méthodologiques.

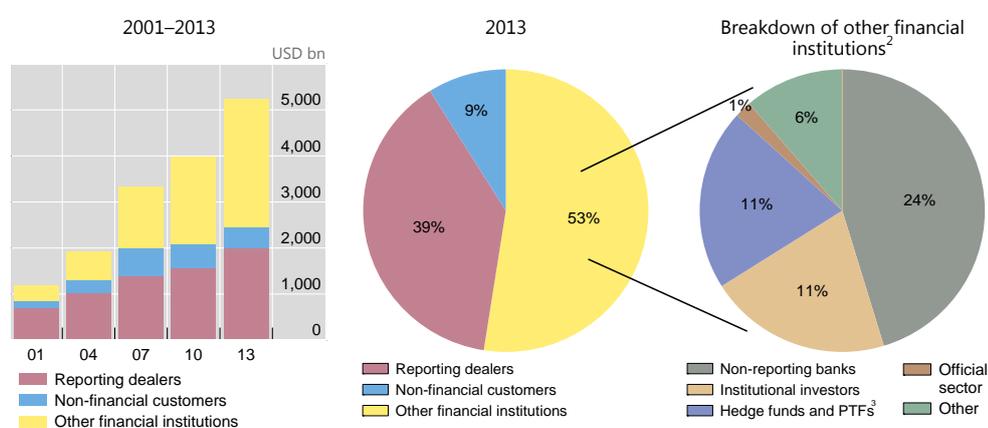
terait d'être la contrepartie de cette transaction. En mutualisant les besoins de nombreux clients, les banques commerciales économisent sur ce coût.

- Les entreprises privées peuvent aussi être amenées à participer directement au marché des changes. C'est particulièrement le cas des entreprises de taille suffisante, qui opèrent dans plusieurs pays. Une telle activité multinationale conduit l'entreprise à payer ses coûts et à encaisser son chiffre d'affaires dans des monnaies différentes, ce qui l'oblige à opérer des transactions de change.
- Les institutions financières non-bancaires participent également à la liquidité du marché des changes. Les transactions opérées par celles-ci sont similaires à celles opérées par les banques commerciales, avec qui elles entrent en compétition.
- Finalement, les banques centrales participent aussi au marché des changes sous la forme d'interventions ponctuelles visant à influencer le taux de change d'équilibre de leur monnaie. Typiquement, ces opérations portent sur des volumes très faibles mais ont un impact massif sur les prix d'équilibre, les interventions étant considérées comme un signal par les investisseurs privés.

En s'appuyant toujours sur le rapport de la BIS, la figure 1.17 montre la part des différents types d'acteurs dans le volume des transactions de change. Les banques commerciales effectuent plus de la moitié des transactions. Comme le montre le graphique de gauche, le poids relatifs des organismes financiers non-bancaires a cependant considérablement augmenté depuis le début des années 2000.

Les transactions sur le marché des changes sont de plusieurs natures. Chaque type de transactions se caractérise par un taux de change d'équilibre. On appelle transactions *spot* les échanges de monnaies délivrables sous 24 heures. Le taux de change négocié pour de telles transactions dépend donc des conditions économiques au moment de la transaction. Les transactions de type *forward* sont des promesses d'échange dans le futur à un taux négocié à l'avance, au moment de la transaction. Le contrat de transaction *forward* se caractérise donc par une maturité (date de livraison), un volume et un prix de vente. Celui-ci dépend de l'offre et de la demande de contrats *forward* au moment de l'échange mais aussi des conditions anticipées de l'économie au moment de la livraison (voir Chapitre 6). Les transactions peuvent aussi se faire sur des contrats d'*option* qui ressemblent à des contrats *forward* à la différence que les contreparties ne sont pas tenues d'effectuer la transaction au moment convenu, elles ont simplement une option pour le faire. En particulier, si le taux de change spot au moment convenu de la livraison est plus favorable que le prix établi sur le contrat, l'acheteur a la possibilité d'ache-

FIGURE 1.17 – Part des différents acteurs dans le volume total des échanges (en %)



¹ Adjusted for local and cross-border inter-dealer double-counting, ie "net-net" basis. ² For definitions of counterparties, see page 19. ³ Proprietary trading firms.

Source : BIS, Triennial Central Bank Survey (<http://www.bis.org/publ/rpfx13fx.pdf>). Les chiffres sont basés sur la moyenne des transactions au cours du mois d'avril de l'année considérée. La catégorie "Reporting dealers" comprend principalement des banques commerciales et des grandes banques d'investissement (y compris les filiales à l'étranger) qui sont soumises spécifiquement à l'enquête de la BIS. La catégorie "Other financial institutions" inclut toutes les autres institutions financières qui n'appartiennent pas à la catégorie précédente, notamment les institutions financières non-bancaires et les banques centrales. Finalement, la catégorie "Non-financial customers" inclut des entreprises privées, des organismes gouvernementaux non financiers et des individus qui investissent sur le marché des changes à titre privé.

TABLE 1.3 – Répartition de la valeur des échanges de devises entre les différents types de contrats (en milliards de dollars)

Global foreign exchange market turnover						
Net-net basis, ¹ daily averages in April, in billions of US dollars						
Instrument	1998	2001	2004	2007	2010	2013
Foreign exchange instruments	1,527	1,239	1,934	3,324	3,971	5,345
Spot transactions	568	386	631	1,005	1,488	2,046
Outright forwards	128	130	209	362	475	680
Foreign exchange swaps	734	656	954	1,714	1,759	2,228
Currency swaps	10	7	21	31	43	54
Options and other products ²	87	60	119	212	207	337
<i>Memo:</i>						
Turnover at April 2013 exchange rates ³	1,718	1,500	2,036	3,376	3,969	5,345
Exchange-traded derivatives ⁴	11	12	26	80	155	160

¹Adjusted for local and cross-border inter-dealer double-counting (ie "net-net" basis). ²The category "other FX products" covers highly leveraged transactions and/or trades whose notional amount is variable and where a decomposition into individual plain vanilla components was impractical or impossible. ³Non-US dollar legs of foreign currency transactions were converted into original currency amounts at average exchange rates for April of each survey year and then reconverted into US dollar amounts at average April 2013 exchange rates. ⁴Sources: FOW TRADEdata; Futures Industry Association; various futures and options exchanges. Foreign exchange futures and options traded worldwide.

Source : BIS, Triennial Central Bank Survey (<http://www.bis.org/publ/rpfx13fx.pdf>). Les chiffres sont basés sur la moyenne des transactions au cours du mois d'avril de l'année considérée.

ter ses devises sur le marché spot. Finalement, les *swaps* de devises sont des contrats relatifs à deux transactions, à deux dates dans le temps. Les contractants échangent des devises à un moment dans le temps, et se rendent ces devises plus tard. Un exemple de *swaps* peut être un accord entre deux entreprises qui se prêtent mutuellement de l'argent dans leur propre monnaie pour couvrir leurs besoins mutuels (temporaires) en monnaie étrangère. Le Tableau 1.3 illustre la répartition des transactions de change entre ces différentes catégories. Les transactions qui dominent très largement le marché sont les opérations *spot* ainsi que les *swaps* de monnaie.

Finalement, notons que les différentes transactions de change peuvent être liées à des motifs très variés. Les agents qui se positionnent sur le marché peuvent simplement avoir besoin de devises étrangères soit immédiatement (transactions *spot*) soit à une date connue (transactions *forward* et options). L'utilisation des marchés *forward* permet aux entreprises de se couvrir contre le risque de change. Les entreprises et institutions financières qui savent qu'elles auront des besoins en devises à une date connue (pour payer un fournisseur ou rembourser une dette en devises étrangères) et qui anticipent une dégradation des conditions auxquelles elles pourront acquérir ces devises

ont intérêt à payer une prime pour pouvoir s'assurer d'un taux de change dans le futur. En ce sens, ces contrats sont des instruments de couverture contre le risque induit par la volatilité des taux de change.

De nombreuses opérations de change sont également liées à des opportunités d'arbitrage. De fait, c'est l'existence même de ces comportements d'arbitrage qui conduit à la convergence des taux de change malgré l'organisation décentralisée du marché. Les traders sur le marché des changes étudient le menu de taux de change bilatéraux d'équilibre observé à un moment dans le temps sur une place donnée et effectuent des transactions de change de façon à tirer un bénéfice des écarts observés. Enfin, une partie importante des transactions relève de comportements spéculatifs. Les investisseurs privés sur le marché des changes forment des anticipations sur la valeur future des taux de change d'équilibre et, en fonction de ces anticipations, définissent des positions (achats et ventes) qui permettent (en principe) de gagner de l'argent sur ces fluctuations. La plupart des transactions de type spéculatif se font sur des variations de change de très petite ampleur, et sur des durées très courtes (intraday).

Chapitre 2

Les modèles néoclassiques de commerce international

“The theory of comparative advantage is the only result in social science that is both true and non-trivial” (attribué à Paul Samuelson)

2.1 Le concept d’avantage comparatif

La notion d’avantage comparatif est attribuable à David Ricardo (1772-1823), économiste et membre du parlement anglais, qui a travaillé sur de nombreuses questions d’économie politique. Dans son essai *On the Principles of Political Economy and Taxation* (1817, 1819, 1821), Ricardo élabore une théorie justifiant le libre-échange et s’oppose aux “*Corn Laws*” interdisant l’importation par l’Angleterre de blé étranger¹. Dans cet essai, il détaille notamment sa “théorie de l’avantage comparatif” selon laquelle tous les pays, même les moins compétitifs, ont un intérêt à participer à l’échange international. En s’ouvrant au commerce, ils peuvent se spécialiser dans la production des biens pour lesquels ils détiennent un avantage productif relatif, et acheter les biens qu’ils produisent avec un désavantage relatif. En ce sens, la théorie ricardienne vient compléter la “Théorie de l’avantage absolu” d’Adam Smith d’après qui il est prudent “de ne jamais essayer de faire chez soi la chose qui coûtera moins à acheter qu’à faire” (*La richesse des nations*, IV.3). Selon Adam Smith, l’échange est mutuellement avantageux dès lors qu’un pays est plus productif dans la production d’un bien tandis qu’un autre pays est meilleur dans la production d’un autre bien. Pour Ricardo, il n’est pas nécessaire que les pays disposent de tels avantages *absolus*. Même si un pays domine strictement l’autre dans la production des deux types de biens, il

1. Les “*Corn Laws*” seront finalement abrogées en Angleterre en 1846.

reste mutuellement avantageux pour lui de se spécialiser dans la production du bien pour lequel son avantage est le plus prononcé et d'importer le bien qu'il produit avec une productivité relativement plus faible.

Pour illustrer son propos, Ricardo développe un exemple basé sur l'échange de vin et de draps entre l'Angleterre et le Portugal. Supposons qu'avec une quantité de travail donnée, le Portugal produise 20 mètres de drap ou 300 litres de vin. Avec la même quantité de travail, l'Angleterre peut quant-à-elle produire 10 mètres de drap et 100 litres de vin. Par rapport au Portugal, l'Angleterre a donc un désavantage absolu dans la production de drap comme dans celle de vin (puisque'elle produit ces biens avec une productivité moindre). Dans la théorie de Smith, les deux pays n'ont pas intérêt à échanger puisque le Portugal est strictement meilleur à produire les deux biens et n'a donc pas intérêt à faire produire à l'étranger ce qu'il peut produire de manière plus efficace domestiquement. Ricardo montre cependant qu'il existe un échange mutuellement avantageux pour les deux pays, même dans cette configuration défavorable à l'Angleterre.

Plus spécifiquement, il montre que chaque pays a intérêt à se spécialiser dans la production du bien pour lequel il a un avantage *comparatif*. Si l'Angleterre est moins productive à produire et le drap et le vin, son désavantage comparatif est moindre dans le secteur du drap : le pays est deux fois moins productif que le Portugal pour produire le drap et trois fois moins productif pour produire le vin. En se spécialisant dans son avantage comparatif, l'Angleterre va pouvoir allouer l'intégralité de sa force de travail à la production de drap. En échangeant 10 mètres de drap avec le Portugal, le pays pourra obtenir jusqu'à 150 litres de vin ($10 \times 300 / 20$). Si, au contraire, le pays utilise le travail équivalent pour produire domestiquement du vin, il n'obtient que 100 litres. L'échange est donc avantageux pour l'Angleterre. Ce qui est moins intuitif (et contraire à l'intuition d'Adam Smith), c'est que le Portugal a également intérêt à participer à un tel échange, bien qu'il dispose d'un avantage absolu dans la production des deux biens. En se spécialisant dans la production de vin, le pays va produire suffisamment pour obtenir plus de drap en Angleterre que ce qu'il aurait pu produire domestiquement. Ainsi, avec 300 litres de vin produit domestiquement, le Portugal pourra obtenir 30 mètres de drap anglais ($300 \times 10 / 100$), au lieu des 20 mètres qu'il aurait produit domestiquement avec la même quantité de travail. Les deux pays ont donc un intérêt commun à échanger.

Pour bien comprendre la différence entre l'avantage absolu et l'avantage comparatif, il est utile de prendre un autre exemple. Considérons les matrices de productivités décrites dans le tableau 2.1. Celles-ci donnent la productivité mensuelle d'un travailleur chinois et d'un travailleur européen dans deux secteurs manufacturiers, le textile et l'automobile. Dans l'exemple (a), la

TABLE 2.1 – Exemple

Production mensuelle par travailleur (en unités de biens)

Exemple (a)			Exemple (b)		
	Chine	UE		Chine	UE
Chemises	200	50	Chemises	400	50
Voitures	5	10	Voitures	20	10

Chine a un avantage absolu dans la production de chemises : Un travailleur y produit 200 chemises quand le travailleur européen n'en produit que 50. L'Europe a cependant un avantage absolu dans la production de voitures : au cours d'un mois, un travailleur européen produit deux fois plus de voitures qu'un travailleur chinois. Dans cette configuration, l'échange est évidemment mutuellement avantageux. La Chine peut se spécialiser dans la production de chemises, dont une partie sera vendue à l'Europe. L'Union Européenne est indifférente entre produire 50 chemises ou 10 voitures. En échange de la production en chemises d'un travailleur chinois, elle est donc prête à offrir un maximum de 40 voitures, quatre fois plus que ce que la Chine produirait en autarcie. En achetant ces 200 chemises à la Chine, l'Europe économise quant-à-elle une quantité de travail qui lui permet de produire 80 voitures. Dans ce cadre, l'échange est donc mutuellement avantageux car il permet aux deux pays d'utiliser sa quantité de travail de manière plus efficace et de consommer plus à l'équilibre.

Dans l'exemple (b), le gain à l'échange est moins évident. En effet, dans cette configuration, la Chine dispose d'un avantage absolu dans la production des chemises comme des voitures. Si l'Europe peut évidemment avoir intérêt à acheter ces biens plutôt que de les produire elle-même, le caractère mutuellement avantageux d'un tel échange n'est pas immédiat. Pour le constater, il faut regarder les chiffres du tableau 2.1.b en termes relatifs et pas en termes absolus. Le coût d'opportunité en termes de voitures de la production d'une chemise supplémentaire est de $1/20$ en Chine et de $1/5$ en Europe². En termes de voitures, il est donc relativement moins coûteux d'augmenter la production de chemises en Chine qu'en Europe. A l'inverse, le coût d'opportunité en termes de chemises de la production d'une voiture supplémentaire est de 20 en Chine et 5 en Europe. En termes de chemises, il est donc relativement moins coûteux d'augmenter la production de voitures en

2. On appelle coût d'opportunité le prix en termes d'unités de biens auxquelles il faut renoncer pour produire une unité supplémentaire d'un autre bien.

Europe qu'en Chine. L'Europe a un *avantage comparatif* dans la production de voitures tandis que la Chine a son avantage comparatif dans la production de chemises. Ces écarts de coûts d'opportunité impliquent qu'il existe un réarrangement mutuellement bénéfique de la production mondiale. En permettant à chaque pays de se spécialiser dans la production de son avantage comparatif, on augmente la production mondiale. Avec plus de biens produits, il existe des répartitions internationales augmentant la consommation d'au moins un pays, sans diminuer celle de l'autre. L'échange est donc Pareto-optimal.

Pour le montrer, il est utile de partir d'une situation d'autarcie dans laquelle chaque pays produit l'intégralité de sa consommation de voitures et de chemises. Supposons que chaque pays dispose d'une quantité L de travail chaque mois. Avec ces travailleurs, la Chine peut produire C chemises et V voitures dans l'ensemble défini par : $L \geq \frac{C}{400} + \frac{V}{20}$, i.e. en saturant la contrainte de capacité : $V = 20L - \frac{C}{20}$. Avec la même quantité de travailleurs, l'Europe peut produire C chemises et V voitures dans l'ensemble défini par : $L \geq \frac{C}{50} + \frac{V}{10}$, i.e. en saturant la contrainte de capacité : $V = 10L - \frac{C}{5}$. Ces "frontières des possibilités de production" sont illustrées sur le graphique 2.1, en gras³. La Chine dispose bien d'un avantage absolu dans les deux biens : au maximum, elle peut produire $20L$ voitures et $400L$ chemises quand l'Europe ne peut produire, au maximum, que $50L$ chemises et $10L$ voitures.

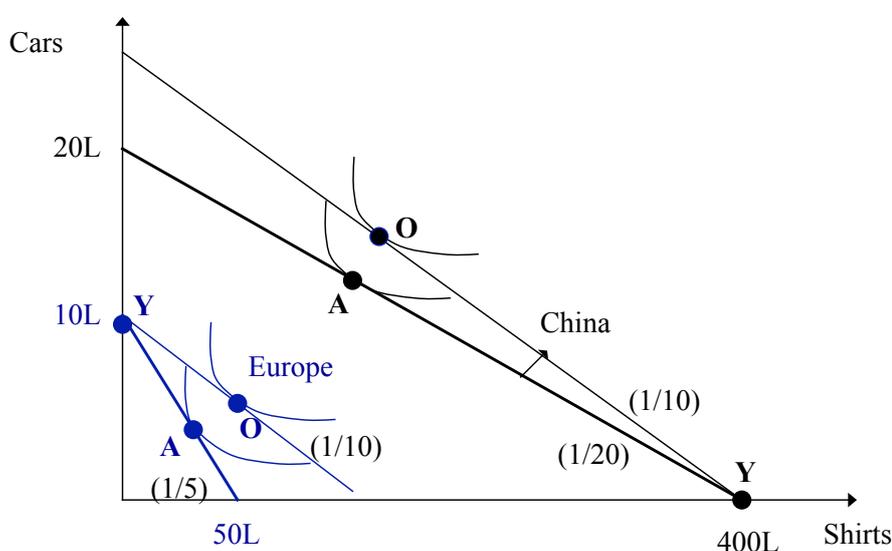
A l'optimum autarcique, chaque pays choisit la combinaison optimale de chemises et de voitures produits. Celle-ci correspond aussi au panier de biens consommé, qui maximise l'utilité des consommateurs sous la contrainte de faisabilité imposée par la frontière des possibilités de production. Le choix optimal correspond aux points A sur le graphique 2.1. En ces points, la frontière des possibilités de production est tangente à la courbe d'indifférence des consommateurs⁴. A l'équilibre concurrentiel, la pente de la courbe d'in-

3. La frontière des possibilités de production délimite l'ensemble des paniers de biens qu'un pays peut produire avec sa dotation en facteurs. Etant donnée une quantité fixe de travail, les producteurs d'un pays font face à un arbitrage : produire plus de voitures implique de réduire sa production de chemises et inversement. Cet arbitrage est représenté par la frontière des possibilités de production, qui est décroissante dans le plan (C, V) . Dans le cas d'un modèle à un facteur de production, la frontière des possibilités de production est une droite dont la pente (en valeur absolue) est égale au coût d'opportunité des chemises en termes de voitures.

4. On appelle courbe d'indifférence l'ensemble des paniers de consommation qui procurent le même niveau d'utilité au consommateur. Dans le plan $(C; V)$, la courbe d'indifférence est convexe et d'autant plus élevée que le niveau d'utilité atteint par la consommation est haut. La maximisation de l'utilité sous la contrainte budgétaire consiste donc à choisir un point de consommation sur la courbe d'indifférence la plus élevée possible, tout en respectant la contrainte.

différence au point de consommation optimale est égale au prix relatif des chemises, en termes de voiture, $1/5$ en Europe et $1/20$ en Chine. Pour ces prix relatifs, les producteurs de chaque pays sont indifférents entre produire des voitures et des chemises, ce qui correspond donc bien à un équilibre.

FIGURE 2.1 – L'allocation optimale des ressources, en autarcie et dans l'économie intégrée



Supposons à présent que les deux pays s'ouvrent au commerce international et que le prix relatif des chemises en termes de voiture s'établit sur les marchés internationaux à $1/10$. Etant donné ce prix relatif, les deux pays ont intérêt à se spécialiser complètement dans la production des biens pour lesquels ils disposent d'un avantage comparatif. La Chine se spécialise complètement dans la production de chemises, produit $400L$ chemises dont une partie est vendue sur les marchés internationaux au prix de $1/10$, i.e. un prix plus élevé que le coût d'opportunité de $1/20$ auquel le pays fait face lorsqu'il réduit sa production de voitures pour produire une chemise supplémentaire à exporter. Parallèlement, l'Europe se spécialise complètement dans la production de voitures et produit $10L$ voitures dont une partie est vendue à la Chine au prix relatif de 10 chemises, plus élevé que le coût d'opportunité domestique de cette production supplémentaire de voiture. Ces allocations productives sont représentées par les points Y sur le graphique 2.1. Ces allocations respectent la contrainte imposée par la frontière des possibilités de

production, qui est évidemment inchangée.

Si les possibilités de production sont inchangées, les possibilités de consommation sont maintenant élargies par rapport au cas autarcique. En produisant $400L$ chemises, la Chine peut maintenant consommer un maximum de $400L$ chemises ou $40L$ ($400L/10$) voitures ainsi que toutes les combinaisons des deux biens qui se situent sur la droite entre ces deux points. En produisant $10L$ voitures, l'Europe peut maintenant consommer $10L$ voitures ou $100L$ chemises, ainsi que toutes les combinaisons entre ces deux points. Ces deux ensembles de possibilités de consommation, en économie ouverte, sont délimités sur le graphe 2.1 par les droites en trait fin. A l'optimum, le consommateur représentatif de chaque pays choisit la combinaison qui maximise son utilité sous cette contrainte budgétaire. Ces choix sont matérialisés par les points notés O sur le graphe, quand la courbe d'indifférence est tangente à la contrainte budgétaire du consommateur représentatif. La pente de la courbe d'indifférence en ce point est égale au prix relatif des chemises sur les marchés internationaux, $1/10$ dans notre exemple.

La comparaison des points de consommation d'autarcie (A) et d'économie ouverte (O) permet d'appréhender l'origine des gains à l'échange. En s'ouvrant au commerce international, chaque pays étend l'ensemble de ses possibilités de consommation, puisqu'il n'est plus contraint de consommer uniquement ce qu'il peut produire. La spécialisation conduit à une utilisation plus efficace des facteurs de production (le travail ici). En se spécialisant complètement dans le bien pour lequel il dispose d'un avantage comparatif, chaque pays maximise l'efficacité de l'utilisation de ses unités de travail. En vendant une partie de sa production sur les marchés internationaux, il peut ensuite consommer un panier de biens en dehors de l'espace des possibilités de production. En ce sens, on peut assimiler le commerce à une méthode indirecte de production, qui est plus efficace que la méthode choisie à l'équilibre en autarcie. Cet exemple montre donc comment l'échange permet à chaque pays de consommer plus d'au moins un des biens et d'augmenter l'utilité du consommateur représentatif. L'échange est Pareto-optimal.

Résumé :

- On dit qu'un pays a un avantage comparatif dans la production d'un bien si le coût d'opportunité à produire ce bien est plus faible que dans un autre pays.
- L'échange entre deux pays est Pareto-optimal si chaque pays se spécialise et exporte le bien pour lequel il a un avantage comparatif

Jusqu'ici, nous avons regardé le problème de manière *ex-post*, sous l'hypothèse d'un prix mondial conduisant chaque pays à préférer acheter d'un bien plutôt que de le produire domestiquement. Nous allons maintenant voir comment, dans un modèle ricardien, les mécanismes de marché conduisent

TABLE 2.2 – Avantage absolu et avantage comparatif dans un modèle à deux biens (X et Y)

		Avantage absolu pour X	
		Domestique ($a_X < a_X^*$)	Etranger ($a_X > a_X^*$)
Avantage absolu pour Y	Domestique ($a_Y < a_Y^*$)	$?$ $\left(\frac{a_X \geq a_X^*}{a_Y < a_Y^*}\right)$	Dom : Y /Etranger : X $\left(\frac{a_X > a_X^*}{a_Y < a_Y^*}\right)$
	Etranger ($a_Y > a_Y^*$)	Dom : X /Etranger : Y $\left(\frac{a_X < a_X^*}{a_Y > a_Y^*}\right)$	$?$ $\left(\frac{a_X \geq a_X^*}{a_Y < a_Y^*}\right)$

a_i (a_i^*) = nombre d'unités de travail nécessaire à la production d'une unité du bien i dans le pays domestique (étranger). La table donne la structure des avantages comparatifs, en fonction de la nature des avantages absolus.

effectivement à une spécialisation des pays ouverts au commerce international.

2.2 Un facteur de production : Le modèle ricardien

2.2.1 Hypothèses

Le modèle ricardien de base est un modèle à deux pays, deux biens et un facteur de production. Dans ce qui suit, les variables relatives à l'économie étrangère se distinguent des équivalents pour l'économie domestique par un astérisque (*). Ces pays partagent la même taille, les mêmes préférences, mais se distinguent par leurs fonctions de production. L'origine de l'avantage comparatif est ici liée à des écarts technologiques entre pays.

L'environnement est parfaitement concurrentiel, ce qui implique que, à l'équilibre, les prix sont égaux au coût marginal de production. Le seul facteur de production, le travail est disponible en quantité L dans chaque pays. Il est parfaitement mobile entre secteurs. La mobilité conduit à la détermination d'un salaire unique pour les travailleurs employés dans les deux secteurs. En revanche, le travail est immobile internationalement et le taux de salaire peut donc être différent dans l'économie domestique et à l'étranger ($w \neq w^*$).

Les fonctions de production sont linéaires en le travail. Dans chaque pays et chaque secteur (X et Y), la technologie se caractérise par un coefficient technique qui mesure le nombre d'unités de travail nécessaire pour produire une unité du bien : a_X , a_X^* , a_Y , a_Y^* . On suppose que les coefficients technologiques satisfont la condition suivante :

$$\frac{a_X}{a_Y} > \frac{a_X^*}{a_Y^*}$$

i.e. que l'économie étrangère a un avantage comparatif dans la production du bien X (il est relativement plus coûteux en termes de production de bien Y de produire une unité supplémentaire de bien X dans le pays domestique qu'à l'étranger \Leftrightarrow le coût d'opportunité de la production de X en termes de bien Y est plus élevé dans l'économie domestique qu'à l'étranger).

Etant donnée cette condition technologique, l'économie domestique est prête à se spécialiser dans la production de bien Y et à importer du bien X dès lors que le prix relatif à laquelle elle fait face sur les marchés internationaux est plus faible que le coût d'opportunité lié à la production domestique du bien, i.e. si :

$$\frac{p_X}{p_Y} < \frac{a_X}{a_Y}$$

où p_X/p_Y est le prix relatif du bien X sur les marchés internationaux. De même, l'économie étrangère est prête à se spécialiser dans la production de bien X et à importer du bien Y dès lors que :

$$\frac{p_Y}{p_X} < \frac{a_Y^*}{a_X^*}$$

Les intérêts mutuels de ces deux économies sont donc cohérents puisque $\frac{a_X}{a_Y} > \frac{a_X^*}{a_Y^*}$. Il existe un prix relatif p_X/p_Y compris entre ces deux limites tel que les pays échangent à l'équilibre.

2.2.2 L'équilibre autarcique

En concurrence parfaite, les prix sectoriels s'établissent au niveau du coût marginal de production (le nombre d'unités de travail nécessaires pour produire une unité du bien multiplié par le salaire unitaire) :

$$\begin{aligned} p_X^a &= a_X w^a & p_Y^a &= a_Y w^a & \Rightarrow & \frac{p_X^a}{p_Y^a} = \frac{a_X}{a_Y} \\ p_X^{*a} &= a_X^* w^{*a} & p_Y^{*a} &= a_Y^* w^{*a} & \Rightarrow & \frac{p_X^{*a}}{p_Y^{*a}} = \frac{a_X^*}{a_Y^*} \end{aligned}$$

L'indice a est utilisé ici pour faire référence à l'équilibre autarcique. Le prix relatif du bien X à l'équilibre autarcique est égal au coût d'opportunité

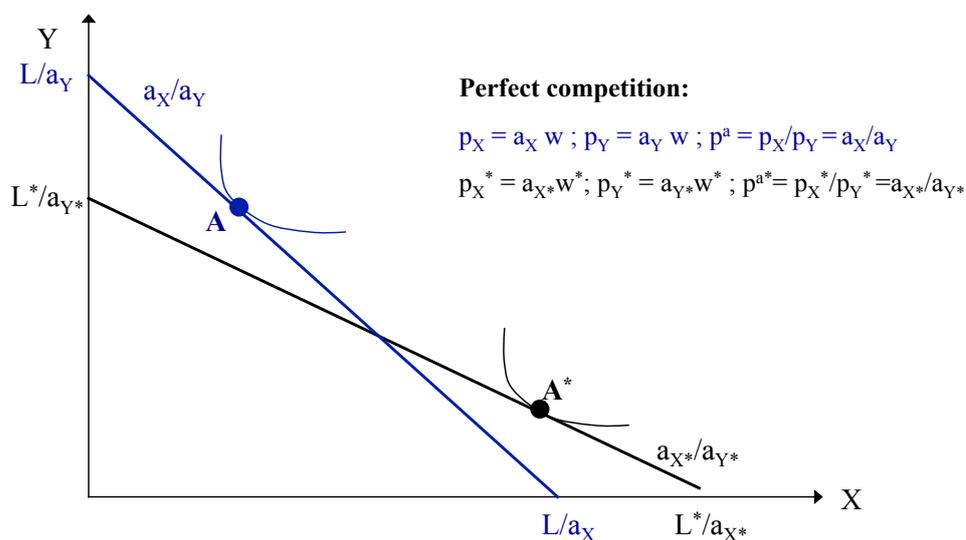
de X en termes de bien Y . A ce prix relatif, les producteurs de chaque pays sont indifférents entre produire X et Y , ce qui caractérise une situation d'équilibre. Le prix relatif du bien X est plus élevé à l'équilibre autarcique dans le pays domestique qu'à l'étranger. C'est une conséquence de l'avantage comparatif de l'économie étrangère dans la production du bien X . C'est cette différence de prix à l'équilibre qui ouvre la porte à un commerce mutuellement avantageux.

Avec des fonctions de production linéaire en un unique facteur, la frontière des possibilités de production s'écrit simplement (respectivement pour l'économie domestique et à l'étranger) :

$$\begin{aligned} L &= a_X X + a_Y Y \\ L^* &= a_X^* X^* + a_Y^* Y^* \end{aligned}$$

Elle est illustrée sur le graphique 2.2, en bleu pour l'économie domestique et en noir pour l'étranger. La pente de la frontière des possibilités de production est égale en valeur absolue au coût relatif de production du bien X . L'avantage comparatif du pays étranger dans la production du bien X se traduit par une frontière des possibilités de production moins pentue dans le plan (X, Y) .

FIGURE 2.2 – L'équilibre autarcique



En autarcie, la frontière des possibilités de production coïncide avec la

frontière des possibilités de consommation (ou contrainte budgétaire)⁵. Le consommateur représentatif choisit le panier de biens qui maximise son utilité étant donnée cette contrainte budgétaire. A l'équilibre, le consommateur représentatif domestique (resp. étranger) choisit le panier noté A (resp. A^*) sur le graphique 2.2. En ces points, la courbe d'indifférence est tangente avec la contrainte budgétaire (i.e. est de pente a_X/a_Y ou a_X^*/a_Y^* respectivement pour le consommateur domestique et le consommateur étranger).

2.2.3 L'équilibre en économie ouverte

A l'équilibre autarcique, le prix relatif du bien X est plus élevé dans le pays domestique qu'à l'étranger. Lorsque les pays s'ouvrent au commerce international, il y a donc une possibilité d'arbitrage pour les consommateurs domestiques, qui peuvent acheter leur consommation en biens X pour un prix relativement plus faible à l'étranger⁶. Inversement, les consommateurs étrangers peuvent trouver leur consommation de bien Y pour un prix relativement plus faible dans le pays domestique. Ce sont ces écarts de prix, résultant des écarts de coûts d'opportunité, qui vont conduire à la spécialisation des pays. Dans chaque pays, la demande relative pour le bien correspondant à son avantage comparatif augmente. En effet, la demande absolue adressée aux producteurs de ce bien augmente du fait de la demande étrangère supplémentaire. Dans le même temps, la demande domestique pour l'autre bien se reporte sur l'étranger. Les facteurs productifs (i.e. le travail) se réallouent entre secteurs pour faire face à cette modification de la demande relative. Il y a donc bien spécialisation dans la production du bien pour lequel le pays dispose d'un avantage comparatif. La spécialisation conduit quant-à-elle à une convergence des prix relatifs, les comportements d'arbitrage ne cessant que lorsque les prix relatifs sont égalisés internationalement. Cette convergence caractérise l'équilibre en économie ouverte.

En économie ouverte, chaque pays se spécialise dans son avantage comparatif si le prix relatif de ce bien est plus élevé que le coût d'opportunité lié à la production supplémentaire pour l'exportation. Tout prix relatif tel que :

$$\frac{a_X^*}{a_Y^*} \leq \frac{p_X}{p_Y} \leq \frac{a_X}{a_Y}$$

5. La frontière des possibilités de production délimite l'ensemble des paniers de biens qu'une économie peut produire étant données les ressources disponibles. La contrainte budgétaire délimite l'ensemble des paniers de consommation qu'un consommateur représentatif peut acheter étant donné son revenu.

6. Ce raisonnement n'est valable que dans une situation où le commerce est parfaitement libre et où il n'y a pas de coût de transport des biens d'une économie à l'autre. Le cas des coûts à l'échange est décrit en Section 2.2.5.

conduit donc à un équilibre avec spécialisation des pays dans leur avantage comparatif. Puisque $\frac{p_X}{p_Y} \geq \frac{p_X^a}{p_Y^a}$, le pays étranger gagne à produire plus de bien X tout en important le bien Y . Puisque $\frac{p_Y}{p_X} \geq \frac{p_Y^a}{p_X^a}$, le pays domestique gagne à se spécialiser dans la production du bien Y et à importer X .

Selon le prix relatif vers lequel l'économie nationale converge, deux situations sont à envisager :

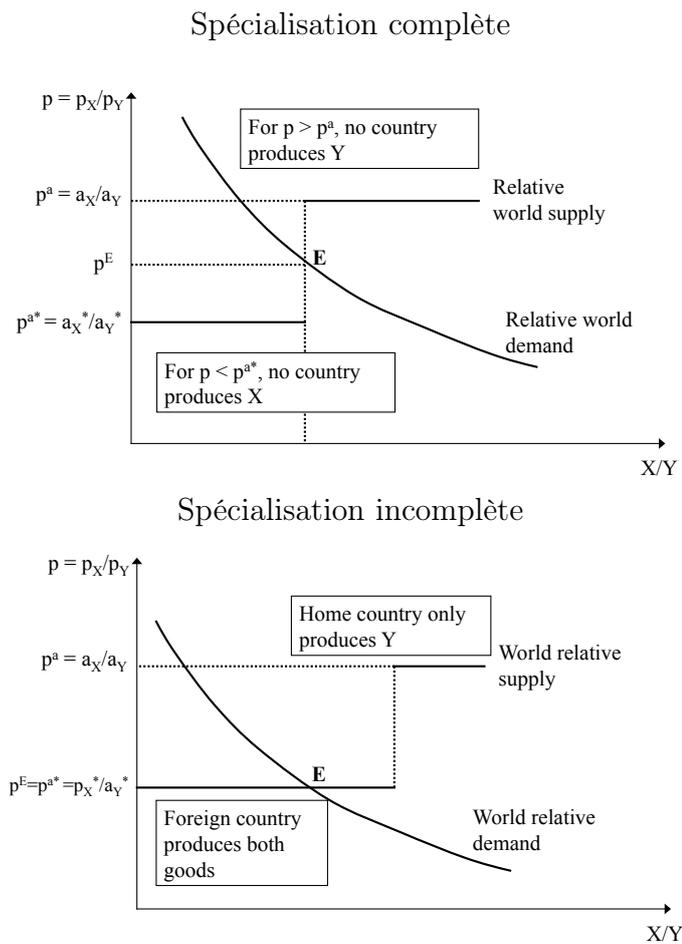
- Si le prix d'équilibre est tel que $\frac{a_X^*}{a_Y^*} < \frac{p_X}{p_Y} < \frac{a_X}{a_Y}$, l'équilibre est caractérisé par une spécialisation complète de chaque pays dans la production de biens pour lesquels il dispose d'un avantage comparatif.
- Si le prix d'équilibre est tel que $\frac{p_X}{p_Y} = \frac{a_X^*}{a_Y^*}$ ou $\frac{p_X}{p_Y} = \frac{a_X}{a_Y}$, l'équilibre est caractérisé par une spécialisation incomplète d'un pays. Si $\frac{p_X}{p_Y} = \frac{a_X^*}{a_Y^*}$, le pays domestique est complètement spécialisé mais le pays étranger continue à produire des deux biens. Si $\frac{p_X}{p_Y} = \frac{a_X}{a_Y}$, le pays étranger est complètement spécialisé mais le pays domestique continue à produire des deux biens.

Pour comprendre l'apparition de ces deux régimes, il est utile d'étudier l'équilibre du marché des biens dans une économie ouverte. Celui-ci est illustré sur la Figure 2.3. Sur chaque graphique sont représentées l'offre et la demande relative de bien X au niveau mondial, en fonction du prix relatif qui s'établit sur les marchés internationaux. Le prix relatif d'équilibre s'établit au point d'intersection des deux courbes, lorsque la demande relative est juste couverte par l'offre relative.

Sur la Figure 2.3, la demande relative est décroissante du prix relatif du fait d'effets de substitution. Plus le prix relatif de X est élevé, plus les consommateurs des deux pays reportent leur consommation sur le bien Y , ce qui implique une diminution de la demande relative de bien X . L'offre relative est croissante, avec plusieurs discontinuités au niveau des prix relatifs d'autarcie. Lorsque le prix d'équilibre est inférieur à a_X^*/a_Y^* , l'offre mondiale relative est nulle car les producteurs des deux pays préfèrent se spécialiser dans la production du bien dont le prix relatif est supérieur au coût d'opportunité (le bien Y ici). Pour $p_X/p_Y = a_X^*/a_Y^*$, les entreprises domestiques continuent à se spécialiser dans la production du bien Y . En revanche, les producteurs étrangers sont indifférents entre produire du bien X ou du bien Y . Ils sont donc prêts à satisfaire la demande relative exprimée sur les marchés internationaux, la courbe d'offre est horizontale. A l'autre extrémité, si $p_X/p_Y > a_X/a_Y$, l'offre relative de bien X tend vers l'infini (puisque aucun des pays ne souhaite produire le bien Y , dont le prix relatif est inférieur au coût d'opportunité). Si $p_X/p_Y = a_X/a_Y$, le pays étranger produit autant que possible de bien X et l'économie domestique est indifférente entre la production

des deux biens. Finalement, il n'y a que quand $a_X^*/a_Y^* < p_X/p_Y < a_X/a_Y$ que l'économie est complètement spécialisée. L'économie domestique ne produit que du bien Y (en quantités L/a_Y), l'économie étrangère ne produit que du bien X (en quantités L/a_X^*) et donc l'offre relative est $\frac{L/a_X^*}{L/a_Y}$.

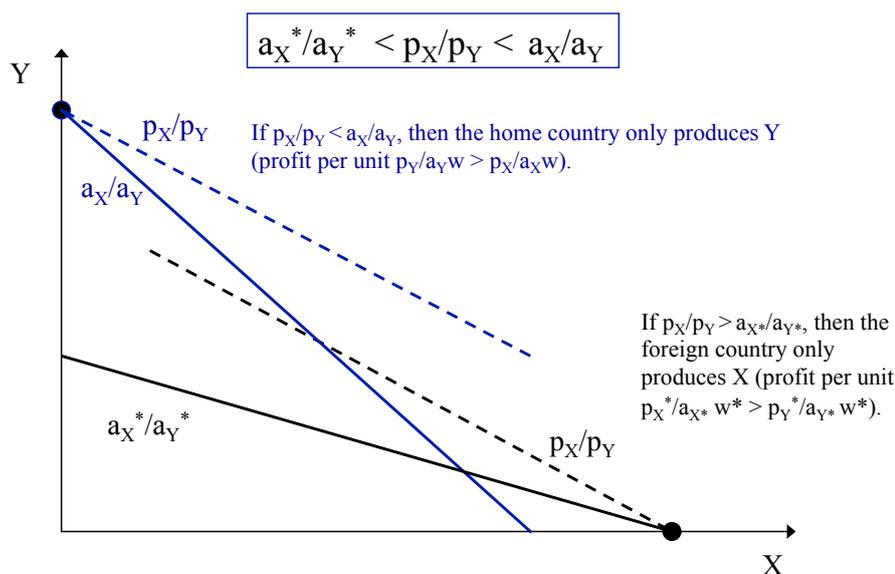
FIGURE 2.3 – L'équilibre en économie ouverte



L'équilibre s'établit au point d'intersection entre la courbe d'offre et la courbe de demande relative en biens X . Sur le graphique du haut de la Figure 2.3, cette intersection se situe à un point où le prix est strictement compris entre les prix relatifs autarciques. Chaque pays est complètement spécialisé dans la production de son avantage comparatif. A l'inverse sur le graphique du bas, la courbe de demande coupe la courbe d'offre dans sa partie horizontale. Il y a spécialisation incomplète. Le pays domestique ne

peut produire toute la demande relative de biens X et le produit étranger produit donc à la fois du bien X et du bien Y .

FIGURE 2.4 – Gains à l'échange en équilibre de spécialisation complète



Considérons à présent l'équilibre avec spécialisation complète. Ici aussi, on peut comprendre l'origine des gains à l'échange en étudiant la manière dont l'ouverture au commerce affecte les possibilités de consommation de chaque pays. Cette dimension est illustrée sur la Figure 2.4. Comme dans la section précédente, le graphique compare, pour chaque pays, l'ensemble des possibilités de production et de consommation, en autarcie et en économie ouverte. En autarcie, la frontière des possibilités de production et la contrainte budgétaire sont confondues. En économie ouverte, la contrainte budgétaire pivote pour atteindre une pente égale au nouveau prix relatif d'équilibre $a_X^*/a_Y^* < p_X/p_Y < a_X/a_Y$. L'économie domestique produit L/a_Y unités de bien Y et consomme un panier de biens situé sur la nouvelle contrainte budgétaire (la ligne en pointillés bleus sur la Figure 2.4). Ce panier de biens se situe sur une courbe d'indifférence plus élevée que n'importe quelle courbe d'indifférence atteinte en autarcie. L'échange améliore donc le bien-être du consommateur représentatif en augmentant les possibilités de consommation. Au lieu de produire sa consommation de bien X , l'économie domestique économise a_X unités de travail à chaque unité de bien X importée. Avec ce travail, elle produit a_X/a_Y unités de bien Y qu'elle vend en échange de $\frac{p_Y}{p_X} \frac{a_X}{a_Y} > 1$

unités de bien X .

2.2.4 La détermination des salaires

Les gains à l'échange discutés dans la section précédente ne dépendent pas du niveau d'équilibre des salaires dans les deux économies. Cette propriété est liée à l'hypothèse de mobilité intersectorielle du travail. Dans la mesure où les salaires sont homogènes entre les secteurs, au sein d'une économie, ils ne modifient pas la structure des avantages comparatifs, qui seule détermine les gains à l'échange dans un modèle ricardien. La question des salaires est cependant importante puisqu'un grand nombre de débats (et d'erreurs) concernant les questions d'économie internationale porte sur la question des différentiels de salaires et de son évolution quand un pays s'ouvre au commerce international.

A l'équilibre autarcique, on a par définition :

$$w^a = \frac{p_X^a}{a_X} = \frac{p_Y^a}{a_Y} \quad \text{et} \quad w^{a*} = \frac{p_X^{a*}}{a_X^*} = \frac{p_Y^{a*}}{a_Y^*}$$

Le salaire d'équilibre est déterminé de façon à équilibrer l'offre et la demande de travail. L'offre est exogène et égale à L . La demande est la somme des demandes émanant des deux secteurs.

A l'équilibre de spécialisation complète, on a :

$$w = \frac{p_Y}{a_Y} \quad \text{et} \quad w^* = \frac{p_X}{a_X^*}$$

Le salaire d'équilibre est toujours défini à l'équilibre du marché du travail. Cependant, l'offre n'émane plus que d'un secteur, celui correspondant à l'avantage comparatif du pays.

On peut comparer ces deux équilibres pour déterminer la manière dont le commerce affecte les salaires des deux économies. La convergence des prix induite par l'ouverture au commerce implique que, dans chaque pays, le prix du bien dans lequel le pays se spécialise augmente. Par rapport à la situation d'autarcie, le salaire augmente donc dans les deux pays : $w > w^a$ car $p_Y > p_Y^a$, $w^* > w^{a*}$ car $p_X^* > p_X^{a*}$. Cette propriété contredit l'idée communément admise selon laquelle l'ouverture au commerce conduirait à du *dumping* social et à une pression à la baisse sur les salaires dans les pays riches. Dans le modèle ricardien, ce n'est pas le cas. Au contraire, les salaires augmentent dans les deux pays. Le travail étant utilisé de manière plus efficace, sa productivité marginale moyenne augmente, ce qui se répercute sur les salaires dans un cadre concurrentiel.

En termes de salaire relatif, on a :

$$\frac{w}{w^*} = \frac{p_Y a_X^*}{p_X a_Y}$$

En utilisant la propriété sur le prix relatif, à l'équilibre de spécialisation complète ($a_Y/a_X < p_Y/p_X < a_Y^*/a_X^*$), on trouve :

$$\frac{a_X^*}{a_X} < \frac{w}{w^*} < \frac{a_Y^*}{a_Y}$$

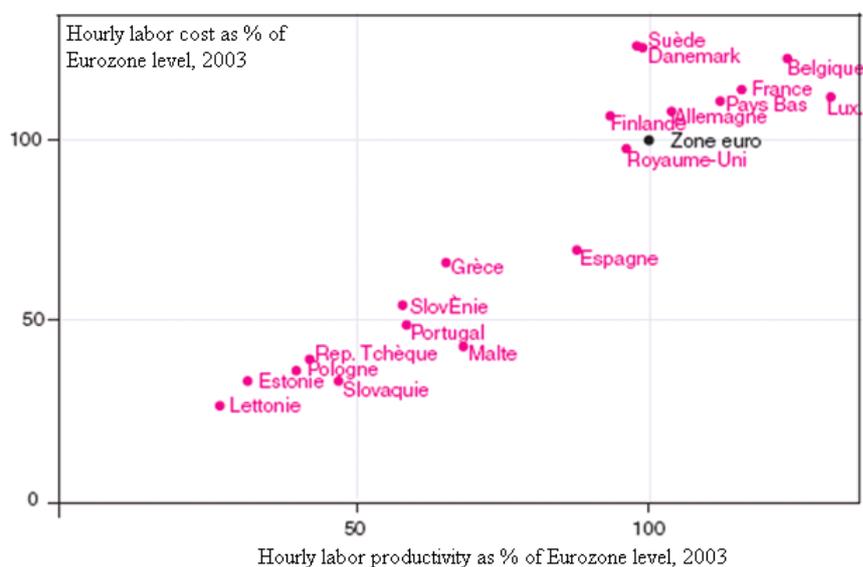
Le salaire relatif, à l'équilibre de spécialisation incomplète, s'établit dans un intervalle dont les extrémités sont déterminées par les avantages *absolus* des deux économies. En particulier, si une des deux économies a un avantage absolu dans la production des deux biens (par exemple, si $\frac{a_X^*}{a_X} > 1$ et $\frac{a_Y^*}{a_Y} > 1$) alors son salaire s'établit à un niveau supérieur à celui de l'autre économie. C'est cette différence salariale qui fait que l'échange est mutuellement avantageux. L'économie qui a un désavantage absolu dans la production des deux biens compense ce manque de compétitivité par un salaire *relatif* plus faible, qui lui permet de vendre son avantage comparatif à un prix compétitif.

Cette prédiction du modèle ricardien est confrontée aux données dans le graphique 2.5. La figure illustre la corrélation entre les salaires nominaux et le niveau de la productivité du travail, pour les pays de l'Union Européenne. La corrélation est positive ce qui signifie que les salaires européens les plus faibles sont associés à une productivité relativement moindre du travail. Cette corrélation contredit l'idée selon laquelle l'ouverture au commerce conduit à du "*dumping social*", idée très largement reprise dans les milieux politiques et médiatiques (voir par exemple les débats autour de la directive Bolkestein en 2006). Dans l'exemple européen, l'intégration des pays d'Europe Centrale et de l'Est à l'union douanière européenne a effectivement ouvert le marché européen à des pays au coût du travail relativement faible. Cette faiblesse est cependant compensée par une moindre productivité relative et ne conduit donc pas à une pression à la baisse sur les salaires d'Europe de l'Ouest.

2.2.5 Extensions et tests empiriques

Le modèle ricardien offre donc une explication des flux internationaux de biens et services en termes de spécialisation sur la base d'avantages comparatifs technologiques. Il prédit notamment que les exportations nettes d'un pays concernent des produits pour lesquels le pays dispose d'un avantage comparatif. Pour pouvoir confronter cette prédiction aux données, il est cependant utile de lever certaines hypothèses du modèle, qui apparaissent particulièrement contraignantes.

FIGURE 2.5 – Salaire et productivité du travail



Source : Eurostat. Ce graphique compare, pour les différents pays de l'Union européenne, la productivité horaire du travail (axe des abscisses, exprimée en pourcentage de la productivité du travail de la zone euro) et le coût unitaire du travail (axe des ordonnées, également exprimé en pourcentage de la moyenne de l'UME).

Modèle multisectoriel : La première hypothèse du modèle qui en fait un cadre d'analyse difficile à tester empiriquement tient à sa structure extrêmement simple à deux pays, deux biens et un facteur de production. Dans la section suivante, nous verrons comment élargir le raisonnement ricardien à un cadre multifactoriel. Les hypothèses de deux pays et deux biens sont beaucoup plus difficiles à lever. En particulier, le modèle s'étend difficilement à plus de deux pays⁷. L'hypothèse de deux secteurs est moins contraignante. En particulier, il est possible de travailler sur un modèle multisectoriel en utilisant les implications qu'a le modèle sur la détermination du salaire relatif. C'est ce qu'ont montré notamment Dornbusch, Fischer & Samuelson (1977).

Considérons un modèle à deux pays et n biens. Pour chaque bien, on peut calculer la productivité relative du pays domestique, et ordonner les n biens

7. Une version "moderne" du modèle ricardien a été proposée par Eaton & Kortum (2002). Celle-ci applique la théorie des avantages comparatifs à un modèle multi-pays avec un continuum de biens. Dans ce modèle, les flux de commerce sont déterminés par la productivité relative de différents pays à produire les différents biens du continuum, que les auteurs modélisent comme une variable aléatoire.

en ordre croissant de cette productivité relative :

$$\frac{a_1^*}{a_1} < \frac{a_2^*}{a_2} < \dots < \frac{a_n^*}{a_n}$$

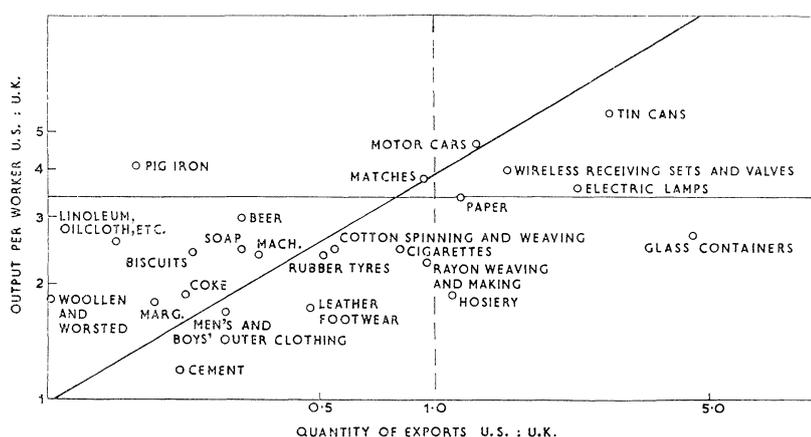
Ici, les n biens sont indexés de façon à ce que le bien 1 corresponde au désavantage comparatif le plus fort de l'économie domestique tandis que le bien n est au contraire un bien qui y est produit avec une technologie très efficiente, en termes relatifs.

A l'équilibre en économie ouverte, les biens sont produits dans le pays où c'est le moins coûteux. Un bien i est produit uniquement dans l'économie domestique si et seulement si :

$$wa_i < w^*a_i^* \Leftrightarrow \frac{a_i^*}{a_i} > \frac{w}{w^*}$$

En localisant w/w^* dans la chaîne constituée par les productivités relatives de n biens, on détermine donc quels sont les biens qui, en théorie, sont produits et exportés par l'économie domestique et par le pays étranger. En particulier, tous les biens i tels que $a_i^*/a_i > w/w^*$ sont produits et exportés par l'économie domestique et tous les biens i tels que $a_i^*/a_i < w/w^*$ sont produits par l'économie étrangère. Pour ces biens en effet, le désavantage relatif de l'économie étrangère en termes de salaire est plus que compensé par un avantage en termes de productivité, ce qui rend le pays compétitif dans ce secteur.

FIGURE 2.6 – Avantages comparatifs et flux de commerce



Source : Stern (1962)

Cette prédiction est testée sur la figure 2.6, tirée de Stern (1962). Ce test s'inspire des travaux empiriques de MacDougall (1951), Stern (1962) et

Balassa (1963) qui “testent” l’intuition du modèle ricardien. Supposons qu’il n’existe que deux pays au monde, les Etats-Unis et le Royaume-Uni. Dans un cadre multisectoriel, on doit s’attendre à ce que chaque pays exporte les biens pour lesquels sa productivité relative du travail est supérieure au salaire relatif. Sur le graphique 2.6, cela se traduit par une corrélation positive entre le ratio des productions par travailleur aux Etats-Unis et au Royaume-Uni (axe des ordonnées) et les exportations relatives des Etats-Unis et du Royaume-Uni (axe des abscisses). Notons que la corrélation valide également l’argument selon lequel ce sont les avantages comparatifs et pas les avantages absolus qui déterminent la structure du commerce international. En effet, à l’époque où le test empirique est élaboré (avec des données relatives à 1950), la productivité du travail en Grande-Bretagne est inférieure à la productivité du travail américaine dans quasiment tous les secteurs (i.e. la productivité relative mesurée sur l’axe des ordonnées est toujours supérieure à 1). Malgré ce désavantage absolu, la Grande-Bretagne exporte vers les Etats-Unis les biens pour lesquels son désavantage comparatif est le moins marqué.

L’intuition ricardienne est donc validée par des données relatives au commerce des Etats-Unis et de la Grande-Bretagne, dans la période post-2ème guerre mondiale. L’intuition reste très largement validée par les données plus récentes. Par exemple, la table 2.3 reporte les résultats d’une estimation proposée par Ruoan & Manying (2002) et citée par Krugman et al. (2012). Dans cette étude, les auteurs testent l’intuition ricardienne sur la structure du commerce de la Chine et de l’Allemagne, qui sont aujourd’hui les deux plus gros exportateurs mondiaux. La table compare la productivité relative de la Chine et sa production relative, en comparaison de l’Allemagne. Pour le secteur manufacturier dans son ensemble, et dans les données de 1995, la Chine souffre d’un désavantage de productivité fort : elle produit avec une productivité par travailleur égale à 1/20ème de celle de l’Allemagne. Ce désavantage absolu est cependant moins fort dans certains secteurs que dans d’autres. En particulier, dans le secteur de l’habillement, le désavantage de productivité n’est que de 1/5. Le secteur de l’habillement n’est pas choisi au hasard ici. Il s’agit d’un des plus gros avantages comparatifs de la Chine sur les marchés internationaux. Cet avantage comparatif explique que la production par la Chine de vêtements est 800% plus élevée que celle de l’Allemagne. Ce résultat est conforme aux prédictions ricardiennes : la Chine profite de son avantage comparatif dans le secteur de l’habillement pour se spécialiser et exporter sur les marchés internationaux, et ce malgré une productivité dans ce secteur qui reste faible.

Coûts à l’échange et biens non échangés : Une autre hypothèse du modèle ricardien qui apparaît très contrefactuelle concerne l’absence de

TABLE 2.3 – Test du modèle ricardien, Chine versus Allemagne (1995)

	Production par tra- vailleur de la Chine (% de l'Allemagne)	Production totale de la Chine (% de l'Alle- magne)
Secteur manufacturier	5.2	71.6
Secteur de l'habillement	19.7	802.2

Source : Ruoen & Manying (2002), cité par Krugman et al. (2012).

coûts à l'échange international. Cette hypothèse est cruciale car les comportements d'arbitrage engendrés par le commerce international sont au coeur de la convergence internationale des prix. Dans un contexte avec coût à l'échange, l'arbitrage ne se fera que pour un différentiel de prix suffisamment grand pour compenser le coût à l'échange. Etant données les nombreuses évidences empiriques de l'existence de coûts à l'échange international, il apparaît utile de lever cette hypothèse du modèle.

Supposons que l'échange international induit un coût multiplicatif τ , qui s'applique au prix unitaire de vente (un exemple de tels coûts multiplicatifs concerne les tarifs *ad valorem* qui s'appliquent à l'importation de certaines marchandises). En présence de coûts à l'échange, les comportements d'arbitrage sont basés sur la comparaison du prix domestique et du prix étranger, augmenté du coût à l'échange. Pour un bien i dont le prix est relativement plus élevé à l'étranger que dans l'économie domestique ($wa_i < w^*a_i^*$), le consommateur représentatif étranger trouvera l'arbitrage profitable si et seulement si le prix domestique augmenté des coûts de transport reste inférieur au prix étranger : $wa_i(1 + \tau) < w^*a_i^*$. Autrement dit, tous les biens tels que : $wa_i < w^*a_i^* < wa_i(1 + \tau)$ ne sont pas échangés à l'équilibre, bien que le pays domestique dispose d'un avantage comparatif dans leur production. De même, les biens tels que $w^*a_i^* < wa_i < w^*a_i^*(1 + \tau)$ sont produits avec un avantage comparatif à l'étranger mais ne sont pas importés par l'économie domestique du fait des coûts de transport.

Le modèle ricardien augmenté de coûts à l'échange prédit donc l'apparition à l'équilibre d'un certain nombre de biens non échangés. En outre, la part des biens non échangés à l'équilibre augmente avec l'ampleur des coûts à l'échange (puisque les intervalles $[wa_i; wa_i(1 + \tau)]$ et $[w^*a_i^*; w^*a_i^*(1 + \tau)]$ s'élargissent quand τ augmente). Ces prédictions améliorent le réalisme du modèle ricardien en reproduisant une caractéristique essentielle des écono-

mies modernes. Dans les pays développés, environ 50% du PIB est en effet composé de biens non échangés (notamment un certain nombre de services). En outre, la part de la consommation portant sur des biens non échangés a tendance à baisser lorsque les coûts à l'échange international diminuent. Ainsi, les nouvelles technologies de télécommunication et d'information ont réduit de manière dramatique certains coûts à l'échange international, notamment sur les services. L'apparition des *call centers*, par exemple, a mis en évidence la manière dont l'apparition de ces technologies pouvaient réduire la part des biens non échangés (car "non échangeables") dans la consommation agrégée. Si ce type de services représente toujours une part relativement faible des échanges internationaux, de nombreux économistes s'accordent sur le fait que les services pourraient dans le futur dépasser les biens manufacturés dans la valeur totale des échanges mondiaux.

Résumé :

- Dans le modèle de Ricardo, l'ouverture au commerce conduit à une spécialisation des pays dans la production des biens pour lesquels ils disposent d'un avantage technologique relatif.
- La spécialisation est induite par l'existence d'écarts de prix relatifs à l'équilibre autarcique, ces écarts offrant au consommateur des possibilités d'arbitrage international une fois les économies ouvertes au commerce.
- La spécialisation est Pareto-améliorante car elle permet aux pays d'utiliser leurs dotations factorielles de manière plus efficace, donc de consommer plus à l'équilibre.
- Dans le modèle ricardien, elle conduit à une augmentation de la rémunération des travailleurs dans les deux pays, ceux-ci profitant des gains d'efficacité vis-à-vis de l'utilisation des ressources.

Limites : Le modèle ricardien est utile pour dériver un certain nombre d'intuitions concernant l'impact de l'ouverture au commerce sur les économies internationales. Cependant, sa structure extrêmement simplifiée induit aussi un certain nombre de prédictions peu réalistes. D'abord, la spécialisation conduit dans ce modèle à des structures productives extrêmes. A l'équilibre, au moins un des pays ne produit que les biens pour lesquels il dispose d'un avantage comparatif tandis que l'intégralité de sa consommation des autres biens est fournie par l'importation. Dans la section suivante, nous verrons comment l'introduction d'un deuxième facteur de production permet d'obtenir, à l'équilibre, des structures productives moins extrêmes, et donc plus conformes à la réalité des données.

Le modèle ricardien propose une vision quasi-idyllique du commerce international, qui augmente uniformément le bien-être des nations engagées sur les marchés internationaux en améliorant l'efficacité avec laquelle les facteurs

sont utilisés. Cette vision néglige la possibilité d'un impact du commerce sur la distribution des revenus à l'intérieur de l'économie nationale. Une des raisons pour laquelle les pressions protectionnistes sont fortes historiquement est que le commerce international ne profite pas à tous les groupes d'intérêt de manière homogène. En particulier, les détenteurs de facteurs utilisés dans les industries qui font face à la concurrence étrangère souffrent de l'ouverture au commerce. L'introduction d'un deuxième facteur de production permet de mettre en avant ces conflits d'intérêt au sein de l'économie nationale.

2.3 Deux facteurs de production : Le modèle Heckscher-Ohlin-Samuelson (HOS)

⁸Le modèle ricardien est utile car il permet de mettre en avant, dans un cadre analytique simple, le rôle des avantages comparatifs dans la détermination des flux de commerce international. Ce modèle repose cependant sur des hypothèses irréalistes. En particulier, les avantages comparatifs sont entièrement déterminés par des différentiels entre pays et secteurs dans la productivité marginale du travail. D'autres sources d'avantages comparatifs interviennent évidemment, notamment des éléments factoriels. Personne ne doute que les conditions climatiques font du Chili un pays relativement plus productif dans la production de vin que le Canada. De même, les exportations mondiales de pétrole sont extrêmement concentrées géographiquement, sur un petit nombre de pays qui disposent de ressources minérales importantes.

Le modèle Heckscher-Ohlin-Samuelson permet de mettre en évidence le rôle des dotations factorielles comme source d'avantages comparatifs. Lorsque la production nécessite plusieurs facteurs productifs, en proportions différentes selon les secteurs, les pays se spécialisent dans la production des biens qui nécessitent des facteurs dont ils disposent en abondance (le beau temps au Chili, le pétrole au Moyen-Orient). Afin d'isoler complètement l'effet des dotations factorielles, il élimine à l'inverse les avantages comparatifs de type ricardien en supposant que tous les pays partagent les mêmes fonctions de production et les mêmes productivités factorielles. Ce modèle permet de mettre en évidence la manière dont les flux de commerce sont influencés par l'interaction de l'abondance relative en facteurs des pays et de l'intensité relative avec laquelle différents secteurs utilisent différents types de facteurs.

8. Le modèle HOS résulte de la compilation des travaux de trois économistes de la première moitié du 20ème siècle, Heckscher (1919), Ohlin (1933) and Samuelson (1947).

2.3.1 Hypothèses du modèle

Le modèle suppose l'existence de deux pays (Domestique et Etranger), deux biens (X et Y) et deux facteurs de productions (le travail L et le capital K)⁹. Comme dans le modèle ricardien, on suppose une mobilité parfaite des facteurs entre les secteurs, qui implique une égalisation des coûts unitaires¹⁰. En revanche, le travail comme le capital sont supposés immobiles internationalement. Dans ce qui suit, on note w (resp. w^*) le coût unitaire du travail dans l'économie domestique (resp. à l'étranger) et r (resp. r^*) le coût unitaire du capital.

Les pays sont similaires en termes de préférences, de structures productives et de productivités. La seule chose qui les différencie concerne leurs dotations factorielles. Dans ce qui suit, on suppose que le pays domestique est relativement riche en capital, en comparaison du pays étranger, i.e.¹¹ :

$$\frac{K}{L} > \frac{K^*}{L^*}$$

L'environnement est parfaitement concurrentiel. Les profits sont donc nuls et les prix égaux aux coûts marginaux de production. Le prix des facteurs de production est défini à l'équilibre du marché des facteurs :

$$\begin{array}{ll} L_X + L_Y = L & \text{et} \quad K_X + K_Y = K \\ L_X^* + L_Y^* = L^* & \text{et} \quad K_X^* + K_Y^* = K \end{array}$$

où L_i et K_i (resp. L_i^* et K_i^*) sont les besoins en travail et en capital du secteur i dans l'économie domestique (resp. à l'étranger).

L'intégralité des revenus factoriels revient à l'équilibre au consommateur représentatif dont le revenu est donc (pour le pays domestique) $R = wL + rK$. Celui-ci choisit sa consommation optimale de biens X et Y de façon à

9. On peut aussi interpréter le modèle HOS comme un modèle dans lequel la production combine du travail qualifié et du travail non qualifié. Avec cette interprétation, on peut utiliser les prédictions du modèle pour comprendre les écarts internationaux de salaires relatifs et l'augmentation des inégalités salariales au cours du temps. Voir l'interprétation proposée dans le paragraphe 2.3.3.

10. Cette hypothèse est probablement plus irréaliste aujourd'hui concernant le facteur capital, qui est relativement mobile internationalement. En revanche, le travail reste encore très largement immobile.

11. A noter que l'"abondance" d'un pays en un facteur de production est définie en termes relatifs. Comme dans le modèle ricardien, les avantages comparatifs sont basés sur des différences relatives. Le pays domestique est dit riche en capital si son ratio K/L est plus élevé que dans le pays étranger et ce même si sa dotation absolue en capital est plus petite ($K < K^*$). A la différence du modèle étudié dans le chapitre suivant, la taille (absolue) d'un pays n'a pas ici d'impact sur la forme de la spécialisation.

maximiser son utilité sous la contrainte budgétaire :

$$R \geq P_X C_X + P_Y C_Y$$

où P_X et P_Y sont les prix d'équilibre des biens et C_X et C_Y les consommations sectorielles. Le même raisonnement s'applique au consommateur représentatif étranger.

Les fonctions de productions, identiques dans les deux pays, combinent des unités de travail et de capital :

$$\begin{array}{lll} X = F_X(L_X, K_X) & \text{et} & Y = F_Y(L_Y, K_Y) \\ X^* = F_X(L_X^*, K_X^*) & \text{et} & Y^* = F_Y(L_Y^*, K_Y^*) \end{array}$$

Elles sont à rendements constants et à productivité marginale des facteurs décroissante.

Dans un cadre à deux facteurs, la frontière des possibilités de production n'est plus une simple droite. Pour en donner l'intuition, considérons le cas simple dans lequel la production combine des unités de travail et de capital qui ne sont pas substituables. Supposons par exemple qu'il faut a_{Ki} unités de capital et a_{Li} unités de travail pour produire une unité de bien i ($i = X/Y$). Dans ce cas, la frontière des possibilités de production dans le plan (X, Y) est déterminée par la confrontation des contraintes de ressources sur les marchés du travail et du capital :

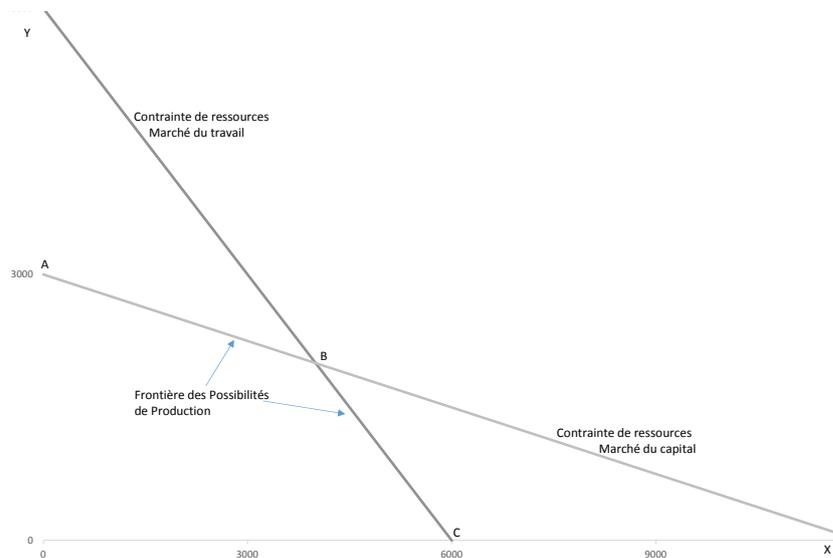
$$\begin{array}{l} L \geq a_{LX}X + a_{LY}Y \\ K \geq a_{KX}X + a_{KY}Y \end{array}$$

Elle est matérialisée sur le graphique 2.7, dans le cas où la production de bien X est relativement intensive en travail ($a_{LX}/a_{LY} > a_{KX}/a_{KY}$). La ligne gris foncé correspond à la contrainte de ressources sur le marché du travail, la ligne gris clair à la contrainte sur le marché du capital. Ces contraintes de ressources imposent de produire une combinaison de biens située sous ces deux droites, i.e. l'aire délimitée par les points A, B et C. Au point A, l'économie ne produit que du bien Y . Elle utilise toutes ses ressources de capital mais certaines unités de travail restent inutilisées. Au point C au contraire, elle est spécialisée complètement dans la production de bien X et n'utilise pas complètement ses ressources en capital. Finalement, il n'y a qu'au point B que l'économie utilise toutes ses ressources factorielles¹². Comme dans le cas ricardien, la pente de la frontière des possibilités de production mesure le

12. Cette spécificité du modèle disparaîtra une fois que les facteurs seront supposés substituables. Dans ce cas, on retrouve une infinité de combinaisons de produits pour lesquelles les dotations en facteurs sont utilisées intégralement.

coût d'opportunité du bien X en terme de bien Y . Cette fois-ci cependant, ce coût d'opportunité n'est pas constant. Il augmente lorsque la production relative de bien X s'intensifie puisqu'alors les ressources en capital deviennent relativement rares. Cette propriété demeure vraie dans le cas plus général où les facteurs de production sont substituables. La frontière des possibilités de production continue à être concave.

FIGURE 2.7 – Frontière des possibilités de production pour des facteurs non substituables



Lorsque les facteurs sont substituables, les coefficients techniques (a_{ij} , $i = X/Y$, $j = K/L$) ne sont plus constants. Ils sont déterminés à l'équilibre en fonction du prix du travail et du capital. En particulier, on a un lien négatif entre l'utilisation relative d'un facteur et son prix relatif (i.e. $a_{iK}/a_{iL} \equiv K_i/L_i$ est décroissant de r/w pour $i = X/Y$). Dans ce qui suit, on suppose cependant que la production du bien X est relativement intensive en travail, i.e. que, à prix des facteurs donnés, le secteur Y emploie relativement plus de capital :

$$\left. \frac{K_Y}{L_Y} \right|_{w/r} > \left. \frac{K_X}{L_X} \right|_{w/r}$$

Ces hypothèses technologiques sont illustrées sur le graphique 2.8. Pour les firmes de chaque secteur, la fonction de production détermine une infinité d'isoquantes, définies comme les différentes combinaisons de facteurs permettant de produire une certaine quantité de biens. Etant donné un objectif de

production, la firme choisit parmi une infinité de combinaisons de facteurs possibles, celle qui minimise ses coûts. Sur le graphique 2.8, les isoquantes permettant d'obtenir la quantité Y_0 de bien Y et la quantité X_0 de bien X sont représentées. Le choix optimal de la firme est déterminé par la maximisation de son profit (ou de manière équivalente, la minimisation de son coût étant donnée une contrainte de production). La firme représentative du secteur X , par exemple, résout le programme suivant :

$$\begin{array}{ll} \text{Max}_{L_X, K_X} & [P_X X - wL_X - rK_X] \\ \text{s.c.} & X = F_X(L_X, K_X) \end{array}$$

qui implique les conditions du premier ordre :

$$P_X \frac{\partial F_X(L_X, K_X)}{\partial L_X} = w \quad \text{et} \quad P_X \frac{\partial F_X(L_X, K_X)}{\partial K_X} = r$$

i.e.

$$\frac{\partial F_X(L_X, K_X)/\partial L_X}{\partial F_X(L_X, K_X)/\partial K_X} = \frac{w}{r}$$

De même dans le secteur Y :

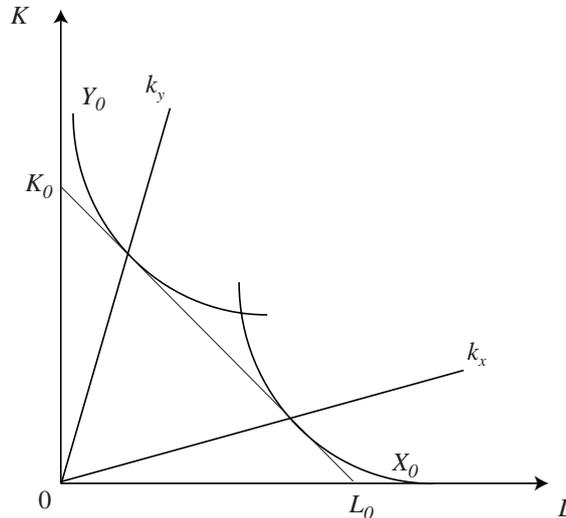
$$\frac{\partial F_Y(L_Y, K_Y)/\partial L_Y}{\partial F_Y(L_Y, K_Y)/\partial K_Y} = \frac{w}{r}$$

A l'optimum, la firme égalise le taux marginal de transformation au prix relatif des facteurs¹³. Autrement dit, la firme choisit sa combinaison optimale de travail et de capital de façon à ce que, à l'équilibre, la productivité marginale relative des facteurs soit juste égale à leur coût relatif. Du fait de la mobilité des facteurs, le coût relatif est égalisé entre les secteurs. A l'optimum, les taux marginaux de transformation sont donc égaux dans les secteurs X et Y , et ce malgré l'indépendance des choix de combinaisons factorielles faits par les entreprises représentatives de chaque secteur.

Ce choix optimal est représenté sur la Figure 2.8 par les droites notées k_X et k_Y . Ces deux droites représentent, pour chaque niveau de production X ou Y , la combinaison optimale de travail et de capital choisie par l'entreprise représentative. Sur cette droite, la pente des isoquantes est exactement égale au prix relatif du travail w/r , ce qui traduit le caractère optimal de ces combinaisons factorielles. Si ces droites ne sont pas suffisantes pour déterminer la répartition intersectorielle des facteurs, elles permettent d'illustrer l'hypothèse relative à l'intensité factorielle des différents secteurs. Plus haut,

13. Le taux marginal de transformation mesure le nombre d'unités de capital que l'entreprise doit employer pour compenser, en termes de production, la baisse de l'emploi de une unité.

FIGURE 2.8 – Détermination de l'utilisation relative des facteurs à l'optimum de la firme



Le graphique illustre les choix sectoriels d'intensité relative en facteurs, dans le plan $(L; K)$. Les courbes X_0 et Y_0 correspondent à une isoquante particulière, pour les secteurs X et Y respectivement. Tous les points de ces isoquantes correspondent à des combinaisons factorielles qui permettent de produire la même quantité de bien. L'entreprise représentative choisit la combinaison qui maximise son profit, lorsque l'isoquante est tangente à une droite de pente w/r . k_x et k_y représentent ces choix pour tout niveau de production, respectivement pour les secteurs X et Y .

on a supposé que la production du secteur X était relativement intensive en travail. Sur le graphique 2.8, cela se traduit par une pente plus faible de la droite k_x dans le plan (L, K) . Pour un niveau de production donné, la firme représentative du secteur X choisit d'allouer une proportion relativement plus importante de ses coûts au travail. Ce choix optimal est représenté sur la Figure 2.8 par les droites notées k_x et k_y . Ces deux droites représentent, pour chaque niveau de production X ou Y , la combinaison optimale de travail et de capital choisie par l'entreprise représentative. Sur cette droite, la pente des isoquantes est exactement égale au prix relatif du travail w/r , ce qui traduit le caractère optimal de ces combinaisons factorielles. Si ces droites ne sont pas suffisantes pour déterminer la répartition intersectorielle des facteurs, elles permettent d'illustrer l'hypothèse relative à l'intensité factorielle des différents secteurs. Plus haut, on a supposé que la production du secteur X était relativement intensive en travail. Sur le graphique 2.8, cela se traduit

par une pente plus faible de la droite k_X dans le plan (L, K) . Pour un niveau de production donné, la firme représentative du secteur X choisit d'allouer une proportion relativement plus importante de ses coûts au travail.

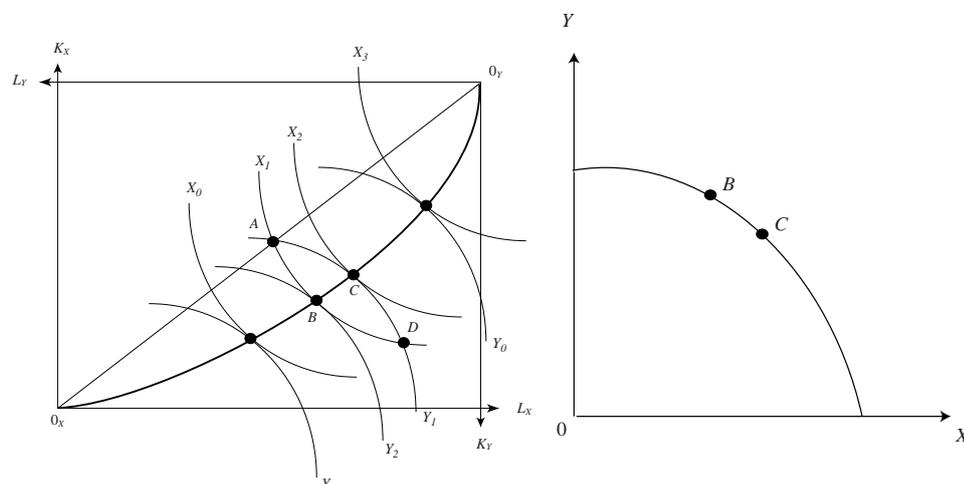
Une autre manière de représenter les combinaisons optimales choisies par les entreprises de chaque secteur consiste à utiliser une “Boîte d'Edgeworth” (graphique de gauche de la Figure 2.9). Comme dans le graphique 2.8 sont ici représentées les choix de combinaisons de facteurs offerts aux entreprises pour un niveau donné de production (les isoquantes notées $X_0... X_3$ et $Y_0... Y_3$ sur le graphique). La boîte d'Edgeworth prend également en compte les contraintes de ressources de l'économie dans son ensemble. En représentant simultanément les demandes en facteurs des deux secteurs et en les contraignant à respecter les contraintes de ressources de l'économie, ce graphique permet de déterminer la forme de la frontière des possibilités de production (graphique de droite sur la figure 2.9)¹⁴. Tous les points de la boîte d'Edgeworth correspondent à des répartitions factorielles entre les secteurs qui sont atteignables étant données les contraintes de ressources. Les points matérialisés sur la courbe en gras sont des points atteignables et respectant la contrainte d'optimalité des entreprises des deux secteurs (i.e. des points pour lesquels les taux marginaux de transformation sont égaux au coût relatif des facteurs). La position de cette courbe sous la diagonale de la boîte caractérise une situation dans laquelle la production du bien X est relativement intensive en travail. A chacun de ses points correspondent un niveau de production de chacun des secteurs, qui est ensuite reporté sur le graphique de droite. Le graphique de droite trace donc la frontière des possibilités de production, qui respecte les conditions d'optimalité des entreprises des deux secteurs.

La frontière des possibilités de production est bien concave, comme dans l'exemple dans lequel les facteurs n'étaient pas substituables¹⁵. Le coût d'opportunité de la production de bien X en termes de bien Y est croissant avec la quantité relative de bien X produite. Comme l'illustre la boîte d'Edgeworth, les ressources factorielles sont entièrement utilisées sur la frontière

14. La boîte d'Edgeworth est constituée de deux graphiques imbriqués. Le premier représente dans le plan (L_X, K_X) les combinaisons de facteurs offertes à l'entreprise du secteur X . Le second représente les choix du secteur Y , dans le plan (L_Y, K_Y) . La valeur maximale prise par (L_X, K_X) et par (L_Y, K_Y) correspond à l'ensemble des ressources disponibles dans l'économie (L, K) . Bien sûr, si toutes les ressources sont utilisées dans un secteur, par exemple $(L_X, K_X) = (L, K)$, elles ne sont plus disponibles pour l'autre secteur (donc $(L_Y, K_Y) = (0, 0)$). Cette interdépendance est prise en compte par le positionnement des deux plans, définissant une boîte. Dans cette boîte, chaque point définit une répartition des facteurs entre les deux secteurs qui est atteignable.

15. Par la suite, nous verrons qu'une modification des ressources factorielles modifie la forme de la frontière des possibilités de production.

FIGURE 2.9 – Boîte d'Edgeworth et frontière des possibilités de production



Le graphique de gauche représente la boîte d'Edgeworth des possibilités de production, pour les secteurs X et Y . Il est composé de deux graphiques imbriqués correspondant aux besoins en facteurs de chaque secteur. Tout point de la boîte correspond à des allocations factorielles des deux secteurs qui sont compatibles avec les contraintes de ressources. Tout point sur la courbe en gras correspond à des allocations qui sont compatibles avec le comportement d'optimisation des entreprises de chaque secteur. Le graphique de droite correspond à la frontière des possibilités de production dans le plan $(X; Y)$, déduite de la boîte d'Edgeworth.

des possibilités de production. Tout panier de productions sectorielles sur cette frontière conduit donc au plein-emploi. Comme dans le cas ricardien, la détermination du point sur la frontière correspondant au panier produit fait intervenir le prix relatif et les comportements de consommation.

Du côté de la demande, le programme du consommateur est standard. Étant donné des préférences relatives à la consommation de biens X et de biens Y (la fonction d'utilité $U(C_X, C_Y)$), le consommateur maximise son utilité sous la contrainte budgétaire :

$$\begin{aligned} \text{Max}_{C_X, C_Y} \quad & U(C_X, C_Y) \\ \text{s.c.} \quad & Y \geq P_X C_X + P_Y C_Y \end{aligned}$$

ce qui implique les conditions du premier ordre :

$$\frac{\partial U(C_X, C_Y)}{\partial C_X} = \lambda P_X \quad \text{et} \quad \frac{\partial U(C_X, C_Y)}{\partial C_Y} = \lambda P_Y$$

où λ est le multiplicateur de Lagrange. En réarrangeant, on trouve la condi-

tion d'optimalité suivante :

$$\frac{\partial U(C_X, C_Y)/\partial C_X}{\partial U(C_X, C_Y)/\partial C_Y} = \frac{P_X}{P_Y}$$

A l'équilibre, le consommateur égalise son taux marginal de substitution au prix relatif des deux biens. Graphiquement, cela revient à choisir le point situé sur la courbe d'indifférence la plus élevée tangent à la contrainte budgétaire.

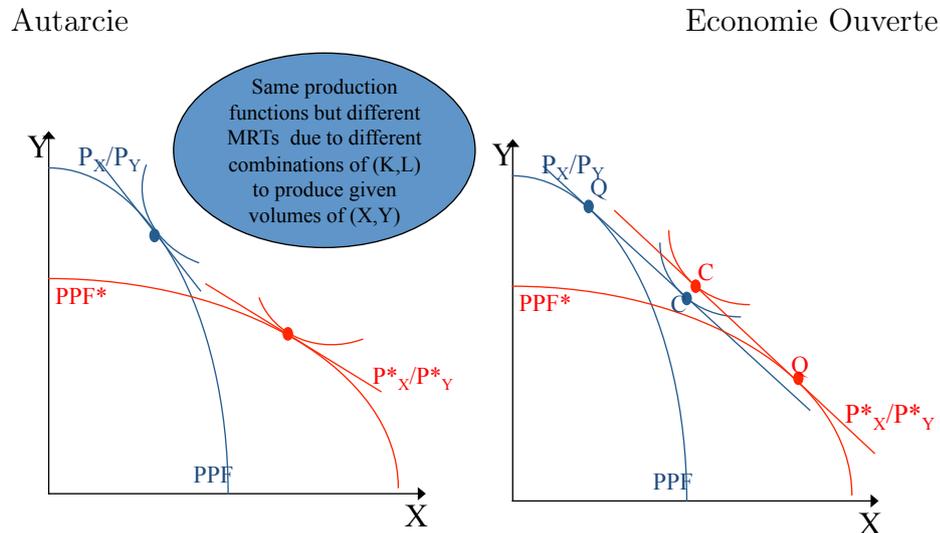
Le bouclage du modèle nécessite finalement de déterminer le niveau des prix relatifs permettant de mettre en cohérence les choix des entreprises représentatives (égalisation des taux marginaux de transformation au prix relatif des facteurs) et les choix du consommateur représentatif (égalisation du taux marginal de substitution et du prix relatif). Notons que le prix relatif d'équilibre est étroitement lié au niveau d'équilibre du salaire relatif. A l'équilibre concurrentiel, les prix relatifs sont en effet déterminés par le prix relatif des facteurs. Dans chaque secteur, le prix s'établit au niveau du coût marginal de production et dépend donc du niveau du salaire et du prix du capital. L'hypothèse selon laquelle le secteur X est relativement intensif en travail implique que le prix relatif de ce bien, à l'équilibre, est une fonction croissante du salaire relatif : $\frac{\partial(P_X/P_Y)}{\partial(w/r)} > 0$.

2.3.2 L'équilibre : Autarcie versus économie ouverte

Les paniers de consommation et de production sont représentés, dans le cas autarcique et en économie ouverte, sur la Figure 2.10. Considérons d'abord le cas autarcique (graphique de gauche). La frontière des possibilités de production de l'économie domestique est représentée en bleu, celle de l'étranger en rouge. Malgré l'absence de différences technologiques entre les deux pays, ces deux frontières de possibilité de production ne sont pas confondues. Cette propriété découle des différentiels de ressources factorielles entre les pays. L'économie domestique est supposée relativement riche en capital, ce qui explique que les possibilités de production sont biaisées en faveur des paniers "riches" en bien Y . Le capital étant relativement abondant dans le pays domestique, le pays peut plus facilement produire des biens qui requièrent une grande quantité de ce facteur.

Cette propriété conduit également à une production et une consommation à l'équilibre autarcique qui sont très différentes dans l'économie domestique et à l'étranger. La relative abondance en capital du pays domestique implique un prix relatif du travail relativement élevé à l'équilibre des marchés, et donc

FIGURE 2.10 – L'équilibre en autarcie et en économie ouverte



un prix relatif élevé pour le bien X , intensif en travail :

$$\frac{P_X}{P_Y} > \frac{P_X^*}{P_Y^*}$$

Ce différentiel de prix relatifs implique que le panier de bien produit et consommé par l'étranger est relativement plus riche en bien X . Dans cette configuration, chaque pays utilise relativement plus le facteur pour lequel sa dotation est la plus élevée, ce qui permet d'atteindre le plein-emploi des facteurs.

Comme dans le modèle ricardien, l'existence de différentiels de prix relatifs à l'équilibre autarcique est à l'origine de gains à l'échange international. Lorsque ces deux économies s'ouvrent au commerce international, il y a des possibilités d'arbitrage conduisant l'économie domestique à importer du bien X et à exporter du bien Y . Chaque pays se spécialise dans la production du bien qui utilise intensivement le facteur qu'il possède en quantités importantes. Ici aussi, il y a spécialisation des pays sur la base de leurs avantages comparatifs. Cependant, à la différence du modèle ricardien, ces avantages comparatifs ne sont pas liés à des différences technologiques entre les pays mais à une répartition hétérogène des ressources factorielles entre les pays. C'est la relative abondance en capital dans l'économie domestique qui réduit le prix relatif du capital, donc le prix relatif du bien Y .

La spécialisation de chaque pays dans son avantage comparatif conduit à une convergence des prix relatifs. L'augmentation de la demande de biens Y dans le pays domestique fait croître le prix relatif de ce bien, tandis que la demande supplémentaire de bien X dans l'économie étrangère conduit à une inflation du prix relatif de ce bien ($P_X/P_Y \downarrow$ et $P_X^*/P_Y^* \uparrow$). La convergence continue jusqu'au point où il n'existe pas de possibilités d'arbitrage, quand les prix relatifs sont égalisés entre les pays. Cet équilibre est représenté sur le graphique de droite de la Figure 2.10. Sur ce graphique, les points notés Q correspondent aux paniers de biens produits à l'équilibre. Par rapport à l'autarcie, chaque pays a intensifié sa production du bien pour lequel il dispose d'un avantage comparatif. Contrairement au cas autarcique, le panier de biens consommé n'est pas déterminé par le panier de biens produit. Les paniers de consommation sont notés C sur le graphique de la Figure 2.10. Ces paniers sont situés en dehors de la frontière des possibilités de production, sur une courbe d'indifférence plus élevée que celle atteinte à l'équilibre autarcique. Le commerce augmente le bien-être des consommateurs en leur permettant de consommer des paniers de biens qu'ils ne peuvent pas produire en autarcie.

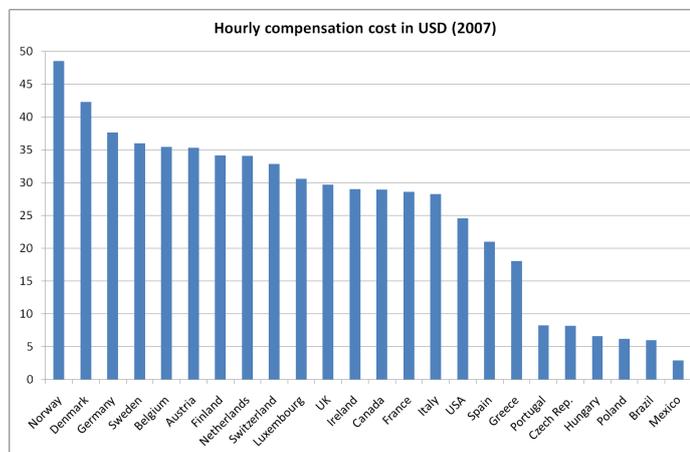
Finalement, le modèle HOS prédit également une convergence internationale du prix relatif des facteurs, via la convergence des prix. En effet, l'augmentation de la production relative de bien Y dans l'économie domestique induit une augmentation de la demande relative de capital qui exerce une pression sur le prix relatif du capital. Celui-ci augmente, tandis qu'à l'inverse, le prix relatif du travail augmente dans le pays étranger ($w/r \downarrow$ et $w^*/r^* \uparrow$). Dans ce cadre, la spécialisation réalloue la production mondiale vers les pays qui disposent en abondance des facteurs utilisés intensivement dans la production. Cette réallocation conduit à une convergence des coûts factoriels relatifs, malgré l'absence de mobilité internationale des facteurs ¹⁶ :

$$\frac{P_X}{P_Y} = \frac{P_X^*}{P_Y^*} \quad \Rightarrow \quad \frac{w}{r} = \frac{w^*}{r^*}$$

Ce résultat est résumé dans le **Théorème de Heckscher-Ohlin-Samuelson** qui stipule que “*le commerce international conduit à une égalisation du prix relatif des facteurs via la convergence internationale des prix relatifs*”. Le commerce joue donc le rôle d'un échange indirect de facteurs entre les pays.

16. Pour comprendre pourquoi la convergence des prix des facteurs est totale, il faut utiliser les hypothèses de base du modèle. La concurrence parfaite implique que, dans chaque secteur et chaque pays, le prix d'équilibre est égal au coût marginal de production. Comme les prix relatifs s'égalisent entre les pays, les coûts marginaux relatifs s'égalisent également. Enfin, l'uniformité des fonctions de production entre les pays implique que l'égalisation des coûts marginaux relatifs conduit à l'égalité du prix relatif des facteurs.

TABLE 2.4 – Ecart salarial au sein de l’OCDE



Source : US Bureau of Labor Statistics (August 2009)

En vendant du bien Y au pays étranger, le pays domestique permet à l'étranger d'utiliser indirectement une partie de sa dotation abondante en capital. Le pays "vend" son capital contre du travail étranger. Et c'est cet échange qui conduit à la convergence du prix relatif des facteurs.

Un "test" empirique simple de cette prédiction consiste à comparer les salaires horaires entre pays ouverts au commerce international. Empiriquement, le capital est relativement mobile internationalement, contrairement à ce que le modèle suppose. Par conséquent, son prix est relativement homogène entre pays. Les écarts internationaux de prix relatifs du travail sont donc principalement attribuables aux écarts de salaires. Cette comparaison est illustrée sur le graphique de la Figure 2.4. Celui-ci reporte le coût horaire du travail en 2007 pour différents pays de l'OCDE. Il montre des écarts élevés entre les pays, qui vont jusqu'à un rapport de 1 à 20 entre le Mexique et la Norvège.

Nous reviendrons dans le paragraphe 2.3.5 sur des tests plus formels du modèle HOS. A ce stade, il est important de noter que la convergence *totale* du prix relatif des facteurs prédite par le modèle repose sur des hypothèses extrêmes qui ne sont certainement pas vérifiées empiriquement. En particulier, pour obtenir ce résultat il faut i) que les deux pays continuent à produire des deux biens (i.e. qu'il n'y ait pas de spécialisation complète), ce qui peut ne pas être le cas si les deux pays ont des dotations factorielles suffisamment différentes, ii) que les pays partagent les mêmes technologies de production de façon à ce que la convergence des prix relatifs implique celle des facteurs,

iii) que la convergence des prix relatifs soit elle-même totale ce qui implique l'absence de barrières à l'échange. Si ces trois hypothèses sont certainement contrefactuelles, l'intuition selon laquelle le commerce conduit à une convergence des prix des facteurs, même partielle, reste pertinente.

2.3.3 Ouverture et inégalités de revenus : le théorème de Stolper-Samuelson

Contrairement au modèle ricardien, le modèle HOS permet d'étudier la manière dont le commerce international affecte la distribution des revenus au sein de l'économie nationale. Si l'échange reste Pareto-améliorant au niveau global, le modèle met aussi en évidence un creusement des inégalités de revenus entre les détenteurs des différents facteurs productifs. Plus spécifiquement, il montre que les "gagnants" de la globalisation sont les détenteurs de facteurs utilisés intensivement dans la production du bien exporté (les détenteurs de capital dans l'économie domestique et les détenteurs de travail dans l'économie étrangère).

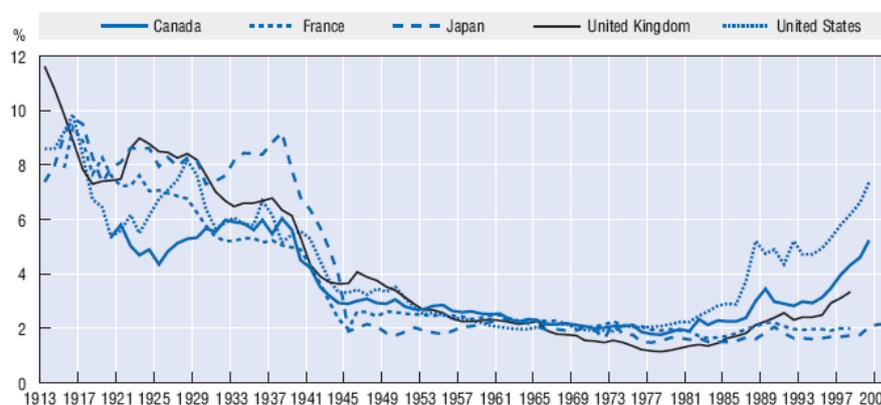
L'origine de ce résultat tient à la manière dont la convergence des prix modifie l'utilisation relative de travail et de capital dans l'économie. Prenons par exemple le cas de l'économie domestique. L'ouverture au commerce conduit à une hausse de la demande relative de bien Y , qui exerce une pression à la hausse sur le prix relatif du capital ($r/w \uparrow$). Dans chaque secteur, l'augmentation du prix relatif du capital conduit les entreprises représentatives à substituer du travail à des unités de capital, afin de rétablir l'égalité entre le taux marginal de transformation et le prix relatif des facteurs. Avec des fonctions de production à productivité marginale décroissante, cette substitution conduit à une augmentation de la productivité marginale du capital et une diminution de la productivité marginale du travail ($\partial F_i(L_i, K_i)/\partial K_i \uparrow$ et $\partial F_i(L_i, K_i)/\partial L_i \downarrow$, $i = X/Y$). Finalement, comme les facteurs sont rémunérés à leur productivité marginale, la substitution conduit à une augmentation de la rémunération réelle du capital et à une baisse de la rémunération réelle du travail ($w/P_i = \partial F_i(L_i, K_i)/\partial L_i \downarrow$ et $r/P_i = \partial F_i(L_i, K_i)/\partial K_i \uparrow$, $i = X/Y$). A l'inverse, dans l'économie étrangère, la rémunération réelle des travailleurs augmente quand celle des capitalistes diminue.

Ce résultat est résumé dans le **Théorème de Stolper-Samuelson** : *"Une augmentation du prix relatif d'un bien augmente la rémunération relative du facteur qui est utilisé intensivement dans la production de ce bien et diminue la rémunération relative de l'autre facteur."* Dans le cas de l'ouverture au commerce, ce théorème implique que les gagnants de la libéralisation commerciale sont les détenteurs du facteur abondant dans l'économie tandis

que les détenteurs du facteur relativement peu abondant y perdent.

C'est ce partage inégal des gains à l'échange qui peut conduire à des conflits d'intérêt entre groupes à l'intérieur d'une économie, et à des pressions protectionnistes. En effet, les groupes d'intérêt exposés à la concurrence des biens importés ont tout intérêt au protectionnisme quand les détenteurs de facteurs abondants dans l'économie militent en faveur de politiques de libre-échange. En théorie, il existe des systèmes de transferts des gagnants vers les perdants qui doivent permettre de compenser ces inégalités, sans avoir à recourir à des politiques protectionnistes, coûteuses au niveau agrégé (puisqu'elles font perdre à l'économie les bénéfices du libre-échange). En pratique cependant, il peut être difficile de mettre en place de telles politiques. Par exemple, pour un pays relativement riche en capital, la politique de redistribution adaptée consiste à taxer les capitalistes pour subventionner les travailleurs qui souffrent de la libéralisation. L'efficacité d'une telle politique est cependant mise en cause dès lors que les capitaux sont mobiles entre les pays (même imparfaitement).

FIGURE 2.11 – L'augmentation des inégalités de revenus : Part dans le revenu national des 0.1% les plus riches pour 5 pays de l'OCDE (1913-2001)



Source : Piketty & Saez (2006).

Le théorème de Stolper-Samuelson a reçu un écho important dans les milieux académiques, notamment parce qu'il était considéré comme une explication de l'augmentation tendancielle des inégalités de revenus à partir des années 70 dans la plupart des pays développés¹⁷. Cette augmentation tendancielle est illustrée, pour 5 pays de l'OCDE, sur le graphique 2.11. Ce

17. Pour une revue de littérature exhaustive sur ces questions, voir Piketty (2013).

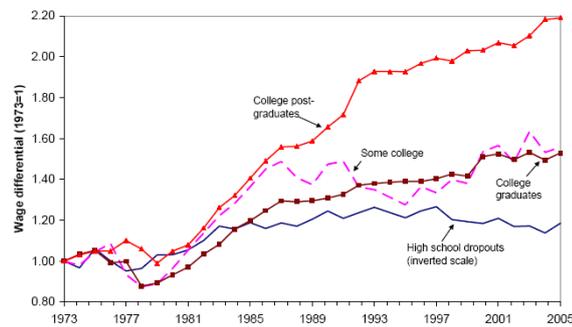
graphique mesure la contribution au revenu national des 0.1% individus les plus riches. Sans surprise, ces individus contribuent disproportionnellement au revenu national (entre 2 et 10% du revenu total). Ce qui est plus intéressant c'est l'augmentation tendancielle de cette contribution à partir de la deuxième moitié des années 70, après une période de stagnation de 30 ans depuis la fin de la deuxième guerre mondiale. Cette tendance a souvent été mise en miroir de la vague de libéralisation commerciale et financière à la même période, qui a fortement intensifié le commerce Nord-Sud (voir discussion en introduction). Ces flux bilatéraux se caractérisent par des différences importantes en termes d'intensités factorielles. Les pays du Sud (principalement les nouveaux pays industrialisés d'Asie) exportent des produits relativement intensifs en travail comme des habits ou des chaussures. Les pays avancés exportent vers ces pays des produits plus sophistiqués, dont la production est intensive en capital et en travail qualifié (produits chimiques, machines).

La concomitance de ces deux tendances conduit souvent les analystes à inférer une relation de cause à effet. Si cette relation n'est pas clairement établie empiriquement, elle est en tout cas conforme au théorème de Stolper-Samuelson dans un modèle HOS où les facteurs de production sont le travail qualifié et le travail non qualifié. Si on admet que les pays développés sont relativement abondants en travail qualifié, l'intégration commerciale de ces économies avec les économies émergentes à partir des années 70 a dû conduire à une spécialisation profitant aux travailleurs qualifiés, dont la rémunération a alors augmenté, au détriment des travailleurs non qualifiés. Quand on regarde plus spécifiquement l'évolution comparée des salaires aux Etats-Unis (graphique du haut de la Figure 2.5), on observe effectivement un tel creusement des inégalités salariales. Entre 1973 et 2005, l'écart de salaires d'un travailleur américain diplômé du lycée et d'un travailleur détenteur d'un diplôme du troisième cycle a augmenté de 120%. La tendance est beaucoup moins marquée pour les salaires français, du fait de rigidités salariales plus importantes. Pour ce pays, le creusement se fait au niveau de la probabilité d'emploi (graphique du bas de la Figure 2.5). En France, l'apparition du chômage de masse dans les années 70 s'est faite principalement dans les catégories de travailleurs peu qualifiés.

Si les évolutions jointes sont effectivement cohérentes avec la théorie néo-classique, les études empiriques concluent cependant que le commerce international n'est pas la cause principale du creusement des inégalités. En effet, plusieurs types d'évidences contredisent l'hypothèse d'un lien direct de cause à effet entre les deux phénomènes. D'abord, l'augmentation du salaire relatif des qualifiés ne s'accompagne pas, dans les économies développées, d'une augmentation du prix relatif des biens intensifs en travail qualifié. Ensuite, on n'observe pas non plus de réduction des inégalités salariales dans les pays du

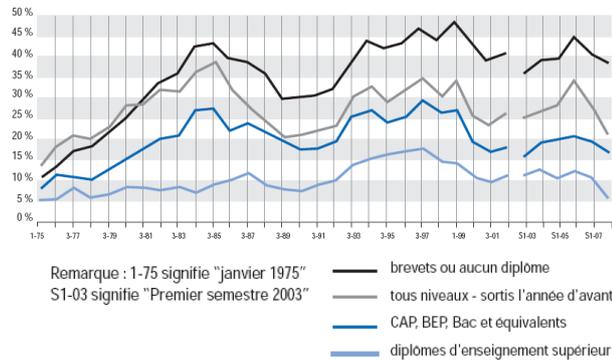
TABLE 2.5 – Inégalités de salaires et inégalités d'emploi

Etats-Unis : Différentiel de salaire horaire (en comparaison des diplômés du lycée, hommes)



Source : Lemieux (2008)

France : Taux de chômage 1-4 ans après la sortie du système éducatif



Source : INSEE

Sud¹⁸. Enfin, le volume relativement faible du commerce en pourcentage de la production nationale implique que la contribution du commerce au creusement des inégalités ne peut au maximum qu’être faible¹⁹. Les économistes expliquent plutôt le creusement des inégalités par une évolution technologique biaisée en faveur des travailleurs qualifiés. Dans le cadre HOS, un tel progrès technique biaisé conduit effectivement les entreprises des deux secteurs et des deux pays à augmenter leur emploi relatif de travailleurs qualifiés, ce qui exerce une pression à la hausse sur leurs salaires. Si le progrès technique ne s’accompagne pas d’une modification des dotations factorielles, les pays du Nord continuent cependant à exporter des produits dont la production utilise intensivement le travail qualifié.

2.3.4 Le rôle des dotations factorielles : le théorème de Rybczynski

La structure du modèle HOS en fait également un bon outil pour étudier la manière dont des changements dans la dotation factorielle d’une économie peuvent affecter la structure de sa production. Une telle analyse s’applique par exemple à des pays en phase d’émergence, qui accumulent du capital grâce à des taux de croissance importants, et souvent plus élevés que la croissance démographique. Un tel phénomène conduit donc le pays à s’enrichir relativement en capital, ce qui modifie la structure de ses avantages comparatifs.

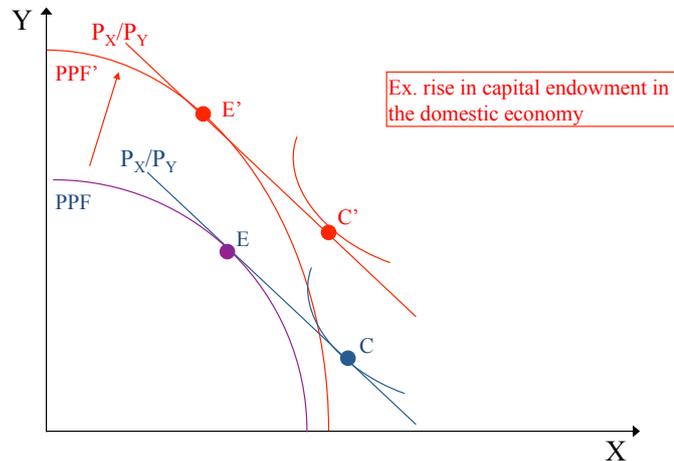
Une telle dynamique peut s’analyser à travers le modèle HOS, en comparant deux équilibres d’économie ouverte de pays aux dotations relatives en facteurs différentes. C’est ce qui est fait dans le graphique 2.12, en supposant pour simplifier l’analyse que le pays est “petit”, i.e. que la modification de sa structure productive n’a pas d’influence sur le prix relatif des biens (P_X/P_Y inchangé). Une augmentation de la dotation relative en capital du pays se traduit par une modification de la forme de sa frontière des possibilités de production, qui passe de la courbe en bleu à la courbe en rouge. L’expansion des possibilités productives est biaisée en faveur du bien intensif en capital.

A prix relatif donné, le pays intensifie sa spécialisation en bien Y : au nouveau point de production E' il produit plus de bien Y et moins de bien X . En effet, le prix relatif du travail ne changeant pas, l’intensité factorielle de chaque secteur reste également la même (i.e. K_X/L_X et K_Y/L_Y inchangés).

18. Piketty (2013) documente une augmentation des inégalités de revenus sur la même période dans certains pays émergents, dont la Chine. Sous l’hypothèse d’un avantage comparatif de ces pays dans la production de biens intensifs en travail peu qualifié, l’évolution est contraire aux prédictions du théorème de Stolper-Samuelson.

19. Pour une revue de littérature sur ces sujets, voir Lawrence (1996).

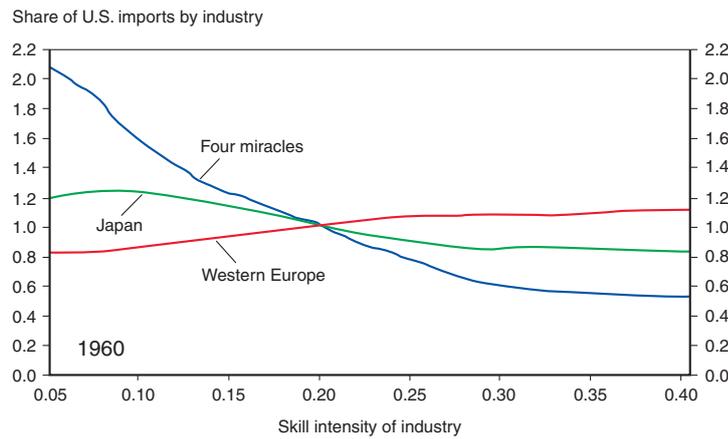
FIGURE 2.12 – Augmentation des dotations relatives en capital



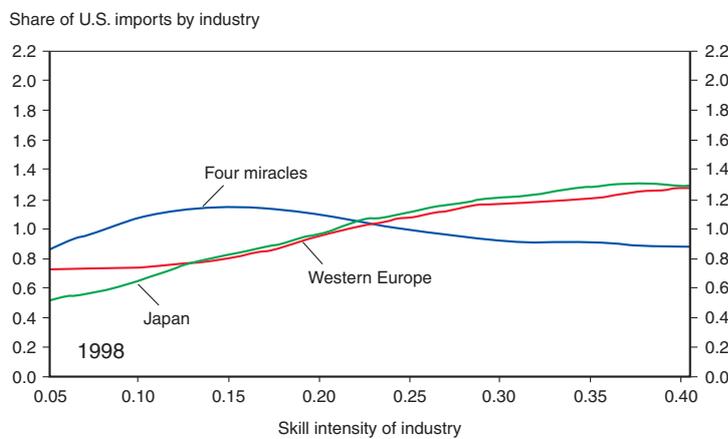
Pour faire face à l'augmentation relative de la dotation en capital, l'économie n'a donc pas d'autre choix que de produire plus du bien qui utilise ce facteur intensivement. Cette modification de la structure productive, en augmentant la spécialisation de l'économie, est une source de gains en bien-être. Ce résultat est résumé par le **Théorème de Rybczynski** : *“Pour un prix relatif donné, une augmentation de la dotation en un facteur de production augmente la production du bien qui est intensif dans ce facteur et diminue la production de l'autre bien.”*

Ce résultat est important parce qu'il implique que les avantages comparatifs ne sont pas une caractéristique inaliénable des pays. Au contraire, la croissance d'un pays peut s'accompagner d'une modification de la structure de ses avantages comparatifs donc de son commerce. Un pays comme la Corée du Sud a connu une période de croissance importante dans les années 80, qui a conduit à une accumulation massive de capital et à une intensification de sa production de biens riches en capital comme les produits électroniques. Cette évolution est illustrée sur la Figure 2.13 tirée de Krugman et al. (2012) et obtenue à partir d'exportations japonaises, européennes et asiatiques vers les Etats-Unis. Le calcul porte sur la part de ces pays dans les importations totales américaines, par secteur, pour 1960 et 1998. L'axe des abscisses mesure l'intensité en travail qualifié de la production des différents biens considérés. En 1960, les importations en provenance d'Europe de l'Ouest sont relativement plus importantes dans les secteurs intensifs en travail qualifié tandis que pour les quatre “dragons” asiatiques (Corée du Sud, Taiwan, Hong Kong

FIGURE 2.13 – Evolution des avantages comparatifs



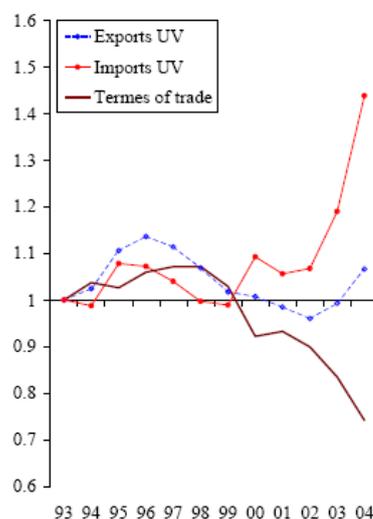
(a) 1960



(a) 1998

Source : Krugman et al. (2012). Le graphique montre la part dans les importations totales américaines de différents pays d'origine (Europe de l'Ouest, Japon et "Four miracles" i.e. Corée du Sud, Taiwan, Hong Kong et Singapour). Le premier graphique correspond aux données de 1960, le second à celles de 1998. Ce calcul est fait sur différents secteurs, classés en fonction de l'intensité en travail qualifié de leur production (axe des abscisses).

FIGURE 2.14 – Evolution des termes de l'échange de la Chine



Source : Gaulier, Lemoine & Unal-Kesenci (2006)

et Singapour) et, dans une moindre mesure, le Japon, l'avantage comparatif est très clairement dans les biens peu intensifs en travail qualifié. En 1998, le graphique est assez différent. En particulier, les dragons asiatiques ont nettement reculé dans les importations américaines de produits peu intensifs en travail qualifié. En revanche, leur poids dans les importations de biens plus intensifs en travail qualifié a augmenté. Cette évolution est cohérente avec les prédictions du théorème de Rybczynski, dans un contexte de développement rapide de ces pays, accompagné d'une accumulation de capital physique et humain.

Avant de conclure ce paragraphe, il est important de noter que le théorème de Rybczynski repose sur une hypothèse forte d'exogénéité des prix relatifs. Lorsqu'un pays comme la Chine, par exemple, accumule du capital sur une période longue et devient un acteur majeur du commerce international, on ne peut pas supposer une absence de réaction des prix relatifs. Dans ce cas de figure par exemple, on observe au cours du temps une détérioration nette de ses termes de l'échange (graphique 2.14). Cette évolution défavorable des termes de l'échange (stabilité du prix des biens exportés mais augmentation prononcée du prix des biens importés) annule au moins partiellement l'effet positif pour le pays d'une croissance tirée par les exportations.

TABLE 2.6 – Contenu en facteurs du commerce américain, 1962

	Imports	Exports
Capital par million de dollars	\$2 132 000	\$1 876 000
Travailleurs par million de dollars	119	131
Ratio Capital-Travail (dollars par travailleur)	\$17 916	\$14 321
Nombre moyen d'années d'étude par travailleur	9,9	10,1
Proportion d'ingénieurs et de scientifiques dans l'emploi	0,0189	0,0255

Source : Baldwin (1971) cité par Krugman et al. (2012)

2.3.5 Tests empiriques : Le paradoxe de Leontief

La validation empirique du modèle HOS repose sur la comparaison des flux sectoriels de commerce et des dotations relatives en facteurs des pays. Dans un cadre multi-pays/multi-produits/multi-facteurs, le théorème de Heckscher-Ohlin se généralise en une corrélation positive entre l'abondance relative d'un facteur dans un pays et ses exportations de biens utilisant intensivement ce facteur. On s'attend à ce que les pays exportent les produits qui utilisent intensivement les facteurs dont ils disposent en abondance.

Leontief (1953) propose une premier "test" empirique de ces prédictions basé sur des données américaines. Ce test n'est pas favorable à la théorie néo-classique du commerce puisqu'il montre que, contrairement aux prédictions du modèle, les Etats-Unis, pays le mieux doté en capital au monde, exportent des produits qui sont plutôt intensifs en travail tandis qu'ils importent des biens produits avec une technologie gourmande en capital. C'est ce qu'on appelle le **paradoxe de Leontief**.

Ce paradoxe est illustré dans la Table 2.6, tirée de Baldwin (1971) et citée par Krugman et al. (2012). Il confirme les résultats plus anciens de Leontief : le ratio capital sur travail est plus élevé dans les importations américaines que dans ses exportations. Cette apparente incohérence des données avec les prédictions du modèle est cependant modérée par une observation plus détaillée. En effet, si le ratio capital/travail est bien plus petit dans les secteurs d'exportation nette, le contenu en travail qualifié (mesuré par le nombre moyen d'années d'étude des travailleurs) ainsi que le contenu technologique (mesuré par l'utilisation d'ingénieurs et de scientifiques) est plus élevé dans les exportations américaines que dans ses importations. Le paradoxe de Leontief pourrait donc venir de l'absence de prise en compte du capital humain dans le modèle de base.

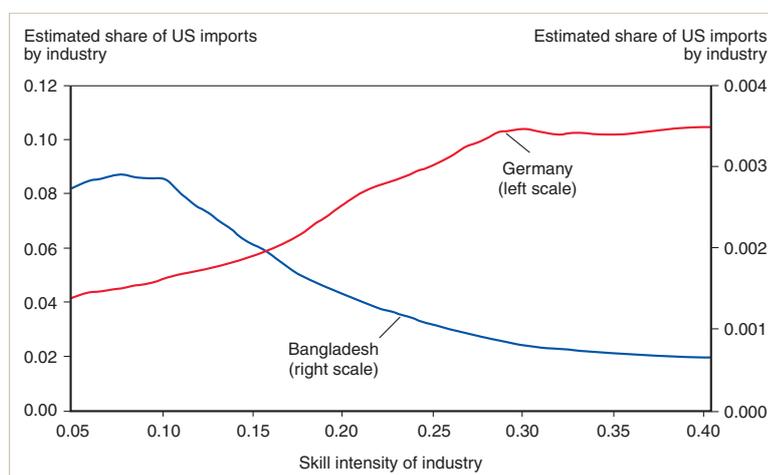
Plusieurs explications ont été proposées pour expliquer le paradoxe de Leontief. Par exemple, le fait d'appliquer l'intuition du modèle HOS à l'ensemble du commerce américain a été critiqué car il ne permet pas de prendre en compte certaines particularités sectorielles (e.g. le fait que certains biens importés par les Etats-Unis n'ont pas de substituts domestiques, par exemple certaines matières premières). Cela conduit aussi à négliger l'hétérogénéité des partenaires des Etats-Unis, alors même que le modèle est basé sur une comparaison de dotations relatives (certains partenaires sont relativement plus dotés en capital que d'autres). Enfin, le modèle repose sur des hypothèses extrêmes comme l'homogénéité des facteurs de production, l'uniformité des technologies productives ou encore la mobilité parfaite des facteurs entre les secteurs.

Pour répondre à ces critiques, des tests plus sophistiqués ont été proposés. L'extension du modèle HOS à n facteurs par Vanek (1968) permet de résoudre le problème de l'hétérogénéité des facteurs. Dans ce cadre d'analyse, l'abondance relative d'un pays dans un facteur de production se mesure par la part de sa dotation dans la dotation mondiale en ce facteur, comparée à la part de ce pays dans le revenu mondial. Cette mesure permet également d'étendre l'analyse à plus d'un pays. C'est ce que font Bowen, Leamer & Sveikauskas (1987) à partir de données relatives à 27 pays et 12 facteurs productifs. La prédiction testée est qu'on devrait observer des exportations nettes positives dans les secteurs utilisant intensivement les facteurs dont le pays est doté en quantité abondante. Bowen et al. (1987) montrent cependant que les données ne sont pas conformes à cette prédiction dans environ 40% des cas ce qui tend à confirmer le paradoxe de Leontief.

Trefler (1995) affirme qu'une partie du problème tient à l'hypothèse d'homogénéité des technologies de production. Dans son test du modèle HOS basé sur des données de 1983 relatives à 9 facteurs de production et 33 pays, il montre que, non seulement la corrélation entre les exportations factorielles nettes et les dotations relatives n'est que de 28%, mais aussi que le volume du commerce est insuffisant, par rapport à ce qui est prédit par le modèle. Par exemple, pour le commerce entre la Chine et les Etats-Unis, le différentiel de dotations relatives en travail est tel qu'on devrait observer un flux d'exportations massif de la Chine vers les Etats-Unis pour des produits intensifs en travail. Le fait que le volume de ce flux soit bien moindre que ce qui est prédit suggère que les Etats-Unis compensent une partie de leur désavantage comparatif par une meilleure productivité du travail de sorte qu'en termes *effectifs*, le différentiel de dotations factorielles est moindre.

Enfin, Romalis (2004) montre que les performances empiriques du modèle HOS sont meilleures lorsque l'échantillon de pays est restreint à des pays suffisamment différents en termes de dotations factorielles. Ainsi, si on compare

FIGURE 2.15 – Importations des Etats-Unis en provenance de l'Allemagne et du Bangladesh, par intensité sectorielle en travail qualifié



Source : Romalis (2004).

deux pays très différents en termes de dotations, comme l'Allemagne et le Bangladesh, on observe que les flux sectoriels d'exportation vers les Etats-Unis sont effectivement conformes aux prédictions du modèle. C'est ce qui est illustré sur la Figure 2.15. La part de l'Allemagne dans les importations américaines est croissante de l'intensité en travail qualifié du secteur. Au contraire pour le Bangladesh, c'est dans les secteurs peu intensifs en travail qualifié que les exportations sont les plus importantes. Ce résultat est cohérent avec la théorie néo-classique, sous l'hypothèse d'un avantage comparatif relatif du Bangladesh dans des productions intensives en travail peu qualifié.

Résumé :

- Dans le modèle HOS, l'ouverture au commerce conduit à une spécialisation des pays dans la production des biens qui utilisent intensivement les facteurs disponibles en relative abondance dans le pays.
- La spécialisation est induite par l'existence d'écarts de prix relatifs à l'équilibre autarcique, ces écarts offrant au consommateur des possibilités d'arbitrage international une fois les économies ouvertes au commerce.
- La spécialisation est Pareto-améliorante car elle permet aux pays d'utiliser leurs dotations factorielles de manière plus efficace, donc de consommer plus à l'équilibre.
- Dans le modèle HOS, elle conduit à une augmentation des inégalités de

revenus entre détenteurs de facteurs. Les gagnants de l'ouverture commerciale sont les détenteurs des facteurs utilisés intensivement dans les secteurs d'exportation tandis que les perdants sont ceux dont les facteurs sont soumis à la concurrence étrangère.

- La structure des avantages comparatifs n'est pas inaliénable. Une croissance inégalement partagée entre les différents types de facteurs conduit à une modification de la structure du commerce d'un pays.

2.4 Conclusion

Dans ce chapitre, nous avons étudié les modèles de commerce de type néo-classique. Ces modèles permettent de comprendre pourquoi et comment des pays différents en termes de technologies ou de dotations factorielles ont intérêt à échanger. Dans un environnement parfaitement concurrentiel, on montre comment le commerce de biens homogènes peut augmenter le bien-être agrégé en améliorant l'efficacité avec laquelle les ressources sont utilisées. Lorsque les pays sont différents en termes de productivité et/ou de dotations factorielles, la spécialisation internationale sur la base d'avantages comparatifs est Pareto-améliorante. En se spécialisant dans la production des biens pour lesquels il dispose d'un avantage comparatif, soit de par sa technologie, soit de par l'abondance des facteurs utilisés dans la production, chaque pays améliore l'efficacité avec laquelle il utilise ses dotations factorielles. L'échange international lui permet ensuite d'échanger ces biens contre des produits que les autres pays fabriquent de manière plus efficace. C'est par ce biais que le pays profite de l'optimisation de l'utilisation des facteurs, qui lui permet de consommer plus que ce qu'il pourrait produire domestiquement.

L'analyse néo-classique du commerce a montré que cette amélioration de l'allocation des facteurs peut se faire par les lois du marché. En effet, la structure des avantages comparatifs fait que les prix relatifs autarciques ne sont pas égalisés entre pays. En cas de libéralisation commerciale, ces écarts de prix relatifs conduisent à un arbitrage international de la part des consommateurs. L'augmentation de la demande relative de biens que cet arbitrage induit explique la spécialisation des structures productives. Dans ce contexte, une politique d'ouverture au commerce et de laissez-faire est donc optimale.

Si les gains à l'échange sont mutuels dans ces modèles, ils ne profitent pas de manière homogène à tous les agents de l'économie nationale. En particulier, le modèle Hecksher-Ohlin-Samuelson montre comment l'échange peut conduire à une dégradation de la situation des détenteurs de facteurs utilisés abondamment dans l'industrie qui s'expose à la concurrence étrangère.

Même si le commerce est globalement Pareto-améliorant, il comporte un coût pour l'économie nationale, lié à l'augmentation des inégalités de revenus entre les différents propriétaires de facteurs. Cette conséquence du commerce est importante car elle peut justifier des politiques interventionnistes dans un monde où le planificateur social se soucie du niveau des inégalités (planificateur Rawlsien). A l'optimum, le libre-échange reste la meilleure solution. Il peut s'accompagner de politiques redistributives de compensation des perdants par une taxe sur les bénéfices de la libéralisation. Dans un monde de *second best*, il n'est pas toujours possible de mettre en place ce type de politiques et le protectionnisme peut devenir la meilleure option.

La théorie néo-classique du commerce a reçu un support empirique mitigé. Si elle ne fonctionne pas très bien pour expliquer le commerce agrégé, ses prédictions sont conformes aux données relatives à l'échange bilatéral de pays suffisamment différents en termes de dotations factorielles et/ou de technologies. Ces résultats conduisent à s'interroger sur les déterminants du commerce entre pays similaires en termes de dotations factorielles, de technologies ou de niveaux de développement. C'est l'objet du chapitre suivant.

Chapitre 3

Commerce international en concurrence imparfaite

Les modèles ricardiens expliquent le commerce de biens *différents* entre pays à la technologie ou aux dotations *différentes*. De la fin de la seconde guerre mondiale au début des années 80, c'est cependant le commerce Nord-Nord - entre pays *similaires* - qui croît. L'intégration européenne conduit en effet à une intensification du commerce intra-européen. Parallèlement, les flux entre l'Europe de l'Ouest et les Etats-Unis continuent à être importants.

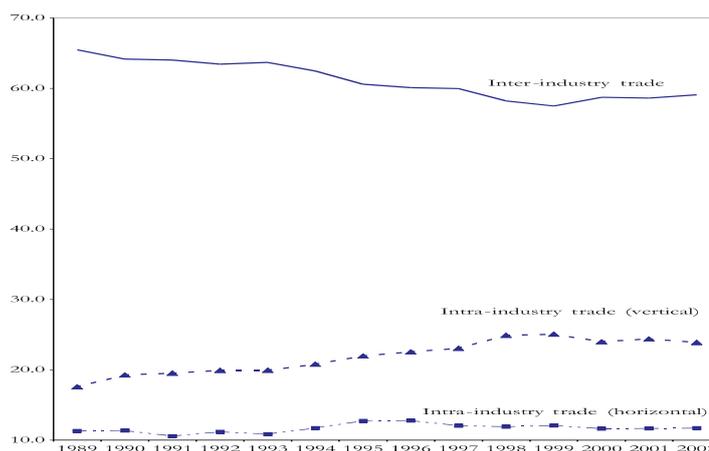
De tels flux Nord-Nord sont difficiles à rationaliser dans le cadre du modèle néo-classique. Et ce d'autant plus que ces échanges sont souvent de type "intra-industriel". Jusqu'ici, l'hypothèse implicite a été celle d'un échange de biens différents, soit par leur technologie soit par leur intensité factorielle. Empiriquement, on constate cependant qu'une part importante du commerce mondial est de type "intra-industriel". Au niveau bilatéral, on observe en effet un volume important d'échanges croisés de biens similaires : la France vend des voitures en Allemagne pour lui acheter ... des voitures.

La part relative du commerce intra- et interindustriel dans le commerce mondial est illustrée sur la Figure 3.1 tirée de Fontagné, Freudenberg & Gaulier (2006)¹. Les auteurs estiment que, sur la période 1986-2002, le commerce interindustriel (échange de biens différents au sein d'une paire de pays) représente environ 65% des flux internationaux. La part du commerce intra-

1. Dans la mesure où il n'existe pas de classification "officielle" dans les données de commerce international, la catégorisation des flux de commerce est basée sur une estimation. En travaillant à un niveau suffisamment fin (ici la nomenclature SH à 6 chiffres qui comprend un peu plus de 5000 biens), Fontagné et al. (2006) distinguent le commerce intra-industriel - défini comme la part des échanges portant sur des flux bilatéraux réciproques pour une paire de pays donnée et un secteur particulier - et le commerce interindustriel qui couvre le résidu.

industriel est donc de 35%, en légère augmentation au cours de la période d'observation.

FIGURE 3.1 – Decomposition du commerce mondial (en pourcentage du total)



Source : Fontagné et al. (2006). Définitions : Le commerce intra-industriel (“Intra-industry trade”) est défini comme l’échange bilatéral simultané de biens dans un sens et dans l’autre, au sein d’une catégorie industrielle fine. Le commerce est dit horizontal lorsque la différenciation des biens échangés se fait sur des caractéristiques subjectives. Il est dit vertical lorsque la différenciation permet de classer les variétés selon un ordre sur lequel les consommateurs peuvent s’entendre (différenciation en qualité). Empiriquement, la distinction est faite en utilisant les écarts observés sur le prix des biens échangés réciproquement.

Si le commerce interindustriel reste majoritaire, la part du commerce intra-industriel est loin d’être négligeable. Pour certaines paires de pays, elle est même très largement majoritaire (Table 3.1). De tels échanges croisés de biens similaires pourraient être à l’origine de la relative faiblesse des performances empiriques du modèle HOS. En effet, ce type de flux bilatéraux ne peut être rationalisé dans un modèle de concurrence parfaite, dans lequel les biens sont homogènes donc produits uniquement par l’entreprise représentative la plus compétitive. La réponse théorique à ce puzzle empirique a été, dans les années 80, l’apparition de modèles de commerce en concurrence imparfaite, expliquant l’échange croisé de biens différenciés entre pays similaires en termes de technologie et de dotations factorielles. C’est ce qu’on appelle les **Nouvelles Théories du Commerce International**, initiées en particulier par Paul Krugman (Krugman, 1979, 1980).

Dans ces modèles, l’existence d’une technologie à rendements croissants

TABLE 3.1 – Part du commerce intra-industriel dans le commerce bilatéral (top-10)

Top total IIT shares (per cent)		
Germany	France	86.20
Netherlands	Belgium and Luxembourg	85.01
France	Belgium and Luxembourg	80.42
France	United Kingdom	77.08
Germany	Switzerland	76.99
Germany	Belgium and Luxembourg	76.83
Austria	Germany	76.63
France	Spain	76.55
Germany	Netherlands	76.01
Canada	United States	73.55

Source : Fontagné et al. (2006). Définitions : Le commerce intra-industriel (“Intra-industry trade”) est défini comme l’échange bilatéral simultané de biens dans un sens et dans l’autre, au sein d’une catégorie industrielle fine. La part de ce type de flux dans le commerce bilatéral est ici reportée pour les 10 paires de pays pour lesquelles elle est maximum.

fait de l’échange une source d’efficience. Le commerce international permet aux pays de se spécialiser dans la production d’un nombre limité de biens donc de produire à un coût moyen plus faible. Dans les modèles que nous allons étudier, la spécialisation se fait sur des biens différenciés horizontalement. Les biens produits sont imparfaitement substituables et la préférence des consommateurs pour la diversité explique les échanges croisés de variétés différenciées. Le consommateur retire de cette spécialisation un gain en bien-être. Grâce au commerce international, il a accès à une plus grande diversité de biens.

La popularité de ces modèles à partir de la fin des années 80 tient à son succès empirique. Le modèle de Krugman a des prédictions qui sont cohérentes avec le modèle de gravité présenté en introduction. Dans ce modèle, les grands pays commercent plus les uns avec les autres car ils produisent plus

de variétés et consomment plus de biens étrangers à l'équilibre. En outre, le rôle de la distance peut s'expliquer par l'existence de coûts à l'échange international qui réduisent le volume de biens échangés à l'équilibre. Dans la mesure où l'équation de gravité est elle-même un succès empirique, on comprend la popularité de ces modèles dans les milieux académiques comme pour les questions de politique économique.

3.1 Modèle à un secteur : Le modèle de Krugman

3.1.1 Ingrédients du modèle

Le modèle de Krugman² est basé sur 4 ingrédients fondamentaux. Avant de décrire les hypothèses du modèle, il est utile de discuter la manière dont ces ingrédients interagissent pour expliquer le commerce à l'équilibre. Ces 4 hypothèses fondamentales sont les suivantes :

- **Économies d'échelle** : La technologie du modèle est à rendements croissants, du fait de l'existence d'un coût fixe à la production. Quelle que soit la quantité produite, l'entreprise doit payer un coût fixe uniforme. La présence de ce coût implique que le coût marginal de production est inférieur au coût moyen qui est lui-même décroissant de la quantité produite (puisque le coût fixe est partagé sur un plus grand nombre d'unités). La présence d'économies d'échelle a deux conséquences. D'abord, elle implique que la tarification au coût marginal n'est pas possible, i.e. qu'on sort de l'environnement de concurrence parfaite. Ensuite, elle explique que l'efficacité augmente avec l'échelle de production. C'est ce qui explique qu'il est efficace pour un pays de s'ouvrir au commerce international et de se spécialiser dans la production d'un nombre limité de variétés, produites à plus grande échelle.
- **Concurrence monopolistique** : Dans un cadre avec économies d'échelle, les grandes unités productives ont un avantage concurrentiel sur les petites. Dans ce contexte, les structures de marché tendent vers l'oligopole ou le monopole. L'étude de marchés oligopolistiques est complexe en équilibre général car elle implique de prendre en compte les interactions stratégiques entre agents (i.e. l'internalisation par les individus de leur pouvoir de marché). Depuis le début des années 80, les macroéconomistes ont adopté une hypothèse à la fois simple et élégante permettant de traiter les questions de concurrence imparfaite. Celle-

2. Nous étudions ici une variante du modèle de Krugman (1980).

ci consiste à supposer que l'environnement se caractérise par de la concurrence monopolistique. Dans un environnement de concurrence monopolistique, les entreprises ont un pouvoir de monopole sur la production de leur propre variété. Chaque entreprise produit une variété différenciée, qui n'est qu'imparfaitement substituable aux variétés produites par les autres entreprises. Ce pouvoir de monopole permet à l'entreprise de tarifier au-dessus de son coût marginal (comportement de marge) et donc de supporter le coût fixe de production. Le pouvoir de monopole est cependant limité par la concurrence exercée par les autres entreprises. Dans la mesure où les variétés différenciées sont substituables, la marge de monopole est limitée et les profits du monopole *contestables*. Une entreprise qui choisirait un taux de marge trop élevé se verrait pénalisée par une demande faible, les consommateurs se reportant sur des variétés substituables. Enfin, lorsque l'hypothèse de concurrence monopolistique est combinée à une hypothèse d'entrée libre, les profits sont tirés vers zéro à l'équilibre du marché. Dans cet équilibre de libre-entrée, chaque firme est supposée suffisamment petite pour que les interactions stratégiques soient négligeables. Avec un nombre suffisant de monopoles, l'impact d'une entreprise sur l'équilibre du marché est petit ce qui implique qu'on peut négliger cet effet quand on résout le programme de la firme.

- **Préférences iso-élastiques** : L'hypothèse d'un agrégat de type CES (*Constant Elasticity of Substitution*) est la forme la plus courante utilisée dans les modèles de concurrence monopolistique³. Elle simplifie fortement la résolution du modèle car elle implique que les taux de marge sont constants à l'équilibre. En effet, l'élasticité prix de la demande est constante lorsque les préférences sont CES. A l'optimum, le monopole choisit donc un taux de marge constant. C'est d'ailleurs cette hypothèse qui nous fait préférer l'étude du modèle de Krugman (1980) à celle de Krugman (1979*b*). Le modèle de 1979 est basé sur une hypothèse de préférences additivement séparables dans laquelle les taux de marge ne sont pas constants mais varient avec le nombre de variétés disponibles dans l'économie.
- **Coûts de transport proportionnels** : L'introduction de coûts de transport est, avec l'hypothèse de rendements d'échelle, l'ingrédient clé pour déterminer les flux internationaux de commerce. En effet, sans coût de transport, la localisation de la production devient complètement neutre. Les entreprises peuvent se positionner n'importe où sans

3. On parle aussi de préférences Dixit-Stiglitz en référence à Dixit & Stiglitz (1977) qui ont introduit cette forme fonctionnelle dans un modèle de concurrence monopolistique.

effet sur leur profitabilité. Dans ce cadre, l'ouverture au commerce est isomorphe à une augmentation de la taille du marché. Avec des coûts à l'échange au contraire, le choix du lieu de localisation des entreprises devient un enjeu. A l'équilibre, le *nombre* d'entreprises présentes dans chaque marché dépend alors de la taille de ce marché.

3.1.2 Hypothèses du modèle

Le modèle considère un équilibre à deux pays qui partagent les mêmes préférences et les mêmes technologies de production. Dans le pays domestique (resp. étranger), un nombre L (resp. L^*) d'individus travaillent, consomment et possèdent les entreprises. Les individus sont symétriques et partagent les mêmes préférences. Il est donc équivalent de considérer un ménage représentatif de taille L (L^*). Le travail est le seul facteur de production. Dans ce qui suit, nous résolvons les programmes individuels et l'équilibre général du modèle autarcique pour l'économie domestique. Les résultats sont symétriques pour l'économie étrangère.

Demande : Le ménage représentatif consomme une masse (endogène) n de variétés d'un bien différencié. Les préférences du consommateur vis-à-vis des différentes variétés $\omega \in [0; n]$ sont résumées par la fonction d'utilité suivante :

$$C = \left(\int_0^n q(\omega)^{\frac{\sigma-1}{\sigma}} d\omega \right)^{\frac{\sigma}{\sigma-1}}$$

où $q(\omega)$ est la quantité consommée de la variété ω . La fonction d'utilité est dite à élasticité constante et $\sigma > 1$ mesure l'élasticité de substitution entre les variétés du bien différencié⁴. Plus σ est grand, plus les variétés sont substituables, i.e. plus le consommateur est prêt à substituer une variété pour une autre lorsque le prix relatif de la première augmente. A la limite, quand σ tend vers plus l'infini, les variétés deviennent parfaitement substituables et on retrouve une situation de concurrence parfaite entre les producteurs de variétés.

Le consommateur représentatif maximise son utilité sous la contrainte

4. On appelle élasticité de substitution l'ampleur du report de la demande d'une variété ω vers une autre variété ω' lorsque le prix relatif de ω augmente de 1% :

$$\sigma = \frac{d \ln q(\omega')/q(\omega)}{d \ln p(\omega)/p(\omega')}$$

où $p(\omega)$ est le prix de la variété ω .

budgétaire :

$$\begin{aligned} \max_{\{q(\omega)\}_{\omega \in [0;n]}} & \left(\int_0^n q(\omega)^{\frac{\sigma-1}{\sigma}} d\omega \right)^{\frac{\sigma}{\sigma-1}} \\ \text{s.c.} & \int_0^n p(\omega)q(\omega)d\omega \leq R \end{aligned}$$

où R est le revenu du consommateur représentatif et $p(\omega)$ le prix de la variété ω . La contrainte budgétaire est saturée lorsque la consommation totale en valeur ($\int_0^n p(\omega)q(\omega)d\omega$) est juste égale au revenu du consommateur représentatif (R). Avec un seul facteur de production, celui-ci est composé des revenus du travail (wL avec w le salaire d'équilibre) et des profits résiduels (nuls à l'équilibre) : $R = wL$.

Pour trouver les fonctions de demande individuelles, on écrit le Lagrangien de ce problème :

$$\mathcal{L} = \left(\int_0^n q(\omega)^{\frac{\sigma-1}{\sigma}} d\omega \right)^{\frac{\sigma}{\sigma-1}} - \mu \left(\int_0^n p(\omega)q(\omega)d\omega - R \right)$$

où μ est le multiplicateur de Lagrange. Les conditions du première ordre sont les suivantes :

$$\begin{aligned} 0 &= \frac{\partial \mathcal{L}}{\partial q(\omega)} = q(\omega)^{\frac{-1}{\sigma}} \left(\int_0^n q(\omega)^{\frac{\sigma-1}{\sigma}} d\omega \right)^{\frac{1}{\sigma-1}} - \mu p(\omega), \quad \forall \omega \in [0;n] \\ 0 &= \frac{\partial \mathcal{L}}{\partial \mu} = \int_0^n p(\omega)q(\omega)d\omega - R \end{aligned}$$

En utilisant les conditions du premier ordre relatives à $\omega \in [0;n]$ et en sommant entre ω , on trouve :

$$\begin{aligned} PC &\equiv \int_0^n p(\omega)q(\omega)d\omega = C\mu^{-\sigma} \int_0^n p(\omega)^{1-\sigma} d\omega \\ C &\equiv \left(\int_0^n q(\omega)^{\frac{\sigma-1}{\sigma}} d\omega \right)^{\frac{\sigma}{\sigma-1}} = C\mu^{-\sigma} \left(\int_0^n p(\omega)^{1-\sigma} d\omega \right)^{\frac{\sigma}{\sigma-1}} \end{aligned}$$

où PC est la consommation nominale tandis que C est la consommation réelle. En combinant les deux expressions, on trouve :

$$P = \left(\int_0^n p(\omega)^{1-\sigma} d\omega \right)^{\frac{1}{1-\sigma}}$$

et

$$q(\omega) = \left(\frac{p(\omega)}{P} \right)^{-\sigma} C$$

P est l'indice de prix associé à l'agrégat CES consommé, autrement appelé "indice de prix idéal". Il est défini comme l'agrégat des prix tel que l'utilité du revenu réel (R/P) soit invariante au niveau général des prix : si à la fois le revenu nominal et l'indice de prix augmentent de $x\%$, l'utilité reste inchangée⁵.

La fonction de demande implique que la demande réelle agrégée (C) est partagée entre les différentes variétés en fonction de leur prix relatif ($p(\omega)/P$). Plus le prix relatif de la variété ω est grand, en comparaison de l'indice de prix, plus la demande relative adressée au producteur de cette variété est faible. Enfin, on peut vérifier que l'élasticité-prix de la demande est bien constante ici :

$$\frac{d \ln q(\omega)}{d \ln p(\omega)} = -\sigma$$

Notons finalement que, à l'optimum du consommateur, on a :

$$P = \left(\int_0^n p(\omega)^{1-\sigma} d\omega \right)^{\frac{1}{1-\sigma}} < \int_0^n p(\omega) d\omega$$

L'indice de prix est inférieur au prix moyen dans l'économie. Cela s'explique par une hypothèse de préférence pour la diversité, implicite dans la fonction Dixit-Stiglitz utilisée ici⁶. A prix donnés, lorsque la diversité de l'offre de variétés augmente ($n \uparrow$), l'utilité du consommateur représentatif augmente également. Tout se passe comme si le revenu réel augmentait du fait d'une baisse des prix. Cette préférence pour la diversité est clé pour comprendre les gains à l'échange dans le modèle de Krugman (1980). En permettant aux consommateurs du pays domestique d'avoir accès à la consommation des variétés différenciées produites par les entreprises du pays étranger, il permet une amélioration de l'utilité de la consommation. Les consommateurs français ont un gain en bien-être lié à l'accès aux variétés allemandes de voitures, et inversement.

5. Cette définition implique également que, à revenu nominal donné, on peut étudier l'impact du commerce sur le bien-être en regardant comment le commerce modifie l'indice de prix P .

6. En supposant que toutes les variétés partagent le même prix p et sont donc consommées en quantité identique ($q(\omega) = \frac{R}{np}$), on a :

$$C = \frac{R}{p} n^{\frac{1}{\sigma-1}}$$

qui est croissante de n , la masse de variétés disponibles dans l'économie. C'est ce qu'on appelle la préférence pour la diversité du consommateur. A noter que cette préférence pour la diversité est d'autant plus forte que les biens sont peu substituables (σ petit).

Offre : On suppose qu'il n'existe pas de coût spécifique lié à la création d'une nouvelle variété. Par conséquent, les entreprises profitent de la préférence pour la diversité des consommateurs en choisissant de produire chacune leur propre variété du bien différencié. De cette manière, elles gardent le monopole sur leur propre variété plutôt que d'entrer en concurrence directe avec les producteurs de variétés existantes. À l'équilibre, la masse des variétés offertes aux consommateurs est donc aussi la masse des firmes actives dans l'économie.

Toutes les entreprises partagent la même fonction de production. Pour produire, elles doivent faire face à une fonction de coût qui se décompose en un coût fixe (indépendant de la quantité produite) et un coût variable constant (sur chaque unité produite). L'existence du coût fixe est à l'origine des rendements croissants. Plus spécifiquement, on suppose que, pour produire $q(\omega)$, le producteur de la variété ω doit employer un volume de travail :

$$l(q(\omega)) = f + \frac{q(\omega)}{\varphi}$$

où φ est la productivité du travail, supposée homogène entre firmes et entre pays.

Le programme de la firme consiste à choisir son prix de façon à maximiser son profit étant donné la demande qui lui est adressée :

$$\begin{aligned} \max_{p(\omega)} & \left[p(\omega)q(\omega) - w \left(f + \frac{q(\omega)}{\varphi} \right) \right] \\ \text{s.c.} \quad & q(\omega) = \left(\frac{p(\omega)}{P} \right)^{-\sigma} C \end{aligned}$$

En concurrence monopolistique, l'entreprise est suffisamment petite pour négliger l'impact de son comportement sur les agrégats du marché ($\partial P / \partial p(\omega) = 0$, $\partial C / \partial p(\omega) = 0$ avec un continuum de firmes). La condition du premier ordre de ce programme s'écrit :

$$0 = (1 - \sigma) \left(\frac{p(\omega)}{P} \right)^{-\sigma} C - \frac{w}{\varphi} \frac{-\sigma}{p(\omega)} \left(\frac{p(\omega)}{P} \right)^{-\sigma} C$$

En réarrangeant, on trouve le prix à l'optimum de la firme :

$$p(\omega) = \frac{\sigma}{\sigma - 1} \frac{w}{\varphi}$$

L'entreprise choisit d'appliquer un taux de marge constant ($\frac{\sigma}{\sigma-1}$) sur son coût marginal de production ($\frac{w}{\varphi}$). De cette façon, elle peut couvrir le coût

supplémentaire induit par la présence du coût fixe. Le taux de marge est une fonction décroissante de l'élasticité de substitution entre les biens, qui est aussi une mesure du degré de concurrence entre les producteurs. Plus les variétés différenciées sont substituables (σ grand), plus les consommateurs sont sensibles aux différentiels de prix entre les variétés, plus la pression concurrentielle perçue par les producteurs de variétés est importante, ce qui tend à réduire les taux de marge d'équilibre.

3.1.3 L'équilibre du modèle

L'équilibre du modèle est obtenu en calculant la masse des firmes et le salaire d'équilibre, conditionnellement aux comportements individuels décrits ci-dessus.

L'équilibre autarcique : En intégrant la stratégie optimale de la firme, on peut réécrire le profit de la manière suivante :

$$\pi(\omega) \equiv p(\omega)q(\omega) - w \left(f + \frac{q(\omega)}{\varphi} \right) = w \left(\frac{q(\omega)}{(\sigma - 1)\varphi} - f \right)$$

On détermine ensuite les quantités produites à l'équilibre en utilisant une hypothèse d'entrée libre. S'il n'y a aucun coût à produire une nouvelle variété, des entreprises vont continuer à entrer sur le marché jusqu'à l'épuisement des profits de monopole. L'équilibre est donc atteint pour :

$$\pi(\omega) = 0 \quad \Rightarrow \quad q(\omega) = (\sigma - 1)\varphi f$$

A l'équilibre de libre entrée, toutes les firmes produisent la même quantité au même prix. La quantité d'équilibre est d'autant plus grande que le coût fixe de production f est important. Pour couvrir un coût fixe plus élevé étant donné un taux de marge constant, il faut atteindre une échelle de production suffisante. La quantité d'équilibre est également croissante de l'élasticité de substitution. Pour σ élevé, le taux de marge est relativement petit et la firme doit donc atteindre une échelle de production plus importante pour couvrir le coût fixe. Finalement, $q(\omega)$ est également croissant de φ . Une hausse de la productivité conduit plus d'entreprises à entrer sur le marché ce qui réduit la quantité produite par chaque producteur.

Une fois la quantité d'équilibre déterminée, la masse des firmes actives à l'équilibre est obtenue en utilisant la condition d'équilibre du marché du travail :

$$n \left(f + \frac{q(\omega)}{\varphi} \right) = L \quad \Rightarrow \quad n = \frac{L}{\sigma f}$$

La masse des firmes actives à l'équilibre est une fonction croissante de la taille du marché et décroissante des coûts fixes et de l'élasticité de substitution. Des coûts fixes plus élevés ou une compétition plus intense impliquent que chaque entreprise doit produire plus pour couvrir le coût fixe. Cet équilibre n'est atteint que pour une masse suffisamment petite de producteurs. La taille de la population active permet quant-à-elle à plus d'entreprises d'entrer sur le marché.

Ces variables d'équilibre peuvent être réinjectées dans l'indice de prix idéal de façon à étudier les déterminants du bien-être, dans une économie en autarcie :

$$P = p(\omega)n^{\frac{1}{1-\sigma}} = \frac{\sigma}{\sigma-1} \frac{w}{\varphi} \left(\frac{L}{\sigma f} \right)^{\frac{1}{1-\sigma}}$$

Du fait de la préférence pour la diversité, le bien-être à l'équilibre autarcique est croissant de la taille du pays (i.e. P est décroissant de L) et décroissant de σ et f . La taille du marché permet à un plus grand nombre de variétés de coexister ce qui augmente l'utilité de la consommation. Inversement, des coûts fixes élevés ou une concurrence accrue sont un facteur de désutilité via leur effet sur l'offre de variétés.

L'équilibre en économie ouverte : Considérons maintenant l'intégration commerciale de deux pays de taille L et L^* . Lorsque le commerce entre les deux pays est parfaitement libre et sans coût, l'intégration revient à augmenter la taille du marché. Comme nous l'avons vu dans le paragraphe précédent, une telle intégration est source de gains en bien-être car elle permet aux consommateurs d'avoir accès à plus de variétés ($n + n^*$) et donc d'augmenter la diversité du panier de biens consommé.

Dans ce qui suit, on suppose cependant que le commerce n'est pas parfaitement libre. Plus spécifiquement, on suppose un coût à l'échange de type iceberg et on note $\tau > 1$ la valeur de ce coût. L'hypothèse de coût iceberg est une simplification souvent utilisée dans les modèles de commerce car elle est à la fois facile d'utilisation et relativement réaliste. Comme nous le verrons dans ce qui suit, elle implique qu'à l'équilibre le coût de transport se répercute multiplicativement sur le prix à l'exportation⁷. On dit que le coût de transport est de type iceberg quand une partie du volume exporté "disparaît"

7. Si on associe la politique commerciale à ces coûts à l'exportation, le coût iceberg s'assimile donc à une taxe *ad valorem* sur les importations. Par exemple, une taxe de 20% sur les importations d'un pays correspond à un coût iceberg de 1.2. Lorsque le prix de vente est p , le prix à l'importation payé par les consommateurs étrangers est $1.2p$ et l'effet est donc bien multiplicatif. La différence avec les instruments de politique commerciale est que ce coût "disparaît" dans les modèles avec coûts iceberg tandis que les revenus des taxes augmentent le budget de l'Etat.

pendant le transport vers le pays de destination. Par exemple, l'expédition d'une unité de biens permet au consommateur final d'acheter $1/\tau < 1$ unité (l'iceberg a "fondu" pendant le transport). Du fait de l'existence de ce coût iceberg, les entreprises qui veulent servir le pays étranger avec $q(\omega)$ unités de biens doivent produire $\tau q(\omega)$ unités. Ce coût se répercute sur le coût marginal de production, qui est lui-même un facteur d'inflation des prix.

Pour le montrer, considérons le programme de maximisation d'un producteur offrant sa variété sur le marché domestique et à l'étranger. Son programme de maximisation est le suivant :

$$\begin{aligned} \max_{p^D(\omega), p^X(\omega)} & \left[p^D(\omega)q^D(\omega) + p^X(\omega)q^X(\omega) - w \left(f + \frac{q^D(\omega) + \tau q^X(\omega)}{\varphi} \right) \right] \\ \text{s.c.} \quad & q^D(\omega) = \left(\frac{p^D(\omega)}{P} \right)^{-\sigma} C \\ & q^X(\omega) = \left(\frac{p^X(\omega)}{P^*} \right)^{-\sigma} C^* \end{aligned}$$

L'existence de barrières au commerce est un facteur de segmentation des marchés. Par conséquent, la firme a la possibilité de fixer un prix différent pour son marché domestique ($p^D(\omega)$) et pour l'étranger ($p^X(\omega)$). Elle le fait en tenant compte du coût marginal supplémentaire induit par l'exportation (τ). En résolvant le programme ci-dessous, on montre qu'à l'optimum de la firme, ce coût supplémentaire est entièrement répercuté sur le consommateur final, via le prix⁸ :

$$p^D(\omega) = \frac{\sigma}{\sigma - 1} \frac{w}{\varphi} = p^D \quad \text{et} \quad p^X(\omega) = \frac{\sigma}{\sigma - 1} \frac{\tau w}{\varphi} = \tau p^D$$

Le profit d'une firme domestique dans l'équilibre avec ouverture aux échanges se réécrit alors de la manière suivante :

$$\begin{aligned} \pi(\omega) &= p^D(\omega)(q^D(\omega) + \tau q^X(\omega)) - w \left(f + \frac{q^D(\omega) + \tau q^X(\omega)}{\varphi} \right) \\ &= w \left(\frac{q^D(\omega) + \tau q^X(\omega)}{(\sigma - 1)\varphi} - f \right) \end{aligned}$$

8. Dans la littérature empirique en commerce international, on distingue habituellement le prix FOB (*Franco On Board*), qui correspond au prix sortie d'usine, et le prix CIF (*Cost Insurance and Freight*), qui correspond au prix à l'entrée dans le pays de destination. Ici, le prix à l'exportation $p^X(\omega)$ correspond au prix CIF, c'est celui qui est effectivement payé par le consommateur final et entre donc dans sa fonction de demande. En termes de prix FOB, le prix domestique et le prix à l'exportation sont les mêmes dans ce modèle, puisque le coût marginal (hors coût de transport) et le taux de marge sont identiques quel que soit le marché sur lequel la firme vend ses produits.

A l'équilibre de libre-entrée, la production totale de la firme est égale à :

$$q^D(\omega) + \tau q^X(\omega) = (\sigma - 1)\varphi f$$

Elle est inchangée par rapport à l'équilibre autarcique. En effet, le coût fixe de production ne changeant pas, le profit opérationnel nécessaire à couvrir ce coût fixe est inchangé, et la quantité produite pour atteindre ce profit opérationnel également. On en déduit finalement le nombre de firmes actives à l'équilibre, qui est lui aussi inchangé (puisque l'offre de travail reste la même en l'absence de mobilité internationale du travail) :

$$n = \frac{L}{\sigma f}$$

Ce résultat est important et différent du résultat obtenu par Krugman (1979*b*). Dans le modèle étudié ici, le commerce n'a pas d'effet pro-compétitif puisqu'il ne modifie pas l'élasticité prix de la demande perçue par la firme (qui est constante et égale à σ). Au contraire, dans le modèle de Krugman (1979*b*), l'élasticité de la demande (donc le taux de marge) varie avec le nombre de concurrents auxquels l'entreprise fait face. Lorsque le pays s'ouvre au commerce international, la concurrence accrue liée à la participation des entreprises étrangères au marché domestique renforce la concurrence, conduit les entreprises à réduire leur taux de marge et, à l'équilibre général, permet à un plus petit nombre d'entreprises d'entrer sur le marché ($n \downarrow$). Dans Krugman (1980) au contraire, lorsque l'élasticité de la demande est invariante au degré de compétition, il n'y a pas d'effet du commerce sur la masse de firmes présentes sur le marché.

Cette propriété n'empêche pas le commerce d'engendrer des gains en bien-être pour le consommateur représentatif. En effet, celui-ci a accès à de nouvelles variétés, produites à l'étranger. L'indice de prix dans l'équilibre avec commerce s'écrit donc :

$$\begin{aligned} P &= \left(\int_0^n p^D(\omega)^{1-\sigma} d\omega + \int_0^{n^*} p^{X^*}(\omega)^{1-\sigma} d\omega \right)^{\frac{1}{1-\sigma}} \\ &= \left(n (p^D)^{1-\sigma} + n^* (\tau p^{D^*})^{1-\sigma} \right)^{\frac{1}{1-\sigma}} \end{aligned}$$

où p^{D^*} et p^{X^*} représentent respectivement le prix domestique et le prix à l'export choisi par les entreprises étrangères.

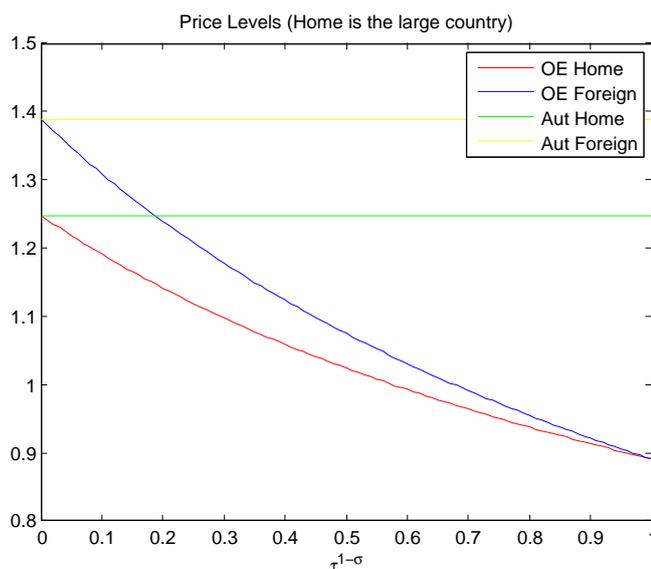
Sans coût à l'exportation et dans un équilibre symétrique ($L = L^* \Rightarrow w = w^* \Rightarrow p^D = p^{D^*}$), l'indice de prix s'écrit :

$$P = P^* = (2np^{D^{1-\sigma}})^{\frac{1}{1-\sigma}} < (np^{D^{1-\sigma}})^{\frac{1}{1-\sigma}}$$

L'ouverture au commerce international conduit donc à des gains à l'échange (à une baisse de P) du fait de l'offre supplémentaire de variétés offerte au consommateur représentatif. En outre, l'indice de prix est égalisé entre les pays, les consommateurs représentatifs achètent les mêmes biens aux mêmes prix.

Dans un cadre plus réaliste avec coût à l'échange, on peut montrer que le commerce réduit également l'indice de prix et donc accroît le bien-être (voir Figure 3.2 qui montre les indices de prix à l'équilibre du modèle en fonction du niveau des coûts à l'échange). Le gain à l'échange est d'autant plus grand que les coûts de transport sont faibles. En effet, pour des coûts de transport élevés, le consommateur profite d'une offre de variétés importante mais ces variétés sont relativement chères, ce qui réduit l'utilité de la consommation.

FIGURE 3.2 – L'indice de prix d'équilibre en autarcie et en économie ouverte



Notes : Ce graphique illustre le niveau de l'indice de prix domestique et étranger, en autarcie et en économie ouverte, en fonction du niveau des coûts à l'échange. Le point d'abscisse 0 correspond à l'autarcie, le point d'abscisse $\tau^{1-\sigma} = 1$ est la limite quand les coûts tendent vers 0 ($\tau \rightarrow 1$). La ligne verte (jaune) correspond à l'autarcie de l'économie domestique (étrangère), utilisée comme référence. On suppose ici que le pays domestique est relativement grand ($L > L^*$). La ligne rouge (bleu) correspond à l'équilibre avec commerce de l'économie domestique (étrangère).

On complète la résolution du modèle avec le salaire relatif des deux pays

en utilisant la condition d'équilibre de la balance commerciale⁹. A l'équilibre de la balance commerciale, la valeur des exportations d'un pays est égale à la valeur de ses importations :

$$np^X q^X = n^* p^{X*} q^{X*}$$

En incorporant les valeurs d'équilibre de n , n^* , p^X , p^{X*} , q^X et q^{X*} , on trouve finalement :

$$\begin{aligned} \frac{w}{w^*} &= \left(\frac{P}{P^*} \right)^{\frac{1-\sigma}{\sigma}} \\ &= \left(\frac{Lw^{1-\sigma} + L^* (\tau w^*)^{1-\sigma}}{L (\tau w)^{1-\sigma} + L^* w^{*1-\sigma}} \right)^{\frac{1}{\sigma}} \end{aligned}$$

Sans coût à l'échange ($\tau \rightarrow 1$), les salaires sont égalisés entre pays. Les consommateurs ont en effet accès aux mêmes variétés, aux mêmes prix. La balance commerciale est donc équilibrée avec des salaires identiques dans les deux pays. Lorsque l'économie tend vers l'autarcie ($\tau \rightarrow +\infty$), le salaire relatif tend vers $\left(\frac{L}{L^*}\right)^{\frac{1}{2\sigma-1}}$. Il est plus élevé dans le grand pays, dans lequel sont produits plus de variétés. Entre les deux, le salaire relatif du grand pays est plus élevé, mais décroissant de la taille des coûts de transport. Avec des coûts de transport, l'indice de prix tend à être plus faible dans le grand pays, qui a accès à un plus grand nombre de variétés relativement bon marché (car non soumises au coût de transport). Cette relative faiblesse de l'indice de prix conduit à une relative faiblesse de la demande d'importations qui, pour que l'équilibre de la balance commerciale soit rétabli, doit être compensée par un coût marginal relativement plus faible à l'étranger ($w/w^* > 1$)¹⁰. Notons pour finir que de telles différences de salaires ne sont possibles que parce que le travail est supposé immobile. Avec mobilité internationale du travail, les travailleurs migreraient vers le grand pays pour profiter de salaires plus

9. Cette stratégie est équivalente à la résolution simultanée de l'équilibre du marché des biens dans les deux pays.

10. Pour le montrer, il est utile de réécrire la demande relative d'importations de l'économie domestique et de l'étranger de la manière suivante :

$$\frac{n^* p^{X*} q^{X*}}{np^X q^X} = \frac{n^*}{n} \left(\frac{\tau p^{D*}/P^*}{\tau p^D/P^*} \right)^{1-\sigma} \frac{wL}{w^*L^*} = \left(\frac{w^*/P^*}{w/P^*} \right)^{1-\sigma} \frac{w}{w^*}$$

Une augmentation de L/L^* conduit à une baisse de P/P^* via l'offre de variétés domestiques. Pour revenir à l'équilibre de la balance commerciale, il faut donc que w/w^* augmente de façon à relancer la demande relative d'importations via un effet de demande agrégée (hausse de w/w^*) et un effet de substitution (baisse de $\frac{w^*/P^*}{w/P^*}$).

élevés. Le grand pays concentrerait donc une part disproportionnée de la production et du travail. Ce type d'ajustements est à la base de la "Nouvelle Economie Géographique".

3.2 Spécialisation : Le modèle de Helpman-Krugman

Jusqu'ici, nous avons présenté un modèle de commerce complètement déconnecté des modèles présentés dans le chapitre précédent. En particulier, l'hypothèse d'un unique secteur producteur d'un unique bien différencié ne permettait pas d'introduire de phénomènes de spécialisation. Dans cette section, nous étendons donc le cadre d'analyse à deux secteurs productifs, de façon à étudier la manière dont le commerce intra-industriel (l'échange de variétés différenciées d'un même bien) peut affecter la spécialisation intersectorielle des pays.

3.2.1 Hypothèses du modèle

Le modèle est constitué de deux pays, deux secteurs (X et Y) et un facteur de production (le travail). Les technologies comme les préférences sont identiques dans les deux pays. La seule chose qui différencie l'économie domestique de l'économie étrangère concerne leur taille, notées L et L^* .

Pour simplifier l'analyse, on suppose que les préférences des agents vis-à-vis des biens produits dans les deux secteurs sont de type Cobb-Douglas. Pour le pays domestique :

$$C = C_X^\mu C_Y^{1-\mu}$$

où C est l'utilité de la consommation agrégée, et C_X et C_Y les consommations sectorielles. Avec des préférences Cobb-Douglas, le consommateur représentatif alloue une part constante (μ) de son revenu à la consommation du bien X .

Le secteur X produit un bien différencié dans les mêmes conditions que celles décrites dans la section précédente. Les consommateurs ont des préférences CES vis-à-vis des variétés différenciées offertes sur le marché. La production de ces variétés est à rendements croissants du fait de l'existence d'un coût fixe de production, qui est couvert par les firmes grâce à une tarification au-dessus du coût marginal. Les entreprises vendent leur variété différenciée sur leur propre marché sans coût supplémentaire et à l'étranger, en faisant face à un coût iceberg. Par conséquent, toutes les entreprises d'un même marché font face à la même demande, et vendent au même prix. Par

exemple, la production d'une entreprise domestique dans le secteur X s'écrit :

$$q(\omega) = q^D(\omega) + \tau q^X(\omega) = \left(\frac{p^D(\omega)}{P} \right)^{-\sigma} \frac{\mu w L}{P} + \tau \left(\frac{\tau p^D(\omega)}{P^*} \right)^{-\sigma} \frac{\mu w^* L^*}{P^*}$$

avec $p^D = \frac{\sigma}{\sigma-1} \frac{w}{\varphi}$ le prix de vente sur le marché domestique.

Le secteur Y produit de son côté un bien homogène, en concurrence parfaite et avec une technologie linéaire dans le travail :

$$Y = L_Y$$

Il est vendu à l'étranger sans coût. On suppose dans la suite que ce bien est produit dans les deux pays, même à l'équilibre en économie ouverte (hypothèse de spécialisation incomplète). Ces hypothèses extrêmes permettent de simplifier la résolution du modèle. En particulier, l'hypothèse de concurrence parfaite implique que le prix de vente du bien Y , à l'étranger comme dans le pays domestique, est égal au coût marginal de production, i.e. le salaire unitaire ($P_Y = w$, $P_Y^* = w^*$). L'hypothèse d'absence de coût de transport implique que le prix de vente de ce bien est égalisé entre les pays ($P_Y = P_Y^* = w = w^*$). En choisissant le bien homogène domestique comme numéraire, on en déduit que le salaire est égalisé entre les pays et égal à 1. Finalement, en supposant la mobilité intersectorielle du travail, on en déduit le salaire s'appliquant aux entreprises du secteur de concurrence imparfaite, égal à 1 dans les deux pays. Les hypothèses extrêmes imposées au secteur producteur du bien homogène permettent donc de s'abstraire de la partie difficile de la résolution du modèle en équilibre général, la détermination du salaire relatif. A partir du moment où un bien homogène est échangé librement entre les deux pays, le salaire relatif ne peut qu'être égal à un. Le salaire relatif étant déterminé, on peut se concentrer sur l'allocation internationale de la production, en économie ouverte.

3.2.2 L'équilibre en économie ouverte

A l'équilibre avec entrée libre, la quantité produite par chacune des entreprises du secteur différencié est égale à $(\sigma - 1)\varphi f$. Comme les technologies sont identiques dans les deux pays, les entreprises y produisent la même quantité de biens :

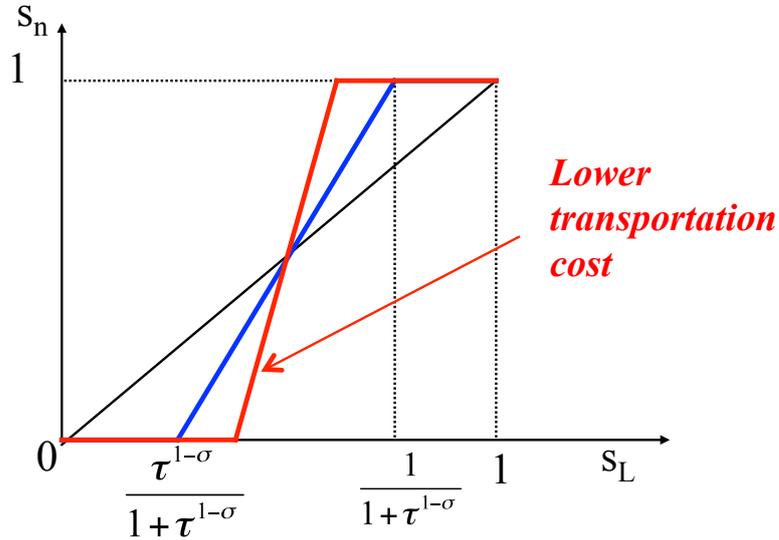
$$\begin{aligned} q(\omega) &= q^*(\omega) \\ \Leftrightarrow \frac{L}{n + n^* \tau^{1-\sigma}} + \frac{\tau^{1-\sigma} L^*}{n^* + n \tau^{1-\sigma}} &= \frac{L^*}{n^* + n \tau^{1-\sigma}} + \frac{\tau^{1-\sigma} L}{n + n^* \tau^{1-\sigma}} \\ \Leftrightarrow n(L^* - \tau^{1-\sigma} L) &= n^*(L - \tau^{1-\sigma} L^*) \end{aligned}$$

Cette équation permet de qualifier l'allocation internationale des producteurs de biens différenciés ($s_n \equiv \frac{n}{n+n^*}$) en fonction de l'allocation internationale de la demande ($s_L \equiv \frac{L}{L+L^*}$) :

$$s_n = \begin{cases} 0, & s_L \leq \frac{\phi}{1+\phi} \\ \frac{s_L(1+\phi)-\phi}{1-\phi}, & s_L \in \left[\frac{\phi}{1+\phi}; \frac{1}{1+\phi} \right] \\ 1, & s_L \geq \frac{1}{1+\phi} \end{cases}$$

où $\phi \equiv \tau^{1-\sigma} \in [0, 1]$ est inversement corrélé à la taille des coûts de transport.

FIGURE 3.3 – Spécialisation dans le secteur du bien différencié



Notes : Ce graphique représente la part des entreprises localisées dans l'économie domestique (s_n en ordonnée) en fonction de la part de la population active localisée dans cette même économie (s_L en abscisse). La ligne en rouge représente un niveau de coût de transport plus faible que sur la ligne en bleu.

Lorsque les deux pays sont de taille très différente ($s_L \leq \frac{\phi}{1+\phi}$ ou $s_L \geq \frac{1}{1+\phi}$), la spécialisation est complète dans le secteur du bien différencié. Le plus grand pays concentre l'intégralité de la production tandis que le "petit" pays ne produit que du bien homogène. Pour des tailles de pays suffisamment proches, des variétés différenciées sont produites dans les deux pays mais le grand pays continue à représenter une part disproportionnée de l'offre de variétés :

$$s_n = s_L + \frac{1}{2} \frac{\phi}{1-\phi} \left(s_L - \frac{1}{2} \right) > s_L \quad \text{si} \quad s_L > \frac{1}{2}$$

En outre, la part du grand pays dans l'offre de variétés croît plus que proportionnellement à sa taille ($\frac{ds_n}{ds_L} > 1$). C'est ce qu'on appelle le "Home Market Effect". Dans un modèle à deux secteurs, le grand pays a un "avantage comparatif" dans la production de biens différenciés, de par sa taille. En effet, la taille du marché domestique permet aux entreprises d'amortir le coût fixe de production, sans perte de compétitivité liée aux coûts de transport. C'est ce qui fait qu'un nombre relativement plus important d'entreprises entre sur le marché du grand pays. Dans le modèle de Helpman-Krugman, cet effet est amplifié par l'ajustement possible du marché du travail via la production de bien homogène. Le grand pays peut allouer une part disproportionnée de son travail à la production de biens différenciés et "acheter" le travail étranger utilisé pour produire du bien homogène. L'échange de biens différenciés contre du bien homogène permet donc de profiter au maximum des économies d'échelle en permettant à plus d'entreprises de se localiser sur le grand marché.

Enfin, il faut noter que cet effet de spécialisation est renforcé par une baisse des coûts de transport. Plus les coûts de transport sont faibles, plus l'intervalle de tailles relatives dans lequel il n'y a pas spécialisation complète dans le secteur du bien différencié est petit, et plus la spécialisation relative du grand pays est importante. C'est ce qui est illustré sur la Figure 3.3.

Ce modèle permet donc d'introduire une nouvelle source de spécialisation des pays dans la production de différents biens (spécialisation intersectorielle). Lorsqu'une partie de la production est à rendements croissants, les grands pays (en termes *absolus*) ont un avantage comparatif. En effet, la taille du marché domestique permet aux producteurs de biens différenciés d'amortir le coût fixe de production auprès de consommateurs qui ne sont pas affectés par la perte de compétitivité induite par les coûts à l'échange international. Ces structures de marché à rendements constants sont également invoquées par les économistes qui prônent l'utilisation de politiques protectionnistes pour permettre aux industries à rendements croissants d'atteindre une échelle de production suffisante avant d'être confrontées à la concurrence internationale (argument de "protection des industries naissantes").

3.3 Evidences empiriques : L'équation de gravité

¹¹Le succès des modèles de concurrence imparfaite tient en partie au fait qu'ils sont compatibles avec l'équation de gravité ¹². Comme nous l'avons évoqué en introduction, l'équation de gravité est d'abord une relation empirique, qui a prouvé une certaine efficacité pour prédire la valeur du commerce bilatéral. Le fait que les nouvelles théories du commerce international soient compatibles avec le modèle de gravité en fait donc *de facto* un succès empirique.

3.3.1 Microfondation de l'équation de gravité

Le modèle de Krugman offre une interprétation structurelle au modèle gravitaire. Pour le montrer, revenons au cadre d'analyse de la section 3.1, que l'on peut généraliser, en équilibre partiel, à un cadre multi-pays. Supposons que le monde est constitué de C pays indicés $i = 1 \dots C$. Chaque pays est constitué d'une masse n_i d'entreprises produisant le bien différencié dans les conditions décrites en section 3.1. L'utilité du consommateur représentatif dans le pays j s'écrit :

$$\begin{aligned} C_j &= \left(\sum_{i=1}^C \int_0^{n_i} q_{ij}(\omega)^{\frac{\sigma-1}{\sigma}} d\omega \right)^{\frac{\sigma}{\sigma-1}} \\ &= \left(\sum_{i=1}^C n_i q_{ij}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}} \end{aligned}$$

où on utilise dans la deuxième ligne l'hypothèse d'homogénéité des entreprises au sein d'une économie donnée ($q_{ij}(\omega) = q_{ij}$, $\forall \omega \in [0, n_i]$).

La demande adressée par le consommateur représentatif de j à l'ensemble des entreprises d'un pays i s'écrit :

$$\begin{aligned} X_{ij} &= n_i p_{ij} q_{ij} \\ &= n_i \left(\frac{\sigma}{\sigma-1} \frac{\tau_{ij} w_i}{\varphi_i P_j} \right)^{1-\sigma} R_j \end{aligned}$$

11. Pour une revue de littérature exhaustive sur le sujet, voir Head & Mayer (2014).

12. L'utilisation de l'équation de gravité est aujourd'hui complètement généralisée en commerce international. Arkolakis, Costinot & Rodriguez-Clare (2012) décrivent une classe de modèles qui conduisent à une prédiction du commerce de type gravitaire. Cette classe s'étend bien au-delà des modèles de commerce en concurrence imparfaite décrits dans ce chapitre. Elle inclut notamment le modèle de commerce de type ricardien proposé par Eaton & Kortum (2002).

où τ_{ij} mesure la taille des coûts de transport entre i et j , $\frac{w_i}{\varphi_i}$ le coût marginal de production dans le pays i , P_j l'indice de prix idéal dans le pays j et R_j son revenu nominal. Par définition :

$$P_j = \left(\sum_{i=1}^C n_i p_{ij}^{1-\sigma} \right)^{\frac{1}{1-\sigma}}$$

En réarrangeant les termes et en log-linéarisant, on trouve une équation expliquant le commerce bilatéral par une forme qui s'apparente à un modèle de gravité :

$$\ln X_{ij} = \underbrace{\ln \left(\frac{\sigma}{\sigma - 1} \right)^{1-\sigma}}_{\text{constante}} + \underbrace{\ln n_i + (1 - \sigma) \ln \frac{w_i}{\varphi_i}}_{i\text{-specific}} + \underbrace{\ln P_j^{\sigma-1} + \ln R_j}_{j\text{-specific}} + \underbrace{(1 - \sigma) \ln \tau_{ij}}_{\text{cout de transport}}$$

Le commerce bilatéral s'explique par :

- une constante,
- un terme spécifique au pays exportateur qui est croissant dans son offre de variétés et décroissant dans son coût marginal de production. Si on ignore le coût de production et en utilisant le résultat selon lequel la masse de variétés produites à l'équilibre est croissante de la taille du pays ($n_i = L_i/(\sigma f)$), on en déduit une relation entre le commerce bilatéral et la taille du pays exportateur.
- un terme spécifique au pays importateur qui croît avec son revenu agrégé et avec le niveau de son indice de prix. Ici aussi, en ignorant l'indice de prix agrégé, on retrouve la relation évoquée en introduction entre le commerce bilatéral et la taille du pays importateur.
- un terme "bilatéral" négativement corrélé aux coûts de transport bilatéraux. En approximant les coûts de transport par une fonction exponentielle de la distance ($\tau_{ij} = dist_{ij}^\delta$ où δ est l'élasticité des coûts de transport à la distance), on retrouve la relation entre commerce et distance.

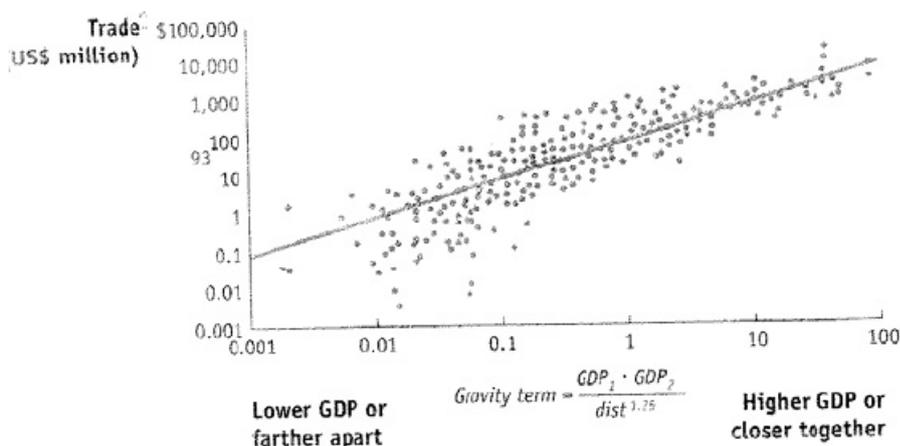
En première approximation, le modèle de Krugman offre donc un cadre analytique qui peut être testé empiriquement en utilisant l'équation suivante :

$$\ln X_{ij} = \alpha \ln L_i + \beta \ln L_j + \gamma \ln dist_{ij} + \varepsilon_{ij}$$

où ε_{ij} est un terme d'erreur, supposé orthogonal aux variables explicatives de l'équation.

Une représentation graphique des résultats d'une telle estimation est reproduite sur la Figure 3.4, issue de Feenstra & Taylor (2011) et relative au commerce entre les Etats-Unis et le Canada. Après estimation, les auteurs

FIGURE 3.4 – L'équation de gravité : Commerce entre les Etats américains et les provinces canadiennes



Source : Feenstra & Taylor (2011).

imposent les valeurs estimées aux coefficients de l'équation, respectivement $\alpha = 1$, $\beta = 1$ et $\gamma = -1.25$ puis comparent les flux de commerce observés dans les données (axe des ordonnées) à ceux prédits par le modèle (axe des abscisses). La corrélation est effectivement très positive et le nuage de points relativement resserré autour de la droite de régression. La valeur du commerce est plus élevée entre les gros Etats américains et les grosses provinces canadiennes et ce d'autant plus que ces zones sont proches géographiquement.

3.3.2 Estimation

Le tableau 3.2 reproduit les résultats de l'estimation d'une équation de gravité, sur des données de commerce agrégé relatives à 208 pays¹³. Pour étudier la stabilité temporelle de l'équation de gravité, les mêmes régressions sont reproduites pour l'année 1970 (colonnes (1)-(3)) et pour 2006 (colonnes (4)-(6)). Les colonnes (1) et (3) reproduisent les résultats de l'estimation la plus simple, où les variables L_i et L_j sont approximées par la population et le PIB par tête du pays tandis que les coûts de transport sont mesurés par la distance entre les pays. Les coefficients sont tous statistiquement significatifs et conformes au modèle structurel. L'effet de la taille du pays est proche de 1, un peu plus élevé dans les données de 2006 pour le pays exportateur. Par

13. Les données sont celles utilisées par Head, Mayer & Ries (2010) et mises à disposition sur le site de Keith Head.

rapport à une mesure de taille économique (ici le PIB), l'utilisation de la taille de la population ne change pas grand-chose puisque la différence entre le coefficient estimé sur la taille et le coefficient estimé sur le PIB par tête est assez proche de 0. Enfin, l'effet de la distance est négatif, très significatif et élevé. Quantitativement, il implique qu'un doublement de la distance entre deux pays partenaires réduit leur commerce de moitié. L'importance de l'effet de la distance permet de mieux comprendre la forte concentration géographique des flux de commerce discutée en introduction de ce cours.

La spécification de l'équation de gravité dans laquelle les coûts à l'échange sont approximatés uniquement par la distance bilatérale se révèle utile pour expliquer la valeur du commerce dans des échantillons de partenaires relativement homogènes, comme c'est le cas par exemple du commerce entre les Etats-Unis et le Canada utilisé pour construire la figure 3.4. Lorsque l'équation de gravité est estimée sur des échantillons de pays plus hétérogènes, il est utile d'ajouter un certain nombre de variables de contrôles permettant de comprendre pourquoi certaines paires de pays échangent plus que ce qui est prédit par une équation de gravité simple. Dans ce cas, l'équation estimée prend la forme suivante :

$$\ln X_{ij} = \alpha \ln L_i + \beta \ln L_j + \gamma \ln dist_{ij} + \chi Controls_{ij} + \varepsilon_{ij}$$

où $Controls_{ij}$ est un ensemble de variables de contrôle et χ un vecteur de coefficients à estimer. Typiquement, le vecteur de contrôles inclut des variables de type culturel (partage d'une même langue, racines coloniales communes), des variables géographiques autres que la distance (partage d'une frontière commune par exemple) et des variables relatives à l'intégration économique des deux pays (appartenance à une même zone de libre-échange, à une union douanière, à une union monétaire)¹⁴. Ce type de variables de contrôle est ajouté dans les régressions des colonnes (2-3-5-6) de la table 3.2.

Comme on peut le constater en comparant les statistiques de R^2 relatives à ces différentes régressions, l'utilisation de variables de contrôle supplémentaires permet d'augmenter la qualité prédictive du modèle. En outre, elle réduit l'effet estimé de la distance, ce qui suggère qu'une partie de l'effet obtenu dans les colonnes (1) et (3) tient à la colinéarité de cette variable

14. Ces variables sont souvent mesurées par des indicateurs booléens. Dans ce cas, l'interprétation des coefficients estimés est directe puisque :

$$\chi_1 = \ln(X_{ij}|Controls_{ij1} = 1) - \ln(X_{ij}|Controls_{ij1} = 0)$$

Toutes choses égales par ailleurs, l'exponentielle du coefficient estimé mesure le supplément (ou la baisse si χ_1 est négatif) de commerce entre deux pays i et j induite par la variable de contrôle $Controls_{ij1}$.

TABLE 3.2 – Différentes spécifications de l'équation de gravité

	Variable dépendante : $\ln X_{ij}$					
	(1)	(2)	(3)	(4)	(5)	(6)
ln Population i	0.799 ^a (.013)	0.823 ^a (.013)		1.185 ^a (.011)	1.191 ^a (.011)	
ln PIB par tête i	1.072 ^a (.016)	1.110 ^a (.017)		1.272 ^a (.012)	1.265 ^a (.012)	
ln Population j	0.723 ^a (.013)	0.740 ^a (.013)		0.896 ^a (.011)	0.900 ^a (.010)	
ln PIB par tête j	1.058 ^a (.017)	1.092 ^a (.018)		0.920 ^a (.012)	0.912 ^a (.012)	
ln Distance	-1.008 ^a (.024)	-0.838 ^a (.026)	-1.000 ^a (.031)	-1.511 ^a (.022)	-1.199 ^a (.029)	-1.619 ^a (.033)
Accord commercial		0.917 ^a (.147)	0.643 ^a (.180)		0.758 ^a (.063)	0.493 ^a (.062)
GATT/OMC		-0.011 (.109)	0.038 (.086)		0.306 ^a (.047)	0.811 ^a (.164)
Monnaie commune		1.470 ^a (.159)	1.460 ^a (.151)		-0.029 (.151)	0.035 (.150)
Frontière commune		0.588 ^a (.109)	0.533 ^a (.118)		1.152 ^a (.108)	0.840 ^a (.113)
Langue commune		0.559 ^a (.057)	0.535 ^a (.060)		1.108 ^a (.056)	0.909 ^a (.060)
Passé colonial		1.376 ^a (.105)	1.277 ^a (.100)		0.672 ^a (.097)	0.889 ^a (.096)
Année	1970	1970	1970	2006	2006	2006
Effets fixes	No	No	Yes	No	No	Yes
# observations	9,035	9,035	9,035	16,936	16,936	16,936
R ²	0.583	0.607	0.710	0.631	0.649	0.741

Notes : Ecart-type robuste entre parenthèses avec ^a, ^b et ^c indiquant une significativité statistique aux seuils respectifs de 1, 5 et 10%. Source : Calculs de l'auteur à partir des données de Head et al. (2010).

avec d'autres mesures de barrières aux échanges. Enfin, elle permet de mieux comprendre les éléments concourant à segmenter les marchés internationaux de biens et services. Ces éléments sont de plusieurs natures. Ils sont d'abord d'ordre tarifaire. Ainsi, les pays participant à un accord régional de libre-échange et, dans une moindre mesure, les pays participant aux négociations internationales sur le libre-échange dans le cadre de l'Organisation Mondiale du Commerce (OMC, anciennement GATT), ont tendance à commercer plus. Ce résultat est cohérent avec l'idée selon laquelle les politiques commerciales protectionnistes ont pour effet de réduire l'ampleur des échanges internationaux.

Ces estimations montrent cependant que les barrières tarifaires ne sont pas la seule source de segmentation internationale des marchés. Ainsi, le partage d'une monnaie commune est, au moins dans les données de 1970, une source de renforcement des échanges commerciaux. L'intensité des échanges commerciaux semble aussi influencée par des éléments de type culturel. Ainsi, le partage d'une langue commune facilite les transactions commerciales. Même en 2006, 40 ans après la dissolution des grands empires coloniaux, les anciennes puissances coloniales continuent à commercer plus avec leurs anciennes colonies qu'avec d'autres pays comparables en termes de taille, de distance, etc. Enfin, les barrières sont aussi géographiques. Des pays mitoyens commercent plus les uns avec les autres, indépendamment de la distance entre ces deux pays. Nous reviendrons sur ces éléments dans le paragraphe suivant, dans lequel nous discuterons un certain nombre de questions relatives à l'intégration internationale, qui ont été traitées en utilisant le cadre gravitaire.

Avant cela, citons une dernière question méthodologique qui a animé les débats concernant l'estimation de l'équation de gravité. Cette question concerne l'approximation des déterminants spécifiques à chaque pays (exportateur ou importateur) par des variables de taille (PIB par tête et population dans les colonnes (1), (2), (4) et (5) de la table 3.2). Si on se réfère à l'équation structurelle sous-jacente, dérivée dans le paragraphe précédent, une telle approximation revient à négliger certains aspects essentiels du modèle, notamment les écarts de compétitivité entre pays, résumés dans les prix à l'exportation (w_i/φ_i) et le niveau de la concurrence dans le pays de destination, corrélé à l'indice de prix idéal ($P_j^{1-\sigma}$). Négliger ces deux déterminants du commerce bilatéral peut conduire à des "erreurs" de prédiction. Pour le montrer, comparons par exemple deux pays de taille similaire comme les Pays-Bas et l'Australie. En 2006, le PIB de l'Australie est 20% plus élevé que celui des Pays-Bas. Pourtant, la valeur des exportations comme des importations de ce pays est moins de un tiers celle des Pays-Bas. La position géographique de l'Australie fait que ce pays est très désavantagé sur les marchés internationaux, par des coûts de transport élevés. Son taux d'ouverture

au commerce international est faible par conséquent mais, conditionnellement à échanger, le pays va interagir avec des pays qui sont très éloignés géographiquement. A l'inverse, la position centrale des Pays-Bas dans l'Union Européenne leur confère un accès à un marché important à une distance géographique faible. Cette situation géographique en fait un pays très ouvert au commerce international qui commerce principalement avec des pays proches géographiquement. Ces deux pays ont donc des positions très dissemblables dans le système de commerce international. Pourtant, l'équation de gravité simplifiée les considère comme comparables, puisque similaires en termes de taille.

Dans la mesure où ces aspects peuvent conduire à une mauvaise estimation des coefficients de l'équation de gravité, il est important d'en tenir compte en utilisant une spécification plus structurelle. C'est ce qui est fait dans les colonnes (3) et (6) du tableau 3.2. Plutôt que de chercher à mesurer les variables structurelles entrant dans l'équation gravitaire, ce qui implique de tester un type de modèles gravitaires spécifiquement, l'approche retenue par la littérature consiste à utiliser la dimension en panel des données sous-jacentes, et à utiliser des effets fixes pour chaque pays exportateur et chaque pays importateur. Ces effets fixes capturent l'impact de toute variable spécifique à un des pays partenaires et influençant le niveau de son commerce. Ils sont donc colinéaires à la taille du pays, qui disparaît de la liste des variables explicatives dans les régressions de la table 3.2. Mais ils tiennent également compte de toute autre variable concourant à la compétitivité relative d'un pays exportateur ou au niveau de la demande d'un pays importateur. Un certain nombre de ces variables étant inobservables, il est nécessairement plus contraignant d'utiliser une telle méthode d'effets fixes que toute autre spécification, aussi riche soit-elle. C'est ce qui explique l'augmentation du pouvoir explicatif de ces régressions (cf comparaison des R^2 des colonnes (2) et (3) ou des colonnes (5) et (6)).

La comparaison des résultats obtenus avec cette spécification et avec la spécification utilisant le PIB par tête et la taille de la population comme contrôles montre que certaines des variables omises dans la spécification des colonnes (2) et (5) sont corrélées à la taille des barrières aux échanges. En particulier, l'effet estimé de la distance est plus important, une fois que les effets fixes sont introduits. Avec les données de 2006, l'élasticité du commerce à la distance est estimée à -1.6, ce qui signifie qu'un doublement de la distance réduit la valeur du commerce par trois !

Aujourd'hui, la littérature empirique sur le sujet s'accorde sur le fait qu'une spécification à effets fixes est préférable à toute autre approximation. La spécification des colonnes (3) et (6) est donc très largement retenue.

3.3.3 Les enseignements de l'équation de gravité

L'importance des questions méthodologiques dans la littérature sur l'équation de gravité s'explique par le rôle qu'a joué cette équation dans les débats de politique économique. A partir du moment où les résultats de l'estimation de cette équation sont utilisés pour justifier telle ou telle politique économique, il est important de s'assurer que les coefficients estimés sont sans biais et correspondent donc à la réalité des données. Ce n'est qu'une fois qu'on s'est assuré de la stabilité de la relation qu'on peut en interpréter les résultats.

Dans ce qui suit, nous résumons les résultats principaux dérivés de l'estimation de l'équation de gravité. La discussion s'appuyera sur nos propres estimations (Table 3.2) et sur les résultats d'une méta-analyse proposée par Disdier & Head (2008) et étendue par Head & Mayer (2014). Cette méta-analyse porte sur 2508 régressions tirées de 159 articles publiés dans des revues scientifiques. Head & Mayer (2014) distinguent deux types d'estimation, les estimations standards et les estimations dites structurelles, qui utilisent une méthode d'effets fixes comme celle décrite dans le paragraphe précédent. Des statistiques descriptives relatives aux différents coefficients estimés sont reportées dans le tableau 3.3, pour l'ensemble de l'échantillon dans la partie de gauche, puis pour l'échantillon restreint aux estimations de type structurel dans la partie de droite.

Rôle de la distance : Un des résultats les plus robustes de la littérature gravitaire concerne le rôle de la distance sur les flux bilatéraux de commerce. Dans la méta-analyse de Head & Mayer (2014), l'élasticité moyenne du commerce à la distance est de 0.93 (1.1 quand on ne tient compte que des estimations de type structurel). Cette élasticité est proche de celle estimée dans le tableau 3.2, pour 1970. Un doublement de la distance entre deux pays partenaires conduit à une réduction de moitié de la valeur de leurs échanges, toutes choses égales par ailleurs.

L'interprétation de cette élasticité est liée à la manière dont les coûts de transport augmentent avec la distance. Hummels (1999) estime ainsi que l'élasticité des coûts de transport à la distance (le coefficient δ dans l'approximation $\tau_{ij} = dist_{ij}^{\delta}$) est de .22 à .46 en fonction du mode de transport. Un doublement de la distance implique une augmentation du coût de transport de 15 à 40%. Au-delà de l'effet direct sur les coûts de transport, la distance à la destination est corrélée avec le temps de transport, qui est également une barrière importante à l'échange. Ainsi, Hummels (2001) estime qu'un temps de transport de 20 jours est équivalent à un tarif de 16% sur l'importation du bien. Enfin, la distance est également corrélée à la distance culturelle entre pays, qui influence l'ampleur du commerce. Ce type de déterminants est évi-

TABLE 3.3 – Résultats d’une méta-analyse sur l’estimation de l’équation de gravité

Estimates:	All Gravity				Structural Gravity			
	median	mean	s.d.	#	median	mean	s.d.	#
Origin GDP	.97	.98	.42	700	.86	.74	.45	31
Destination GDP	.85	.84	.28	671	.67	.58	.41	29
Distance	-.89	-.93	.4	1835	-1.14	-1.1	.41	328
Contiguity	.49	.53	.57	1066	.52	.66	.65	266
Common language	.49	.54	.44	680	.33	.39	.29	205
Colonial link	.91	.92	.61	147	.84	.75	.49	60
RTA/FTA	.47	.59	.5	257	.28	.36	.42	108
EU	.23	.14	.56	329	.19	.16	.5	26
CUSA/NAFTA	.39	.43	.67	94	.53	.76	.64	17
Common currency	.87	.79	.48	104	.98	.86	.39	37
Home	1.93	1.96	1.28	279	1.55	1.9	1.68	71

Notes: The number of estimates is 2508, obtained from 159 papers. Structural gravity refers here to some use of country fixed effects or ratio-type method.

Source : Head & Mayer (2014)

demment difficile à mesurer mais certaines évidences indirectes montrent que les facteurs culturels participent au développement du commerce international. Par exemple, Rauch (2001) met en évidence le rôle des relations sociales en utilisant des données sur les migrations internationales. Il montre que la présence d'une communauté suffisante de migrants issus d'un pays particulier augmente l'intensité des échanges entre ce pays et le pays d'accueil.

FIGURE 3.5 – Estimation de l'élasticité à la distance, au cours du temps



Source : Calculs de l'auteur à partir des données de Head et al. (2010). La ligne pleine correspond à la valeur absolue du coefficient estimé sur la distance dans l'équation :

$$\ln X_{ij} = FE_i + FE_j + \gamma \ln dist_{ij} + \chi Controls_{ij} + \varepsilon_{ij}$$

Les lignes en pointillés représentent l'intervalle de confiance au seuil de significativité de 5%.

Aussi étonnant que cela puisse paraître, le rôle de la distance n'a pas tendance à diminuer au cours du temps, malgré les progrès technologiques facilitant le transport de biens et services. C'est ce que montre notamment la méta-analyse de Disdier & Head (2008). Toutes choses égales par ailleurs, l'élasticité du commerce à la distance a tendance à augmenter au cours du temps. Ce résultat est reproduit sur la Figure 3.5 qui montre l'évolution de l'élasticité estimée du commerce à la distance (i.e. $|\hat{\gamma}|$) où le coefficient γ est estimé en utilisant la spécification des colonnes (3) et (6) de la Table 3.2 pour chaque année de 1948 à 2006). A part au tout début de l'échantillon, i.e. dans

l'immédiat après-guerre, l'élasticité du commerce à la distance augmente sur la période observée, passant de 0.8 en début de période à 1.6 en 2006. Ce doublement de l'effet distance a des implications importantes puisqu'il implique une concentration géographique des flux internationaux de commerce international au cours de la seconde moitié du 20ème siècle.

Ce résultat reste un puzzle que la littérature en commerce international peine à expliquer. Plusieurs explications ont été avancées, sans que la littérature ne tranche définitivement en faveur de l'une ou de l'autre. Il semble cependant que les changements technologiques de la deuxième moitié du 20ème siècle, qui étaient supposés réduire fortement les barrières à l'échange, n'ont pas eu un effet très prononcé. Ce résultat tient probablement à la nature de ces changements, qui ont surtout réduit les coûts de télécommunications, et pas tellement les coûts induits par le transport physique des produits (Hummels, 2007). Cette stagnation des coûts de l'échange international n'est pas suffisante cependant pour expliquer une *augmentation* de l'effet distance. Hummels (2001) et Deardorff (2003) suggèrent que le commerce international est devenu plus sensible au temps consacré au transport des marchandises, ce qui pénalise les flux sur de longues distances. En particulier, la réorganisation des structures productives en faveur de modes de production à flux tendus, dans lesquels les stocks sont limités au maximum pour pouvoir répondre au plus vite à l'évolution de la demande, fait de la logistique une activité clé des systèmes productifs, qui dépend crucialement de la vitesse avec laquelle les unités productives peuvent se réapprovisionner. Dans ce contexte, la distance aux fournisseurs est extrêmement pénalisante.

Effet de l'intégration commerciale : De longue date, les économistes ont utilisé l'équation de gravité pour estimer, et tenter de prédire, l'effet sur le commerce des accords régionaux de libre-échange. Dans la Table 3.3, les lignes "RTA/FTA", "EU", "CUSA/NAFTA" et "Common currency" donnent une idée de l'ampleur des effets estimés. La ligne "RTA/FTA" mesure l'effet sur le commerce des accords régionaux de libéralisation, sans distinguer différents types d'accords. L'effet moyen estimé est de 0,36 ce qui signifie que le commerce est 40% plus élevé entre les membres d'un accord régional de libéralisation qu'entre deux pays similaires ne participant pas à l'accord. De nombreuses estimations ont cependant été proposées, qui estiment l'effet spécifique de certaines expériences d'intégration importantes, en particulier l'intégration européenne et l'intégration nord-américaine. Dans le tableau 3.3, la ligne "EU" mesure l'effet spécifique supplémentaire de l'intégration européenne. Celui-ci est de 0,16 en moyenne ce qui signifie que les pays européens ont un commerce 17% plus élevé que les membres d'une autre zone d'intégration régionale, comparables en termes de taille. La ligne "CUSA/NAFTA"

mesure l'effet spécifique de la zone de libre-échange nord-américaine. Bien que le degré d'intégration économique dans cette zone soit nettement moins important que dans l'Union Européenne, l'effet supplémentaire estimé est très important, 0,76 en moyenne ce qui correspond à un doublement de l'intensité du commerce, par rapport à la zone de libre-échange "moyenne"¹⁵. Finalement, l'effet supplémentaire de l'intégration monétaire est mesuré dans la ligne "Common currency". Celui-ci est également très important, 0,86 en moyenne ce qui signifie que les membres d'une union monétaire échangent 2,3 fois plus que des pays similaires n'appartenant pas à une zone monétaire.

La grande variabilité des coefficients estimés pour différentes zones d'intégration régionale et la difficulté à interpréter cette variabilité ont conduit à de nombreux débats dans les milieux économiques. La question centrale est de savoir si on peut utiliser ces résultats empiriques pour inférer quelque chose sur l'effet potentiel d'une *nouvelle* expérience d'intégration. Ces débats ont été particulièrement nombreux autour de la question de l'intégration monétaire. En effet, un des bénéfices attendus de l'intégration monétaire européenne était un renforcement du commerce intra-européen. L'ampleur de ce renforcement devait (ou pas) compenser les coûts induits par la perte de souveraineté monétaire des économies de la zone euro. Au moment où l'euro a été introduit, après une période de convergence coûteuse pour les économies européennes, la réalisation de ces bénéfices était une question importante politiquement.

L'effet estimé initialement par Rose (2000) à partir de données relatives à de nombreuses unions monétaires est jugé beaucoup trop élevé pour s'appliquer au commerce de la zone euro. Rose estime en effet que les membres d'une union monétaire commercent trois fois plus que des partenaires similaires non membres d'une union monétaire. Ce résultat très (trop ?) élevé a conduit à de nombreux débats méthodologiques sur l'estimation de l'équation de gravité. En particulier, il a conduit à s'interroger sur le caractère endogène de l'intégration commerciale. De tels coefficients élevés pourraient s'expliquer par une causalité inverse. Les pays qui choisissent de signer des accords d'intégration ne sont pas tirés au hasard mais sont ceux qui commercent beaucoup

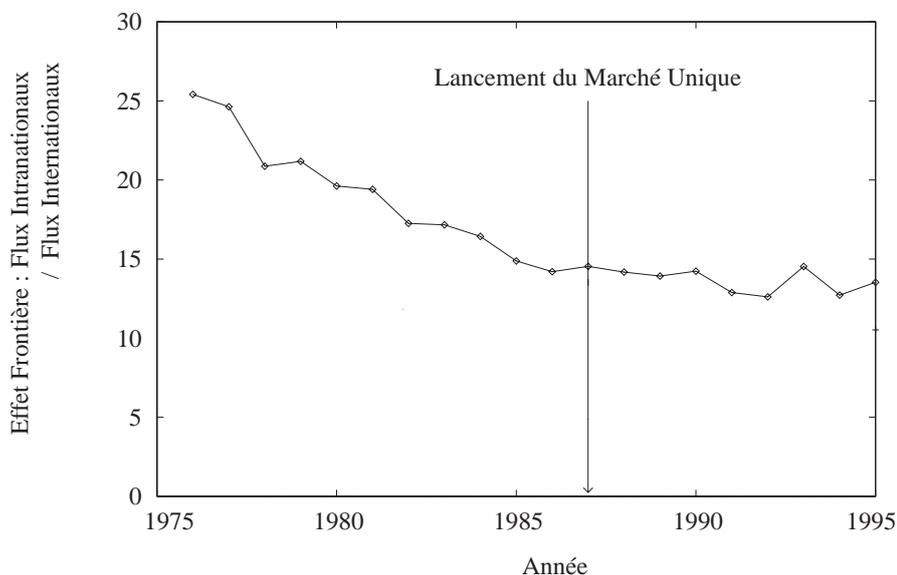
15. Depuis 1993, l'Union Européenne (UE) est un "marché unique", ce qui signifie que les biens, les personnes et les capitaux peuvent circuler librement entre les différentes économies de l'UE. Avant la signature de l'Acte Unique Européen, la Communauté Economique Européenne (CEE) était une union douanière, i.e. que les tarifs douaniers étaient nuls à l'intérieur de l'Europe et étaient unifiés pour toutes les relations des économies européennes avec le reste du monde. L'Accord de Libre-Echange Nord-Américain (ALENA ou NAFTA) fait du Canada, des Etats-Unis et du Mexique une zone de libre-échange, dans laquelle les tarifs internes sont nuls mais les tarifs appliqués par chaque économie au reste du monde restent décidés souverainement. Le marché unique est donc l'étape la plus avancée de l'intégration des marchés réels.

initialement. L'effet estimé important de l'intégration serait une conséquence de cette causalité inverse. Les estimations proposées depuis, qui tiennent compte de cette possible endogénéité, semblent converger vers un effet plus faible de l'intégration monétaire, autour de 0,7, effet qui serait même encore plus limité pour l'Union Monétaire Européenne.

Effet frontière : Dans la littérature sur l'intégration régionale, les zones de libre-échange sont comparées au reste du monde, supposé moins intégré. Mais l'expérience d'intégration qu'ont connu les économies européennes au cours des 60 dernières années a conduit ces économies à être aussi intégrées que peuvent l'être les différentes régions d'une économie nationale (en principe au moins). Tous les flux de biens, de services, de personnes et de capitaux sont parfaitement libres au sein de l'Union Européenne. En outre, l'introduction d'une monnaie commune à une grande partie des membres de cette union a encore réduit les coûts à l'échange intra-zone en supprimant les problèmes de conversion monétaire et les coûts induits par l'incertitude sur les parités de change. Pour une telle zone d'intégration, le point de comparaison devrait donc être l'économie nationale et pas le reste du monde, dont les économies sont fortement segmentées. McCallum (1995) est le premier à mettre en avant cet argument et à proposer un test empirique. Si l'intégration est complète, les frontières à l'intérieur de la zone intégrée ne devraient plus avoir aucun impact. L'estimation de l'"effet frontière" permet donc de mesurer à quelle distance l'économie se trouve de l'intégration totale.

McCallum applique ce raisonnement à la zone de libre-échange nord-américaine, et plus spécifiquement au commerce entre les Etats-Unis et le Canada. L'intérêt de cet exemple est que les pays sont à la fois extrêmement intégrés économiquement mais aussi proches culturellement, partageant la même langue, les mêmes institutions, et un certain nombre de références culturelles. En outre, on dispose pour cette zone géographique de données de commerce au niveau régional. L'institut statistique canadien offre en effet des statistiques sur le commerce entre provinces canadiennes et sur le commerce entre les différentes provinces canadiennes et les Etats américains. Si l'intégration économique était totale entre les Etats-Unis et le Canada, les déterminants du commerce entre deux provinces canadiennes devraient être les mêmes que pour le commerce entre une province canadienne et un Etat américain. A distance équivalente, il n'y a aucune raison que les provinces canadiennes commercent plus l'une avec l'autre. Pourtant c'est le cas, comme le montre McCallum (1995). En particulier, à distance donnée et pour des régions de tailles similaires, le commerce entre provinces canadiennes est 20 fois plus élevé qu'entre une province canadienne et un Etat américain. L'"effet-frontière" est substantiel.

FIGURE 3.6 – L'effet frontière à l'intérieur de l'Union Européenne : Ratio des flux infranationaux aux flux internationaux



Source : Head & Mayer (2000).

Head & Mayer (2000) reproduisent le même type d'exercice, en considérant le commerce entre pays membres de l'Union Européenne et le commerce "interne" aux économies européennes. Le résultat est illustré sur la Figure 3.6. Les échanges à l'intérieur des économies européennes sont en moyenne 20 fois plus importants que le commerce entre les différents pays européens. Comme on peut s'y attendre, l'effet frontière a tendance à diminuer au cours du temps mais reste très élevé et difficile à expliquer.

La taille de l'effet frontière reste un puzzle empirique en commerce international. De nombreuses explications ont été avancées, bien qu'aucune ne soit suffisante pour reproduire le fait stylisé. La présence de barrières non tarifaires aux échanges pourrait réduire le commerce transnational, par rapport au commerce infranational. L'utilisation de monnaies différentes est également un facteur de segmentation des marchés, comme discuté plus haut. Les consommateurs ont probablement des préférences biaisées en faveur des produits domestiques. C'est particulièrement le cas de la consommation publique. Enfin, le rôle des réseaux (sociaux ou d'affaires), qui sont plus denses à l'intérieur de l'économie nationale, pourrait participer à la taille de l'effet frontière.

3.3.4 Limites

Sur données agrégées, l'équation de gravité s'est révélé un outil empirique utile permettant de prédire la taille des échanges bilatéraux et de quantifier les bénéfices de l'intégration économique. Cependant, ce cadre empirique se révèle incapable de reproduire un fait stylisé important, la présence de "zéros" dans la matrice du commerce bilatéral. Une illustration de ce fait stylisé se retrouve dans le nombre d'observations utilisées pour estimer l'équation de gravité dans le tableau 3.2. Pour 1970, 9 035 relations bilatérales sont utilisées, sur un maximum de 31 506 dans un échantillon de 178 pays. Seulement 29% des paires de pays reportent un flux de commerce strictement positif¹⁶. En 2006, ce ratio est plus important (41% sur les 41 006 flux bilatéraux potentiels) mais reste faible. La moitié des paires de pays n'échangent pas du tout, un fait stylisé que ni l'équation de gravité, ni le modèle de Krugman ne permettent d'expliquer.

Dans le modèle de Krugman, l'hypothèse de préférences CES implique que les consommateurs consomment une quantité strictement positive de toutes les variétés produites dans le monde, même celles exportées de pays très éloignés, qui souffrent d'un déficit concurrentiel important. Pourtant, on observe empiriquement qu'un grand nombre de pays n'échangent pas du tout. En outre, ce nombre augmente dramatiquement si on regarde le commerce de biens particuliers, et pas le commerce agrégé. A la limite, au niveau individuel, une fraction très faible des entreprises exportent tandis qu'une partie importante du commerce agrégé est attribuable à un petit nombre de "grandes" entreprises. Par exemple, dans les données françaises, 17,4% des entreprises n'exportent pas du tout et 34% des exportateurs ne servent qu'un marché étranger (très souvent la Belgique ou l'Allemagne). A l'autre extrême, 1% des entreprises françaises sont responsables de 68% des exportations agrégées.

La littérature récente en commerce international s'est concentrée sur ce fait stylisé. A la suite de Melitz (2003), la littérature théorique s'est orientée vers des modèles dans lesquels des entreprises hétérogènes en termes de productivité font des choix hétérogènes en termes de participation au commerce international. Ce type de modèles permet d'expliquer la présence de "zéros" dans la matrice de commerce, au niveau individuel et dans les données agrégées. Au niveau individuel, les entreprises qui sont insuffisamment productives réaliseraient à l'étranger un chiffre d'affaire qui n'est pas suffisant pour couvrir les coûts à l'exportation¹⁷. A l'optimum, elles choisissent dont

16. L'équation de gravité étant une spécification en log, le nombre d'observations correspond aussi au nombre de flux strictement positifs dans l'échantillon.

17. Le modèle de Melitz (2003) est basé sur la structure du modèle de Krugman (1980).

de ne servir que le marché domestique. A l'autre extrémité de la distribution, les entreprises les plus productives ont un avantage concurrentiel important et réalisent donc des ventes, sur leur marché domestique et à l'étranger, qui sont plus importantes que leurs concurrents moins productifs. C'est ce qui explique l'extrême concentration des ventes sur les marchés internationaux, observée dans les données désagrégées.

Ce modèle permet également d'expliquer la présence de "zéros" dans la matrice du commerce agrégé. Au niveau agrégé, des coûts fixes très élevés à exporter vers des destinations "exotiques" peuvent expliquer qu'aucune entreprise d'un pays ne décide d'exporter vers ces destinations. Ce sera d'autant plus le cas que le coût moyen de production dans le pays d'origine est élevé (i.e. que w_i/φ_i est élevé), que la destination a un potentiel de marché faible ($R_j P_j^{\sigma-1}$ petit) et que la distance entre les deux pays est importante (τ_{ij} élevé). Ce modèle permet donc d'intégrer les zéros dans le cadre de l'équation de gravité, en fournissant des fondements théoriques à la probabilité d'occurrence d'un tel flux nul. Helpman, Melitz & Rubinstein (2008) ont montré comment l'intégration de ce cadre analytique à l'équation de gravité pouvait augmenter la qualité prédictive de ce cadre empirique.

3.4 Conclusion

Dans ce chapitre, nous avons étudié une classe de modèles qui permettent d'expliquer pourquoi des pays similaires en termes de niveaux de développement, de technologies ou de dotations factorielles peuvent avoir un intérêt mutuel à l'échange international. Cette classe de modèles repose sur l'hypothèse d'une production à rendements d'échelle croissants. Les rendements croissants impliquent un gain à la spécialisation dans la production d'un nombre limité de biens. Lorsque ces économies d'échelle portent sur la production de biens différenciés, et qu'en outre les consommateurs ont un goût pour la diversité, l'ouverture au commerce conduit à une spécialisation des économies dans la production de variétés différentes de biens différenciés et à l'échange croisé de ces variétés. Dans ce cadre d'analyse, on observe donc des échanges croisés intra-sectoriels. Ce type de modèles permet donc d'expliquer les 40% du commerce mondial portant sur des flux croisés intra-sectoriels, mentionnés en introduction.

Dans ce cadre d'analyse, la valeur agrégée des échanges dépend de la taille de la demande, de la taille du pays exportateur, qui détermine l'offre de

A ce cadre d'analyse sont ajoutées deux hypothèses : i) une hétérogénéité des firmes en termes de productivité (i.e. des productivités du travail φ différentes), ii) un coût fixe à l'exportation.

variétés qu'il peut vendre sur les marchés internationaux, et de l'ampleur des coûts à l'échange international. En effet, le cadre d'analyse tient compte de l'existence de barrières à l'échange (tarifs mais aussi barrières non tarifaires) qui réduit la compétitivité-prix des variétés produites à l'étranger, donc le volume de la demande d'importations. Le commerce est freiné par les coûts à l'échange et l'hétérogénéité géographique de ces coûts peut expliquer la répartition spatiale des flux bilatéraux de commerce.

Enfin, en étendant ce modèle à un cadre à deux secteurs, on a montré que l'échange de biens différenciés pouvait conduire à de la spécialisation *intersectorielle*. En particulier, le rôle des économies d'échelle, combiné à l'effet des coûts à l'échange sur la compétitivité-prix des producteurs domestiques, fait de la taille du pays une source d'avantage comparatif dans la production du bien différencié. A l'équilibre, les grands pays se "spécialisent" dans la production de biens différenciés ce qui permet de maximiser les gains d'économie d'échelle. Ils produisent une part plus que proportionnelle de l'offre mondiale de variétés tandis que les "petits" pays se spécialisent dans la production de biens homogènes, non soumis aux économies d'échelle.

Le succès de ces modèles repose en partie sur le support empirique qu'ont reçu ces prédictions. En particulier, le modèle de Krugman offre un support théorique à l'équation de gravité, ce qui en fait *de facto* un succès empirique. La dernière partie de ce chapitre résume les enseignements tirés de la littérature empirique sur l'équation de gravité. Cette littérature a permis de formaliser de façon plus précise l'origine des coûts à l'échange. Ces coûts sont en partie tarifaire, ce qui explique que l'intégration régionale conduit effectivement à un renforcement de l'intensité des échanges. Ils sont aussi non tarifaires, notamment liés à des éléments culturels qui sont plus difficiles à mesurer, et à comprendre. L'intensification potentielle des échanges au-delà des niveaux atteints aujourd'hui repose en grande partie sur la suppression de ces barrières "inobservables" à l'échange.

Chapitre 4

Approche intertemporelle du Solde Courant

Dans les chapitres précédents, nous avons étudié différentes théories expliquant le commerce international. Dans les modèles statiques étudiés, la balance commerciale est toujours équilibrée (la valeur des exportations est égale à la valeur des importations), et les capitaux sont par hypothèse immobiles entre pays. Dans un des modèles (HOS) le commerce égalise les prix des facteurs, de sorte que les flux de capitaux seraient même inutiles s'ils étaient possibles. Cependant, dans la réalité, très peu de pays ont un commerce équilibré, et la plupart des pays échangent à la fois des biens et des services et des actifs financiers avec les mêmes partenaires. Pour comprendre les échanges de capitaux internationaux, nous devons donc acquérir de nouveaux outils théoriques.

Dans ce chapitre, nous commencerons par approfondir la présentation de la Balance des Paiements, commencée dans le chapitre introductif. Ce document, qui résume les transactions d'un pays avec le reste du monde, aide à comprendre que chaque transaction commerciale a pour contrepartie une transaction financière, et que ces transactions reflètent des décisions d'épargne et d'investissement nationales. Comme les décisions d'épargne et d'investissement engagent le futur, nous sortirons du cadre statique des modèles précédents et présenterons des modèles simples de choix intertemporel. Ces modèles rationalisent l'existence de déséquilibres temporaires de la balance courante. Enfin, nous introduirons la notion de taux de change réel et son rôle dans l'ajustement des déséquilibres courants, dont nous poursuivrons l'étude dans les chapitres suivants.

4.1 Définitions : Balance des Paiements et taux de change

4.1.1 La Balance des Paiements

Il est impossible de comprendre les échanges commerciaux et financiers entre pays sans consulter leurs balances des paiements et particulièrement leurs soldes courants.

La *balance des paiements (BdP)* est un document comptable qui recense l'ensemble des transactions entre résidents et non-résidents d'un pays au cours d'une période de temps, typiquement une année. Elle mesure donc des flux, contrairement à la *position extérieure nette*, qui enregistre les stocks de créances et de dettes des résidents vis-à-vis des non-résidents.

La BdP enregistre chaque transaction deux fois, une fois au crédit et une fois au débit, suivant le principe de comptabilité en partie double. Par conséquent, sauf erreurs et/ou omissions, la construction de la BdP implique un solde égal à zéro. L'information capturée par la BdP réside dans le signe et l'ampleur des soldes des comptes qui la composent ¹.

Il existe trois grands comptes. Le *compte de transactions courantes* recense les échanges de biens et de services, les revenus primaires versés à l'étranger et reçus de l'étranger, et les transferts (tels que les sommes versées et reçues en provenance du budget de l'UE). Le *compte de capital* regroupe les transferts en capital (remise de dette, pertes sur créance...) et les cessions/acquisitions d'actifs non financiers non produits (brevets, marques, etc.). Le *compte financier* résume les opérations financières avec l'étranger, incluant les investissements directs, les investissements de portefeuille, les instruments financiers dérivés, les autres investissements (prêts et emprunts bancaires) et les avoirs de réserve.

La BdP fait le lien entre flux commerciaux et flux financiers. Comme mentionné en introduction, tout pays ayant un excédent commercial doit automatiquement acquérir des actifs étrangers de valeur égale, que ce soit des liquidités en devise étrangère, des créances bancaires ou des actifs financiers. Inversement un pays en déficit commercial accumule un passif vis-à-vis de l'étranger. Comme le solde du compte de capital est généralement petit, le solde financier représente en première approximation la contrepartie du solde courant.

1. En réalité, les statistiques de la BdP ne donnent pas généralement pas un solde exactement égal à zéro. On crée donc un compte Erreurs et Omissions Nettes représentant les incohérences entre sources statistiques et les erreurs. Certains économistes utilisent également ce compte pour estimer les activités financières illicites. Pour plus de détails sur la construction des BdP on peut consulter [une page de la Banque Centrale Européenne](#)

Pour bien comprendre comment la BdP est construite, il est utile de présenter quelques exemples introductifs. Par construction, toute opération est enregistrée deux fois, une fois à l'actif avec un signe positif (paiement de l'étranger) et une fois au passif avec un signe négatif (paiement vers l'étranger). Dans le chapitre introductif, nous avons mentionné l'exemple d'une firme française important un bien d'une firme américaine et réglant la transaction par virement sur le compte bancaire aux Etats-Unis du fournisseur. Supposons que la valeur de l'achat soit 100.000 euros. Dans ce cas la BdP française enregistrera les opérations suivantes (en millions d'euros) :

I. Compte Courant	-0,1
A.1.Biens	-0,1
III. Compte Financier	+0,1
D.Avoirs de Réserve	+0,1

Un signe positif représente une écriture au crédit tandis qu'un signe négatif représente une écriture au débit. Dans les échanges de biens et services, une importation est enregistrée au débit tandis qu'une exportation est enregistrée au crédit. Dans les échanges d'actifs, y compris la monnaie, une sortie de capitaux (hausse des créances sur les non-résidents ou baisse des dettes envers les non-résidents) est enregistrée au débit, tandis qu'une entrée de capitaux (baisse des créances sur les non-résidents ou hausse des dettes envers les non-résidents) est enregistrée au crédit.

Dans ce premier exemple, l'importation est enregistrée au débit de la France. La vente d'euros contre des dollars pour effectuer le virement bancaire entre au crédit de la France, en tant que hausse de la dette envers les non-résidents (la monnaie est une dette de la Banque Centrale envers ses détenteurs). Notons que seules les transactions entre résidents et non-résidents sont enregistrées dans la BdP, de sorte que les transactions intermédiaires entre les entreprises et leurs banques (résidentes du même pays) n'apparaissent pas.

Considérons un deuxième exemple, représentant de façon simplifiée la situation d'un pays émergent dans les années 90 : le pays paie intérêt et principal sur la dette passée, encourt un déficit commercial et finance l'ensemble en émettant de nouvelles obligations :

Le déficit commercial de 10 apparaît logiquement avec un signe négatif. Le paiement du principal de la dette passée apparaît au compte financier avec un signe négatif (-20), comme réduction de la dette envers les non-résidents. Le paiement des intérêts apparaît avec un signe négatif pour les mêmes raisons, mais dans le compte courant, étant un revenu primaire versé

I. Compte Courant	-20
A.1.Biens	-10
B.Revenus	-10
III. Compte Financier	+20
B.Investissements de portefeuille	-20+40

aux non-résidents. Enfin, l'émission d'obligations nouvelles correspond à une hausse de la dette envers les non-résidents ou entrée de capitaux (+40). Notons que dans cet exemple, si le pays perdait subitement la confiance des investisseurs étrangers, il ne pourrait pas se refinancer de 40 : il faudrait alors par exemple dégager un excédent commercial de 30 en dévaluant la monnaie, réduire les réserves de 30, emprunter cette somme auprès du FMI, etc.

Le Tableau 4.1 présente de façon simplifiée les BdP de la Chine, de la zone euro et des Etats-Unis en 2012. On observe que la Chine et la zone euro sont exportateurs nets et ont un solde courant positif, contrairement aux Etats-Unis. On voit également qu'en contrepartie de leur excédent courant la Chine et la zone euro ont un solde financier négatif, mais de nature différente : en Chine, les autres investissements et avoirs de réserves augmentent, tandis que dans la zone euro ce sont les investissements directs à l'étranger qui augmentent.

En conclusion, on peut retenir plusieurs leçons importantes de la simple lecture de la Balance des Paiements. Premièrement, son solde est toujours nul, mais les soldes des comptes courant et financier ne le sont généralement pas, et ce sont eux qui apportent l'information majeure de ce document. Deuxièmement, aux déficits courants d'un pays correspondent les excédents courants d'un autre pays (aux erreurs et omissions près). Les transferts unilatéraux (aide publique, remise de dette) sont les seuls éléments de la BDP issues d'une décision unilatérale. Ainsi, les déséquilibres courants résultent en principe d'échanges librement consentis entre deux ou plusieurs pays, et leurs sources sont à chercher dans la situation économique de l'ensemble de ces pays.

4.1.2 Solde courant et épargne nette

Nous avons vu en introduction une deuxième identité comptable utile (équation 1.1) pour comprendre la détermination du solde courant (en négligeant les revenus primaires nets). Cette identité se réécrit pour faire apparaître l'épargne et l'investissement nationaux.

$$(Y - T - C - I) + (T - G) = X - M$$

	Chine	Zone Euro	Etats-Unis
I. Compte Courant	+193,1	+166,5	-440,4
A. Biens et Services	+231,8	+242,2	-534,7
1. Biens	+321,6	+125,2	-741,5
2. Services	-89,7	+117	+206,8
B. Revenus	-42,1	+65,1	+223,9
C. Transferts Courants	+3,4	-140,9	-129,7
II. Compte de Capital	+4,3	+6,6	+7
III. Compte Financier	-117,7	-185,9	+439,4
A. Investissements Directs	+191,1	-4,8	-222
B. Investissements de Portefeuille	+47,8	+95,4	+586,8
C. Autres Investissements	-260	-258,3	+70,2
D. Avoirs de Réserve	-96,6	-18,3	+4,5
IV. Erreurs et Omissions Nettes	-79,8	+12,9	-5,9

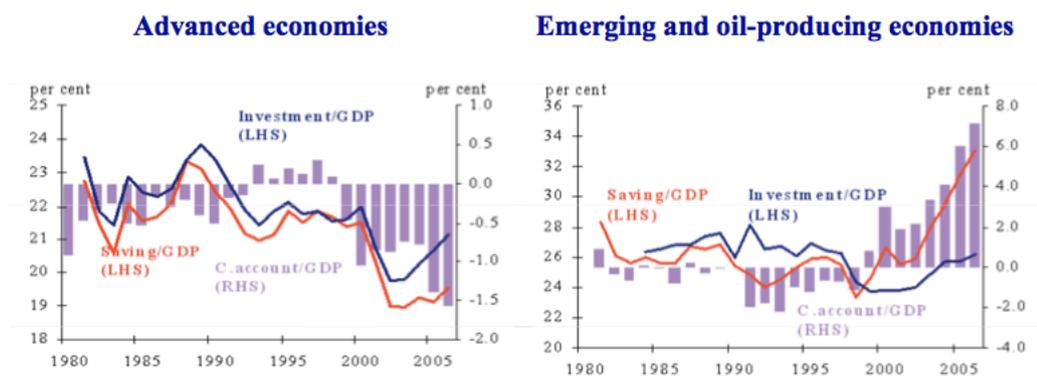
TABLE 4.1 – Balance des Paiements simplifiées de la Chine, de la zone euro, et des Etats-Unis en 2012, en milliards de dollars US. Sources : SAFE (Chine), BEA (Etats-Unis), BCE.

où T désigne les revenus fiscaux du gouvernement et G désigne les dépenses publiques. Le terme de gauche, qui correspond au terme $S - I$ de l'équation 1.1, décrit l'épargne nette nationale comme la somme de l'épargne nette privée et publique.

On appelle *épargne nette* la différence entre épargne et investissement nationaux. Selon cette identité comptable, le solde courant est donc égal à l'épargne nette, c'est-à-dire la somme de l'épargne nette privée et de l'épargne nette publique. Un pays ayant un excédent courant est un pays ayant une épargne nette positive, qu'elle vienne du secteur privé ou public. Les déterminants du solde courant d'un pays seront liés à ceux de son épargne et de son investissement.

La Figure 4.1 illustre bien le lien entre solde courant et épargne nette issu de l'identité comptable. On constate aussi depuis la fin des années 90 un fort déficit courant des économies développées et un fort excédent courant des pays émergents. Il n'est pas étonnant qu'après l'accélération de l'ouverture financière à cette période on observe une plus grande déconnexion entre épargne et investissement nationaux. Après tout, dans un monde idéal sans coûts de transaction internationaux et sans asymétries d'information, l'épargne et l'investissement nationaux devraient être *déconnectés*, et tout investissement rentable devrait trouver financement quelle que soit la prove-

FIGURE 4.1 – Épargne, investissement et solde courant en pourcentage du PIB dans les économies développées (gauche) et dans les pays exportateurs de pétrole et émergents (droite) de 1980 à 2007



Source: IMF.

nance de l'épargne. Cependant, on peut penser qu'un pays ne peut indéfiniment accumuler une dette ou une créance sur le reste du monde. Dans le reste de ce chapitre, nous proposerons une approche intertemporelle du solde courant, modélisant l'idée que les déficits courants sont fondés sur la promesse d'excédents futurs. Dans le chapitre 5, nous verrons comment amender ces modèles pour mieux rendre compte de déséquilibres persistants.

4.1.3 La position extérieure nette

Comme mentionné en introduction, la position extérieure nette représente la richesse ou l'endettement nets des résidents vis-à-vis de l'étranger à un moment dans le temps. Lorsque la position extérieure est négative cela signifie que les résidents empruntent davantage au reste du monde qu'ils ne lui prêtent. Elle est donc le pendant de la BdP, mesurant des stocks lorsque la BdP mesure des flux.

Nous avons vu avec la Figure 1.1 que la position extérieure nette peut être approchée par la somme des soldes courants cumulés au cours du temps. Cependant, si effectivement le solde courant de l'année vient ajouter de nouvelles dettes et créances à la position extérieure nette, celle-ci varie également sous l'effet de changements de valorisation des dettes ou créances existantes. Ces changements viennent de changements de rendement liés aux fluctua-

tions des marchés financiers, ou de fluctuations des taux de change quand les dettes ou créances sont libellées en monnaie étrangère. De plus, la variation ne sera pas nécessairement de même signe côté dettes et côté créances. En pratique, ces effets de valorisation sont négligeables pour la plupart des pays, mais pas pour les Etats-Unis, comme le suggère la Figure 1.1. En effet, pour des raisons historiques liées au système de Bretton Woods, les Etats-Unis ont la capacité d'émettre facilement de la dette à court terme peu risquée (notamment publique) en dollars, tout en investissant à l'étranger dans des actifs plus risqués, à plus long terme et en monnaie étrangère. Ainsi, dans le cas des Etats-Unis, l'incidence des déficits courants cumulés sur la position extérieure nette a été tempérée par l'écart de rendement en faveur des Etats-Unis, surtout lors des épisodes de dépréciation du dollar. Cette différence entre soldes courants cumulés et position extérieure nette entre 1970 et 2010 atteignait 36% du PIB américain de 2010.

Dans ce qui suivra, on fera le choix simplificateur de négliger ces effets de valorisation. Nous traiterons les actifs financiers nationaux et étrangers comme parfaitement substituables (une fois leurs rendements convertis dans la même monnaie), et utiliserons de façon interchangeable les soldes courants cumulés et la position extérieure nette. On notera simplement que ces questions de substitution imparfaites entre actifs nationaux et étrangers sont abordées par les travaux récents sur les Etats-Unis.

4.1.4 Les taux de change nominal et réel

Avec le solde du compte courant, le second grand objet d'étude des chapitres à venir sera le taux de change. Plus précisément, nous étudierons deux types de taux de change, nominal et réel. Les chapitres 6 et 7 sont plus particulièrement consacrés à leurs déterminants.

Le *taux de change nominal* indique la valeur d'une monnaie en unités d'une autre monnaie, et correspond au "taux de change" du langage courant. Il existe deux conventions pour définir le taux de change. Dans ce qui suit nous définirons généralement le taux de change en unités de monnaie étrangère (définition dite "au certain"). Ainsi une hausse du taux de change S_t correspond à une appréciation de la monnaie nationale. Par exemple pour un Européen, un taux de change de 1 EUR = 1,25 USD est noté $S_t = 1,25$ et une hausse de S_t correspond à une appréciation. Néanmoins, la définition du taux de change "à l'incertain", qu'on peut noter $E_t = \frac{1}{S_t}$, est au moins autant utilisée par les économistes. Il est utile de se familiariser avec les deux conventions.

Le *taux de change réel* (Real Exchange Rate en anglais, TCR ci-après) exprime le prix des biens nationaux relativement au prix des biens étrangers.

Par analogie avec le taux de change nominal, il exprime la valeur d'un panier de biens d'un pays en unités de panier de bien (identique) d'un autre pays. Ce panier de biens peut être celui utilisé dans la construction de l'indice des prix à la consommation mensuel des instituts statistiques nationaux, le Big Mac de McDonald's dont les prix sont recueillis par le journal *The Economist*, ou un panier standardisé dont les prix sont recueillis tous les 5 ans par le groupe International Comparison Project (ICOP) de la Banque Mondiale. Dans le cas du Big Mac, le panier se réfère implicitement au prix des ingrédients, comme nous le verrons dans le chapitre 6.

4.2 L'approche intertemporelle du solde courant

Tant les décisions d'épargne que d'investissement engagent le futur. La modélisation de ces décisions requiert donc un cadre dynamique dans lequel les agents prévoient leur consommation et leur investissement. Il en résulte une approche intertemporelle du solde courant, où la dynamique de celui-ci répond à un choix intertemporel optimal compte tenu des anticipations des agents économiques.

Nous allons présenter deux modèles simples suivant cette approche intertemporelle. Un modèle à un seul bien, un seul actif et deux périodes, suffit à générer une prédiction simple sur la dynamique du solde courant. Pour aller plus loin et relier la dynamique du solde courant aux taux de change réel et nominal, on introduit un modèle à deux biens : un bien échangeable et un bien non échangeable.

Ces modèles illustreront une source fondamentale de gain à l'échange financier : la possibilité de lisser sa consommation au cours du temps malgré des chocs sur le revenu. Dans la PC associée à ce chapitre et dans le chapitre 5, nous verrons également comment l'ouverture financière permet une allocation mondiale du capital plus efficiente².

2. Dans l'introduction, nous avons mentionné un autre gain à l'échange fondamental, la diversification des risques propres à chaque pays. Ce gain est généralement abordé dans des modèles de choix de portefeuille. Pour des raisons d'espace nous n'aborderons pas ce type de modèle dans ce cours. On peut cependant considérer les questions de partage de risque et les questions de choix intertemporel comme indépendantes.

4.2.1 Le modèle de Fisher : un bien, un actif, deux périodes

Le modèle de Fisher (1930) à un bien, un actif sans risque et deux périodes offre le cadre le plus simple pour prédire la dynamique du solde courant ³.

Considérons une petite économie ouverte, c'est-à-dire une économie qui est affectée par le reste du monde mais qui ne l'affecte pas. Cette économie peut échanger des biens et services et prêter/emprunter au taux d'intérêt mondial r , qu'elle traite comme une variable exogène. Pour simplifier, on suppose aussi que la production est exogène (cette hypothèse sera relâchée dans un exercice de PC) et qu'il n'y a pas de gouvernement. On se concentre ainsi sur les décisions de consommation des ménages.

L'agent représentatif national reçoit à chaque période l'utilité de sa consommation $u(c_t)$ supposée séparable dans le temps ⁴. La préférence pour le présent $t = 1$ par rapport au futur $t = 2$ est capturée par le facteur d'escompte $0 < \beta < 1$. L'utilité intertemporelle est ainsi mesurée en $t = 1$ par

$$U(c_1, c_2) = u(c_1) + \beta u(c_2)$$

Le consommateur fait face à une contrainte de budget à chaque période. Pour simplifier les calculs on normalise à 1 le prix du bien de consommation à chaque période. Si le pays était en autarcie financière, l'agent ne pourrait emprunter et ferait face aux contraintes $c_1 \leq y_1$ et $c_2 \leq y_2$. Cependant, l'agent peut prêter ou d'emprunter auprès du reste du monde au taux r . Notons Z_1 l'épargne nette à $t = 1$, qui peut être positive ou négative. A la période 2 l'agent devra rembourser sa dette ou récupérer sa créance, de sorte que

$$Z_2 = -(1 + r)Z_1$$

On peut réécrire les contraintes de budget

$$\begin{aligned} c_1 + Z_1 &\leq y_1 \\ c_2 - Z_1(1 + r) &\leq y_2 \end{aligned}$$

ce qui implique une *contrainte de budget intertemporelle*

$$c_1 + \frac{c_2}{1 + r} \leq \underbrace{y_1 + \frac{y_2}{1 + r}}_Y$$

3. Les élèves ayant suivi ECO432 reconnaîtront une extension du modèle de la section 3.4.2. du photocopié de ce cours.

4. On suppose aussi que $u(\cdot)$ est strictement concave pour garantir un maximum unique.

Le terme de droite, qui représente la somme actualisée des revenus présent et futur, est ici exogène. On le note Y pour alléger la notation.

Ces hypothèses étant posées, on commence la résolution du modèle par le programme du consommateur

$$\begin{aligned} \max_{c_1, c_2} u(c_1) + \beta u(c_2) \text{ sous contrainte } c_1 + \frac{c_2}{1+r} &\leq Y \\ \Rightarrow \frac{u'(c_1)}{u'(c_2)} &= \beta(1+r) \end{aligned} \quad (\text{Euler})$$

Cette condition de premier ordre est aussi appelée *équation d'Euler de la consommation*. Cette équation d'Euler et la contrainte de budget intertemporelle suffisent à définir les valeurs optimales de c_1 et c_2 en fonction de r , Y et les paramètres de la fonction d'utilité.

Intuitivement, l'équation d'Euler dit que le consommateur alloue son revenu entre les deux périodes (en prêtant/empruntant) pour que le rapport des utilités marginales en $t = 1$ et $t = 2$, $\frac{u'(c_1)}{\beta u'(c_2)}$, soit égal au prix relatif de la consommation en $t = 2$, $1 + r$. La logique est exactement la même que dans le problème d'un consommateur allouant son revenu entre deux biens. Si le consommateur est impatient, au sens où sa préférence pour le présent est plus forte que le facteur d'actualisation, $\beta < \frac{1}{1+r}$, alors il consomme plus en première période, et inversement quand le consommateur est patient.

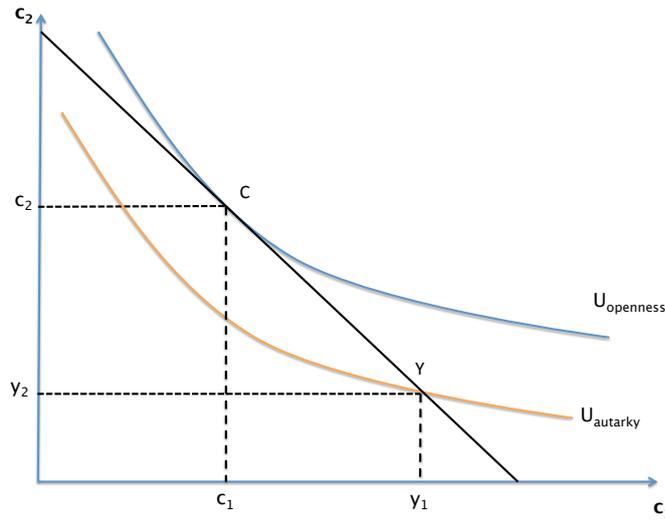
On note que le choix de consommation ne dépend que de Y et pas de la répartition entre y_1 et y_2 : le consommateur "lisse" sa consommation dans le temps, indépendamment de la chronologie de ses flux de revenus. Dans ce modèle, c'est même la raison pour laquelle l'ouverture financière augmente le bien-être. En prêtant ou empruntant au reste du monde au taux r , le consommateur peut atteindre des allocations intertemporelles de la consommation et des niveaux d'utilité correspondants inaccessibles en autarcie financière. La Figure 4.2 illustre cette propriété.

La Figure 4.2 illustre la différence d'utilité entre l'autarcie financière (consommation nationale égale à la production nationale à chaque période) et le libre-échange de capitaux. Le taux d'intérêt r , considéré comme exogène, détermine une contrainte de budget intertemporelle représentée par une droite décroissante dans l'espace $\{c_1, c_2\}$. La consommation maximisant l'utilité intertemporelle sous cette contrainte de budget n'est généralement pas au point Y (autarcie financière). La Figure 4.2 représente un exemple où le pays gagne à avoir un excédent courant en période 1 et un déficit en période 2.

Dans le cas particulier d'une fonction d'utilité logarithmique, l'équation d'Euler s'écrit

$$\frac{c_2}{c_1} = \beta(1+r)$$

FIGURE 4.2 – Gain d'utilité issu du lissage intertemporel permis par l'ouverture financière



Le point Y représente les revenus de périodes 1 et 2. En autarcie financière, ce sont aussi les niveaux de consommation maximaux à ces deux périodes. En ouverture financière, le consommateur fait face à la contrainte intertemporelle représentée par la droite (CY), de pente $1 + r$. Dans le cas représenté ici, le consommateur subit une forte baisse de revenu à la période 2, et trouve optimal d'épargner d'abord pour lisser sa consommation. Il gagne donc de l'utilité par rapport au cas d'autarcie.

qui avec la contrainte intertemporelle

$$c_1 + \frac{c_2}{1+r} = c_1(1+\beta) = Y$$

implique

$$c_1 = \frac{Y}{1+\beta}$$

$$c_2 = \frac{\beta(1+r)Y}{1+\beta}$$

La consommation est constante. Dans ce cas particulier le pays aura un surplus courant si $y_1 > \frac{\beta(1+r)Y}{1+\beta}$ et un déficit courant sinon. Bien que le cas d'utilité logarithmique soit peu réaliste, il nous servira de modèle de référence simple par la suite.

Au total, le modèle de Fisher à un bien, un actif et deux périodes, suffit malgré sa simplicité à illustrer des propriétés importantes de la finance internationale : l'ouverture financière permet le lissage intertemporel de la consommation, ce qui augmente l'utilité; tout surplus courant actuel implique un déficit courant futur, et inversement ($Z_2 = -(1+r)Z_1$). On peut facilement généraliser ce modèle à un nombre fini de périodes et obtenir le même résultat. On peut également le généraliser en rendant la production endogène et/ou en traitant le cas d'un grand pays dont l'activité affecte le taux d'intérêt r (voir exercices de PC).

4.2.2 Deux biens, un actif, deux périodes : solde courant et change réel

Le modèle précédent supposait un seul bien de consommation, parfaitement échangeable : par exemple si $y_1 > c_1$, la totalité de $y_1 - c_1$ est exportée à $t = 1$ en échange de créances sur le reste du monde, permettant d'augmenter la consommation future de $(y_1 - c_1)(1+r)$ dans le futur. Dans la réalité, seule une fraction des biens est échangeable, même si cette fraction peut évoluer au gré des innovations technologiques réduisant les coûts du commerce international. Nous allons donc étendre l'analyse précédente à deux biens : un bien échangeable (T) et un bien non-échangeable (N). Ces hypothèses vont nous permettre d'étudier le lien entre solde courant et prix relatif des échangeables et non-échangeables. Ce lien sera approfondi au chapitre 6.

On suppose que le bien T est échangeable sans aucun coût tandis que le bien N est complètement non-échangeable. On note p_t^T et p_t^N les prix de ces biens et on suppose pour simplifier que p_t^T est constant dans le temps. On note S_t le taux de change nominal, défini en unités de monnaie étrangère (au certain).

On suppose que l'agent représentatif a une fonction d'utilité Cobb-Douglas :

$$U(c_1^T, c_1^N, c_2^T, c_2^N) = \gamma \ln c_1^T + (1 - \gamma) \ln c_1^N + \beta (\gamma \ln c_2^T + (1 - \gamma) \ln c_2^N)$$

qu'il maximise sous une contrainte de budget intertemporelle similaire à celle de la section précédente

$$c_1^T + Q_1 c_1^N + \frac{c_2^T + Q_2 c_2^N}{1+r} \leq \underbrace{Q_1 y_1^N + y_1^T + \frac{Q_1 y_2^N + y_2^T}{1+r}}_Y$$

où on a défini par $Q_t \equiv \frac{p_t^N}{p_t^T}$ le prix relatif du bien non-échangeable.

On suppose qu'il n'y a aucun coût du commerce et que les marchés des produits sont parfaitement concurrentiels. Cela implique que les prix des

biens échangeables étrangers et nationaux exprimés dans la même monnaie sont égaux (ce qu'on appelle parfois Loi du prix unique), ou

$$p_t^T = \frac{p_t^{T*}}{S_t}$$

Définissons l'indice des prix national comme $P_t = (p_t^T)^\gamma (p_t^N)^{1-\gamma}$. Cette définition rappelle l'indice de prix idéal vu aux chapitres précédents, au sens où une hausse de cet indice requiert une hausse de même proportion du revenu pour que l'utilité reste constante. On peut réécrire $P_t = p_t^T (Q_t)^{1-\gamma}$, et suivant la "loi du prix unique"

$$P_t = \frac{p_t^{T*}}{S_t} (Q_t)^{1-\gamma} = \frac{P_t^* (Q_t)^{1-\gamma}}{S_t (Q_t^*)^{1-\gamma}} \Rightarrow \underbrace{S_t \frac{P_t}{P_t^*}}_{\text{Taux de Change Réel}} = \left(\frac{Q_t}{Q_t^*} \right)^{1-\gamma}$$

Le taux de change réel (TCR) est donc une fonction croissante du prix relatif des non-échangeables, ici avec une élasticité constante γ d'autant plus grande que la part des non-échangeables dans la consommation totale est élevée.

On cherche maintenant à résoudre le problème du consommateur et en déduire l'évolution du solde courant. Notons y_t^N et y_t^T la production de biens non-échangeables et échangeables, respectivement. Notons x_t les exportations nettes de bien T, et remarquons qu'elles sont ici égales au solde courant en l'absence de commerce international de facteurs de production. L'équilibre sur les marchés des deux biens s'écrit

$$\begin{aligned} y_t^N &= c_t^N \\ y_t^T &= c_t^T + x_t \end{aligned}$$

Le consommateur résout

$$\begin{aligned} &\max_{c_1^T, c_1^N, c_2^T, c_2^N} \{ \gamma \ln c_1^T + (1-\gamma) \ln c_1^N + \beta (\gamma \ln c_2^T + (1-\gamma) \ln c_2^N) \} \\ &\text{sous contrainte } c_1^T + Q_1 c_1^N + \frac{c_2^T + Q_2 c_2^N}{1+r} \leq \underbrace{Q_1 y_1^N + y_1^T + \frac{Q_2 y_2^N + y_2^T}{1+r}}_Y \end{aligned}$$

Les conditions du premier ordre de maximisation s'écrivent

$$\frac{c_2^T}{c_1^T} = \beta(1+r); \quad \frac{c_2^N}{c_1^N} = \beta(1+r) \frac{Q_1}{Q_2}; \quad \frac{c_1^T}{c_1^N} = \frac{\gamma}{1-\gamma} Q_1; \quad \frac{c_2^T}{c_2^N} = \frac{\gamma}{1-\gamma} Q_2$$

En utilisant ces conditions et la contrainte intertemporelle on déduit

$$\begin{aligned}c_1^T &= \frac{\gamma}{1+\beta}Y \\c_1^N &= \frac{1-\gamma}{Q_1(1+\beta)}Y \\c_2^T &= \frac{\beta(1+r)\gamma}{1+\beta}Y \\c_2^N &= \frac{\beta(1+r)(1-\gamma)}{Q_2(1+\beta)}Y\end{aligned}$$

L'équilibre sur le marché des non-échangeables implique $c_i^N = y_i^N$, d'où la relation

$$\frac{Q_2}{Q_1} = \beta(1+r)\frac{y_1^N}{y_2^N}$$

On peut montrer que

$$\begin{aligned}Q_1 &= \alpha \left[\frac{y_1^T + y_2^T}{y_1^N} \right] \\Q_2 &= \alpha\beta(1+r) \left[\frac{y_1^T + y_2^T}{y_2^N} \right]\end{aligned}$$

avec $\alpha \equiv \frac{1-\gamma}{\gamma(1+\beta)}$. Ainsi

$$Y = (1 + \alpha + \alpha\beta)(y_1^T + \frac{y_2^T}{1+r}) = \frac{1}{\gamma}(y_1^T + \frac{y_2^T}{1+r})$$

et

$$\begin{aligned}x_1 &= y_1^T - c_1^T = y_1^T - \frac{1}{1+\beta}(y_1^T + \frac{y_2^T}{1+r}) = \frac{\beta}{1+\beta}y_1^T - \frac{1}{(1+r)(1+\beta)}y_2^T \\x_2 &= -(1+r)x_1\end{aligned}$$

On peut maintenant étudier l'évolution du solde courant au cours du temps, et l'impact de chocs sur le change réel et le solde courant.

Premièrement, comme dans le modèle de Fisher, l'évolution du solde courant dépend de la patience des consommateurs. Plus un pays est "impatient" (β faible), plus il allouera de consommation à $t = 1$ plutôt que $t = 2$: en particulier lorsque $\beta < \frac{1}{1+r}$ on a $c_1^T > c_2^T$. Donc plus un pays sera impatient et plus il est probable qu'il aura d'abord un déficit courant puis un excédent. La nouveauté de ce modèle est qu'il prédit que l'ajustement se fera par une dépréciation réelle : $\beta < \frac{1}{1+r} \Rightarrow Q_2 < Q_1$. Dans un pays "patient", au contraire,

l'ajustement nécessitera une appréciation réelle. On peut également noter qu'une hausse du taux d'intérêt mondial r implique une réallocation du revenu vers la consommation future et une réduction du déficit courant ou une hausse de l'excédent courant⁵.

Deuxièmement, on peut étudier l'impact de chocs transitoires sectoriels sur le solde courant et le change réel par comparaison avec la situation de référence sans chocs. Un choc positif transitoire sur la production d'échangeables ($dy_1^T > 0, dy_2^T = 0$) augmente la consommation aux deux périodes, mais moins vite que la production, de sorte que le solde courant s'améliore à $t = 1$ relativement au cas de référence. La demande de N augmente aussi, et puisque l'offre est fixe le prix de N doit augmenter aux deux périodes. Le change réel est plus élevé que dans la situation de référence (mais comme précédemment, son évolution dans le temps dépend de la patience).

Un choc positif transitoire futur sur la production de T ($dy_1^T = 0, dy_2^T > 0$) encourage la consommation aux deux périodes. De la même façon le prix de N augmente aux deux périodes pour accommoder la hausse de demande, c'est-à-dire un taux de change réel plus élevé que dans la situation initiale. Par un raisonnement symétrique à celui du choc précédent, la consommation d'échangeables croît moins vite que la production à $t = 2$, d'où un solde courant à $t = 1$ inférieur et à $t = 2$ supérieur à leurs valeurs de référence. Par un raisonnement similaire, une économie avec un secteur T en croissance ($y_2^T > y_1^T$) a plus de chances d'être d'abord en déficit courant puis en excédent qu'une économie où ce secteur ne croît pas.

Enfin, un choc positif transitoire sur la production de N ($dy_1^N > 0, dy_2^N = 0$) implique une baisse de Q_1 , c'est-à-dire une dépréciation réelle par rapport à la situation initiale. Sans surprise, on consomme plus de non-échangeables à $t = 1$. Cependant la consommation de T ne varie pas et le solde courant non plus⁶. En $t = 2$ le change réel garde son niveau de référence en l'absence de choc.

5. Dans le cas d'un grand pays ouvert, un déficit courant en première période augmenterait le taux d'intérêt mondial r , de sorte que ce déficit courant serait plus faible que dans le cas d'un petit pays identique par ailleurs.

6. A noter qu'avec une autre fonction d'utilité que Cobb-Douglas on pourrait avoir un effet sur la consommation de T. Le choc sur N a trois effets sur la consommation de T : un effet substitution (négatif) et un effet revenu (positif) via la hausse de son prix relatif $\frac{1}{Q}$, ainsi qu'un effet revenu (aussi positif) via la hausse directe de Y . Avec les formes fonctionnelles choisies ici, les trois effets s'annulent parfaitement.

4.3 Conclusion

Dans ce chapitre, nous avons vu plusieurs enseignements fondamentaux de la macroéconomie internationale.

Premièrement, les flux de commerce et les flux financiers sont étroitement liés, puisque les uns sont la contrepartie des autres, comme l'illustre la Balance des paiements. Un pays exportateur net achète en contrepartie des créances sur le reste du monde. Un pays importateur net vend en contrepartie de la dette au reste du monde.

Deuxièmement, une autre identité comptable nous apprend que le solde courant est étroitement lié aux choix d'épargne et d'investissement dans le pays. Un pays en excédent courant est un pays avec un excès d'épargne. Un pays avec déficit courant est un pays avec excès d'investissement. Un pays ne peut avoir de déficit courant que si un autre pays accepte de le financer avec son excès d'épargne : le monde dans son ensemble est une économie fermée. Autrement dit, les déficits et excédents courants sont le résultat des décisions prises par au moins deux pays.

Troisièmement, tant l'épargne que l'investissement engagent des choix intertemporels. Le lissage intertemporel de la consommation (dans ce chapitre) ou les différences de taux de croissance et de rendement du capital (dans le prochain chapitre et en PC) expliquent l'existence de déficits et d'excédents courants. Les pays plus patients ou mieux dotés en capital exportent du capital aujourd'hui contre du capital demain, avec des gains à l'échange conceptuellement similaires aux gains à l'échange de biens et services. En principe, le solde courant sera généralement non nul, et aux déficits d'aujourd'hui correspondent les excédents de demain.

Quatrièmement, l'ajustement, c'est-à-dire la variation du solde courant vers un niveau compatible avec les contraintes de ressources intertemporelles, implique une variation du taux de change réel. Ceci requiert une variation des prix et/ou du taux de change nominal. Par exemple, un pays en déficit courant doit connaître une dépréciation réelle (dépréciation nominale et/ou baisse des prix nationaux). Nous verrons dans les chapitres à venir comment l'ajustement varie selon que le régime de change nominal est fixe ou flexible.

Finalement, le solde courant est lié à la demande pour les biens échangeables et non-échangeables. La corrélation entre les prix des échangeables d'un pays à l'autre fait que le change réel dépend étroitement du prix relatif des non-échangeables dans chaque pays. Des chocs asymétriques sur ces 2 secteurs ont un impact sur le change réel et sur l'évolution du solde courant.

Chapitre 5

Déséquilibres structurels

Dans le chapitre précédent nous avons étudié la notion de solde courant et des modèles simples permettant d'expliquer l'existence de déficits ou excédents temporaires par le désir de lisser la consommation.

Dans ce chapitre, nous verrons un deuxième motif de transactions financières internationales, lié aux différences de productivité marginale du capital entre pays. Le modèle de croissance néoclassique étendu à une économie ouverte nous apprendra que les pays connaissant une forte croissance de la productivité doivent recevoir des capitaux des autres pays. Cependant, nous verrons que cette prédiction contredit l'observation de déséquilibres structurels mondiaux avec des pays émergents (notamment asiatiques) nettement exportateurs nets de capitaux. Nous proposerons par la suite plusieurs explications à ces déséquilibres, dont certaines peuvent s'imbriquer dans une extension du modèle de croissance néoclassique.

5.1 Croissance et flux financiers internationaux à long terme

Dans cette section on introduit des flux financiers internationaux dans un modèle autrement semblable au modèle de croissance néoclassique, appelé modèle de Ramsey-Cass-Koopmans. Comme dans le modèle de Solow, les moteurs principaux de la croissance sont l'accumulation de capital et le progrès technique. Mais à la différence de Solow, le taux d'épargne est fixé de façon endogène pour atteindre le degré optimal de réallocation intertemporelle compte tenu des ressources et des connaissances technologiques¹.

1. Pour une présentation du modèle de Solow, voir le polycopié du cours EC0432, chapitre 1. Cette présentation du modèle de Ramsey-Cass-Koopmans suit la section 3 de Gourinchas & Rey (2014) et le chapitre 2 du manuel de Blanchard & Fischer (1989),

Le modèle est en temps continu. On considère un agent représentatif de durée de vie infinie (une dynastie) avec une fonction d'utilité dite CRRA :

$$U_t = \int_t^{+\infty} e^{-\rho(s-t)} N_s u(c_s) ds = \int_t^{+\infty} e^{-\rho(s-t)} N_s \frac{(c_s)^{1-\frac{1}{\sigma}}}{1-\frac{1}{\sigma}} ds \quad (5.1)$$

où c_t désigne la consommation entre t et $t + dt$, $\sigma > 0$ est l'élasticité de substitution intertemporelle, $\rho > 0$ paramètre la préférence pour le présent, N_t désigne la taille de la population en t . Pour simplifier on suppose que la population est constante² ou $N_t = N, \forall t$ et que chaque travailleur offre son temps de façon inélastique au salaire, de sorte que N est également l'offre de travail.

La production requiert du capital et du travail suivant la fonction de production

$$Y_t = F(K_t, N) \equiv (K_t)^\alpha (\xi_t N)^{1-\alpha} \quad (5.2)$$

où ξ_t croît au taux constant et exogène $g > 0$ et $\xi_0 > 0$. La croissance de Y viendra donc de l'accumulation de K et de l'évolution de ξ .

On suppose que le stock de capital initial est $K_0 > 0$ et que le capital se déprécie au cours du temps au taux δ , de sorte que

$$\forall t > 0, \dot{K}_t = I_t - \delta K_t \quad (5.3)$$

où I_t désigne l'investissement. On suppose que le marché du capital est concurrentiel.

En l'absence de gouvernement et en économie fermée, l'identité comptable de l'équilibre emploi-ressources national s'écrit :

$$Y_t = C_t + I_t \quad (5.4)$$

avec $C_t = N c_t$. On a normalisé les prix du bien de consommation et du bien d'investissement à 1 pour alléger les calculs.

5.1.1 Equilibre en économie fermée

On peut montrer que (mais on admettra ici) que *la solution décentralisée de ce problème est la même que la solution centralisée*³, dans laquelle un planificateur social bienveillant choisirait la séquence des $\{c_s\}$ pour maximiser l'utilité, compte tenu des valeurs initiales non-négatives K_0, ξ_0, N .

auxquels on peut se référer pour un traitement plus complet.

2. Le cas avec croissance de la population sera traité en PC.

3. Voir Blanchard & Fischer (1989), section 2.2.

Le planificateur résout

$$\max_{\{c_s\}} \left\{ \int_t^{+\infty} e^{-\rho(s-t)} Nu(c_s) ds \right\}$$

sous la contrainte

$$\forall t, (K_t)^\alpha (\xi_t N)^{1-\alpha} = Nc_t + \dot{K}_t + \delta K_t \quad (5.5)$$

On peut résoudre ce problème d'optimisation dynamique en utilisant la théorie du contrôle optimal. Définissons le Hamiltonien

$$H_t = e^{-\rho t} [Nu(c_t) + \lambda_t ((K_t)^\alpha (\xi_t N)^\alpha - \delta K_t - Nc_t)]$$

où λ_t est défini pour que $\mu_t \equiv e^{-\rho t} \lambda_t$ désigne le multiplicateur de Lagrange associé à la contrainte de la date t .

Dans ce genre de problème, on appelle c_s la variable de contrôle (choisie pour maximiser la fonction objectif), K_s la variable d'état et λ_s la variable de coétat. Selon la théorie du contrôle optimal, des conditions nécessaires et suffisantes pour que $\{c_s\}$ soit solution de ce problème sont

$$\begin{aligned} \frac{\partial H_t}{\partial c_t} &= 0 \\ \frac{\partial H_t}{\partial K_t} &= -\frac{\partial \mu_t}{\partial t} \\ \lim_{t \rightarrow +\infty} K_t \mu_t &= 0 \end{aligned}$$

ce qui implique

$$\begin{aligned} u'(c_t) &= \lambda_t \\ -\frac{\partial \lambda_t}{\partial t} &= \lambda_t \left(\frac{\partial F(K_t, N)}{\partial K_t} - \delta - \rho \right) \\ \lim_{t \rightarrow +\infty} K_t u'(c_t) e^{-\rho t} &= 0 \end{aligned}$$

En remplaçant λ_t on obtient l'équation d'Euler

$$-\frac{\partial u'(c_t)}{\partial t} = u'(c_t) \left(\frac{\partial F(K_t, N)}{\partial K_t} - \delta - \rho \right) \quad (5.6)$$

qui équivaut à

$$\frac{\dot{c}_t u''(c_t) c_t}{c_t u'(c_t)} = \frac{\partial F(K_t, N)}{\partial K_t} - \delta - \rho$$

Donc le modèle se réduit à la contrainte (5.5) et les équations

$$\frac{\dot{c}_t}{c_t} = \sigma \left(\alpha \left(\frac{K_t}{\xi_t N} \right)^{\alpha-1} - \delta - \rho \right) \quad (5.7)$$

$$\lim_{t \rightarrow +\infty} K_t u'(c_t) e^{-\rho t} = 0 \quad (5.8)$$

Dans la première équation, $\alpha \left(\frac{K_t}{\xi_t N} \right)^{\alpha-1} - \delta$ représente la productivité marginale nette du capital, c'est-à-dire la production supplémentaire nette de la dépréciation issue d'une unité de capital supplémentaire. L'équation signifie que lorsque la productivité marginale nette du capital $\alpha \left(\frac{K_t}{\xi_t N} \right)^{\alpha-1} - \delta$ est supérieure à ρ , le taux d'escompte subjectif, la consommation augmente dans le temps. Inversement, lorsque la productivité marginale nette est inférieure à ρ , la consommation diminue. Intuitivement, le consommateur choisit de consommer moins aujourd'hui que demain si le rendement d'une unité épargnée et investie aujourd'hui (consommation supplémentaire demain) est supérieur au taux d'escompte s'appliquant à cette consommation de demain. Dans ce cas sa consommation croît.⁴

On remarque que cette grandeur $\alpha \left(\frac{K_t}{\xi_t N} \right)^{\alpha-1} - \delta$ s'interprète comme le *taux d'intérêt en autarcie*. En effet, dans un marché du capital concurrentiel en autarcie, le taux d'intérêt d'équilibre doit être égal à la productivité marginale du capital, le taux maximal que les producteurs peuvent supporter. Comme par hypothèse la fonction de production a des rendements marginaux décroissants (productivité marginale décroissante), un pays ayant un stock de capital faible aura une productivité du capital élevée et donc un taux d'intérêt autarcique élevé. L'équation d'Euler implique qu'en autarcie ce pays épargnera initialement beaucoup pour consommer plus dans le futur. Au fur et à mesure que le pays accumulera du capital, la productivité marginale et avec elle l'incitation à épargner baissera. La vitesse de décroissance dépend de l'élasticité de substitution intertemporelle σ : plus elle est élevée, plus un changement de taux d'intérêt aura un impact fort sur la consommation.

La deuxième équation, appelée condition de transversalité, a l'intuition suivante. Tant que l'utilité marginale de la consommation est positive, il n'y a aucun intérêt à laisser un stock de capital terminal positif, plutôt que de consommer plus. Dès que cette utilité marginale vaut zéro (satiété), le

4. Blanchard & Fischer (1989) montrent que cette équation d'Euler est équivalente à l'égalité entre le taux marginal de substitution entre consommation en t et $t + dt$, et le taux marginal de transformation entre production en t et production en $t + dt$. Cette équivalence est connue sous le nom de règle de Keynes-Ramsey.

consommateur est indifférent entre consommer et laisser un stock de capital terminal non nul.

Nous allons maintenant résoudre le modèle à l'état stationnaire. Notons sans indice de temps les valeurs d'état stationnaire des variables (par exemple, c est la valeur d'état stationnaire de c_t). Comme dans le modèle de Solow avec croissance de la productivité, il est utile de redéfinir plusieurs variables en "unités de travail efficace", c'est-à-dire de les diviser par la quantité de "travail efficace" $\xi_t N$. Appelons capital par unité de travail efficace la grandeur $\tilde{k}_t = \frac{K_t}{\xi_t N}$ et revenu par unité de travail efficace la grandeur $\tilde{y}_t = \frac{Y_t}{\xi_t N}$. La fonction de production se réécrit

$$\tilde{y}_t = (\tilde{k}_t)^\alpha$$

On peut ainsi réécrire le modèle formé par (5.5), (5.7) et (5.8) en unités de travail efficace. Ce modèle admet un état stationnaire auquel le capital par unité de travail efficace \tilde{k}_t et la consommation par unité de travail efficace \tilde{c}_t sont constants (donc un sentier de croissance équilibré du modèle initial où les variables endogènes k_t et c_t croissent au taux g). On peut réécrire (5.7) à cet état stationnaire :

$$g = \sigma(\alpha \tilde{k}^{\alpha-1} - \delta - \rho) \Leftrightarrow \tilde{k} = \left(\frac{\alpha}{\rho + \frac{g}{\sigma} + \delta} \right)^{\frac{1}{1-\alpha}}$$

Donc à l'état stationnaire autarcique la productivité marginale nette du capital, égale au taux d'intérêt autarcique, s'écrit

$$r^a = \rho + \frac{g}{\sigma} \quad (5.9)$$

On peut montrer que le système dynamique admet un seul état stationnaire satisfaisant (5.5), (5.7) et (5.8) et qu'il n'existe qu'une trajectoire de point selle menant à cet équilibre⁵.

En conclusion, à l'état stationnaire autarcique, un pays ayant une plus forte croissance de la productivité aura un taux d'intérêt autarcique plus élevé. Si un tel pays s'ouvre aux flux de capitaux internationaux, on s'attend à ce qu'il importe des capitaux des pays ayant une croissance de la productivité moins forte.

5.1.2 Equilibre dans une petite économie ouverte

On considère maintenant le cas d'une petite économie ouverte, c'est-à-dire d'un pays trop petit pour que les variables nationales affectent les variables

5. Voir Blanchard & Fischer (1989) pp 45-47 et 75-76.

étrangères, qui seront considérées exogènes dans ce qui suit. En particulier, le taux d'intérêt d'équilibre mondial r est pris comme donnée par le petit pays.

L'équilibre sur le marché du capital implique que

$$\alpha \tilde{k}_t^{\alpha-1} - \delta = r$$

Avec un taux d'intérêt constant, le stock de capital en unités de travail efficace doit être constant. On le note $\tilde{k}(r)$. Par conséquent, la production par unité de travail efficace doit aussi être constante. On la note $\tilde{y}(r)$. Ainsi

$$\tilde{k}(r) = \left(\frac{\alpha}{r + \delta} \right)^{\frac{1}{1-\alpha}} \Rightarrow \tilde{y}(r) = (\tilde{k}(r))^\alpha$$

Donc Y_t et K_t croissent uniquement sous l'effet de la croissance de la productivité, au taux constant g . Le stock de capital atteint un niveau plus élevé qu'à l'état stationnaire d'autarcie quand $r^a > r$, et plus faible sinon.

En utilisant (5.7) et l'égalité entre taux d'intérêt et productivité marginale nette du capital on obtient

$$\frac{\dot{c}_t}{c_t} = \sigma(r - \rho)$$

ou encore, en utilisant (5.9),

$$\frac{\dot{c}_t}{c_t} = g + \sigma(r - r^a)$$

La consommation croît à un taux constant tout au long du processus d'accumulation de capital. Ce taux est plus élevé que la croissance de la production si le taux d'intérêt mondial est supérieur au taux autarcique, et moins élevé sinon. Ainsi, le marché financier international permet une déconnexion entre la croissance de la production et de la consommation, sauf dans le cas particulier où $r^a = r$.

La dynamique du solde courant dépend de ce différentiel de croissance. Notons B_t la position extérieure nette du pays, et définissons la richesse nette W_t comme

$$W_t = B_t + K_t$$

en notant \tilde{w}_t la richesse nette par unité de travail efficace.

On peut écrire simplement

$$\dot{\tilde{w}}_t = (r - g)\tilde{w}_t + (1 - \alpha)\tilde{y}(r) - \tilde{c}_t$$

Le terme $(r-g)\tilde{w}_t$ décrit la croissance de l'épargne existante, qui croît au taux r , mais qui exprimée en unités de travail efficace décroît avec la croissance de la productivité. Le terme $(1-\alpha)\tilde{y}(r) - \tilde{c}_t$ décrit l'épargne nette (par unité de travail efficace) disponible pour le marché financier international. Celle-ci est égale au revenu total moins la part rémunérant le capital national, $\alpha\tilde{y}(r)$, moins la consommation⁶.

On peut utiliser le modèle pour décrire l'évolution de la richesse nette nationale et de la position extérieure nette (donc indirectement du solde courant). Faisons l'hypothèse que le reste du monde a atteint l'état stationnaire et croît au taux du progrès technique g^w . Le monde étant une économie fermée, et le pays considéré étant petit, ceci implique que $r = \rho + \frac{g^w}{\sigma}$ de sorte que

$$\frac{\dot{c}_t}{c_t} = g^w$$

Dans tous les pays, la consommation du pays croît au taux g^w indépendamment de la croissance de la production nationale. La différence de croissance de productivité entre le pays considéré et le reste du monde dictera seulement la vitesse à laquelle le pays accumulera des dettes ou des créances sur le reste du monde.

On peut montrer après des calculs fastidieux⁷ (mais ici on admettra) que lorsque $r > \max g, g^w$ la consommation par unité de travail efficace s'écrit :

$$\tilde{c}_t = (r - g^w) \left(\tilde{w}_t + \frac{(1-\alpha)\tilde{y}}{r-g} \right)$$

On remarque que le terme entre parenthèses représente la richesse totale⁸, à savoir la somme de la richesse nette (financière) et la somme actualisée future des revenus du travail (capital humain). Cette relation dit que le consommateur consomme une part constante de cette richesse totale, quoique cette richesse totale varie au cours du temps avec l'accumulation de capital et l'évolution de la position extérieure nette.

On en déduit

$$\dot{\tilde{w}}_t = \sigma(r - r^a) \left(\tilde{w}_t + \frac{(1-\alpha)\tilde{y}}{r-g} \right)$$

6. On rappelle qu'à l'optimum de production en concurrence parfaite avec une fonction Cobb-Douglas l'exposant du facteur de production (ici α) est égal à la part de la rémunération de ce facteur dans le revenu national.

7. Par exemple on peut conjecturer que \tilde{c}_t prend une forme affine en \tilde{w}_t et déduire les coefficients de l'expression de \tilde{w}_t et de l'équation précédente.

8. Rappelons que si la variable x_t croît au taux constant γ , la somme actualisée à l'infini des x_t s'écrit $\int_0^{+\infty} e^{-rs} x_s ds = \frac{x_0}{r-\gamma}$.

En résolvant cette équation différentielle affine en \tilde{w}_t et en remarquant que $\tilde{b}_t = \tilde{w}_t - \tilde{k}(r)$ on obtient la position extérieure nette exprimée en unités de travail efficace

$$\tilde{b}_t = \left(\tilde{w}_0 + \frac{(1-\alpha)\tilde{y}}{r-g} \right) e^{\sigma(r-r^a)t} - \frac{(1-\alpha)\tilde{y}}{r-g} - \tilde{k}(r)$$

Comme r est constant dans le temps, et que les créances et les dettes croissent au même taux, la position extérieure nette ne varie que sous l'effet des soldes courants accumulés (pas d'effets de valorisation mentionnés dans le chapitre introductif). Ainsi, on peut déduire directement le solde courant en unités de travail efficace de l'évolution de \tilde{b}_t :

$$SC_t = \dot{B}_t \Rightarrow \tilde{s}c_t = g\tilde{b}_t + \dot{\tilde{b}}_t$$

d'où

$$\tilde{s}c_t = g^w \left[\left(\tilde{w}_0 + \frac{(1-\alpha)\tilde{y}}{r-g} \right) e^{\sigma(r-r^a)t} - \frac{(1-\alpha)\tilde{y}}{r-g} - \tilde{k} \right] - g \left[\frac{(1-\alpha)\tilde{y}}{r-g} + \tilde{k} \right]$$

Nous pouvons finalement énoncer les prédictions du modèle. Celles-ci dépendent de la différence entre le taux d'intérêt mondial et le taux d'intérêt autarcique :

- lorsque $r^a < r$, c'est-à-dire quand $g < g^w$, le pays tend à accumuler des excédents courants et devient créancier du reste du monde. en d'autres termes il exporte du capital.
- lorsque $r^a = r$, c'est-à-dire quand $g = g^w$, le solde courant dépend des conditions initiales, autrement dit des valeurs de \tilde{w}_0 et du stock de capital autarcique \tilde{k} . Si ces deux valeurs sont proches, par exemple parce qu'au moment de l'ouverture la position extérieure nette est zéro, alors la balance courante reste constamment équilibrée.
- lorsque $r^a > r$, c'est-à-dire quand $g > g^w$, le pays tend à accumuler des déficits courants et devient débiteur du reste du monde, en d'autres termes il importe du capital.

En conclusion, le modèle prédit une corrélation négative entre croissance de la productivité et solde courant. On notera qu'en raisons d'hypothèses simplificatrices ce modèle n'est pas adapté pour décrire la dynamique précise des mouvements de capitaux. Dans ce modèle un pays passant de l'autarcie totale à l'ouverture totale atteint immédiatement le stock de capital et la position extérieure nette optimaux et cohérents avec la condition de transversalité, ce qui peut impliquer de forts mouvements instantanés de la position nette. En revanche, le modèle est adapté pour décrire la trajectoire générale des mouvements de capitaux entre pays hétérogènes.

5.2 Déséquilibres structurels mondiaux

Le modèle de croissance néoclassique prédit donc que les pays à faible stock de capital et/ou à forte croissance de la productivité reçoivent plus de capitaux. Robert Lucas, dans un article de 1990 intitulé "Why Doesn't Capital Flow From Rich To Poor Countries?", relève que les différences de productivité marginale du capital entre pays riches (Nord) et pauvres (Sud) devraient causer des flux considérables de capitaux du Nord vers le Sud. Dans le modèle que nous venons de considérer, le rapport entre productivité marginale de deux pays appelés N et S *de même technologie* est égal à $\left(\frac{y^N}{y^S}\right)^{1-\frac{1}{\alpha}}$. En notant que le PIB/habitant des Etats-Unis (N) en 1988 était 15 fois supérieur à celui de l'Inde (S) et supposant une part du capital dans le revenu plausible égale à 0,4, Lucas trouve un ratio des productivités marginales égal à 58! Si on acceptait littéralement le modèle, les Etats-Unis devraient envoyer quasiment tout leur capital en Inde!

Lucas lui-même et d'autres économistes ont répondu à ce paradoxe en invoquant deux types de raisons. Premièrement, ce calcul ne tient pas compte des différences technologiques entre pays, ni d'autres caractéristiques des pays affectant la croissance et donc le PIB/habitant, en particulier le capital humain et les institutions politiques (confiance dans la justice, respect des droits de propriété, stabilité politique notamment). Deuxièmement, il néglige aussi différentes distorsions dans l'allocation du capital entre pays qui affectent le calcul du rendement, comme la fiscalité, les coûts de transaction, les différences de prix des biens d'équipement, éventuellement la corruption ou le risque d'expropriation.

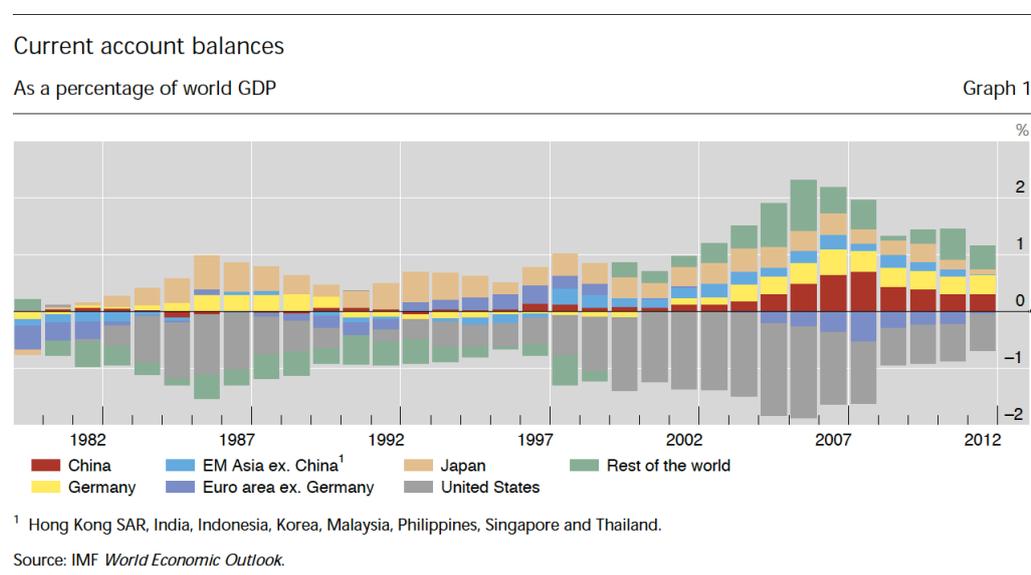
Par ailleurs, tant le modèle précédent (avec la condition de transversalité) que le modèle à deux périodes du chapitre 4 prédisent que les déficits courants doivent être suivis par des excédents courants, et inversement. Nous allons voir dans cette section que cela est loin d'être le cas, et que depuis 30 ans plusieurs pays affichent des déséquilibres courants structurels. Cette section aura pour but de documenter ce phénomène et d'apporter des explications, dont certaines peuvent intégrer notre modèle de croissance.

5.2.1 Un aperçu des déséquilibres mondiaux dans les données

Les Figures 5.1 et 5.2 présentent les déséquilibres dans le monde et dans la zone euro, respectivement. La Figure 5.1 montre la persistance du déficit courant américain et de l'excédent courant japonais depuis le début des années 1980, ainsi que la persistance des excédents courants chinois et allemand

depuis le milieu des années 1990 et le début des années 2000, respectivement. La Figure 5.2 montre que ce phénomène existe aussi dans la zone euro, où l'Allemagne, les Pays-Bas, l'Autriche, la Finlande enregistrent des excédents courants depuis le début des années 2000 tandis que l'Espagne, l'Italie, la Grèce et le Portugal enregistrent des déficits courants. La France est un des rares pays dont le solde courant change de signe pendant la période.

FIGURE 5.1 – Soldes courants exprimés en pourcentage du PIB mondial.

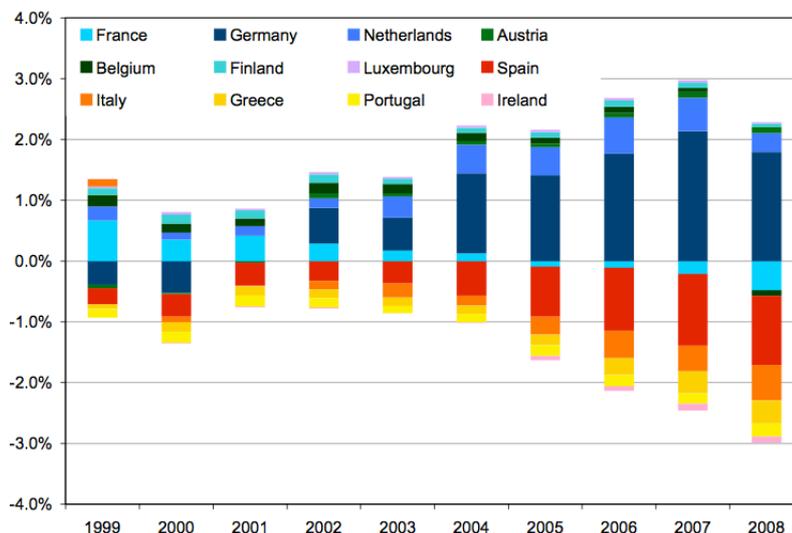


Source : FMI.

Ces Figures appellent deux observations. Premièrement, la persistance des signes des soldes courants semble aller à l'encontre de l'approche intertemporelle du solde courant vu au chapitre précédent. Deuxièmement, le fait que plusieurs pays émergents (Chine, autres pays émergents asiatiques, pays exportateurs de pétrole) soit en excédent courant tandis que les Etats-Unis, le Royaume-Uni et d'autres pays développés sont en déficit courant semble aller à l'encontre de la théorie que nous venons de présenter et confirmer le paradoxe de Lucas.

Pour aller plus loin, on peut étudier plus en détail les balances des paiements. La Figure 5.3 montre comment les excédents chinois sont la contrepartie des déficits américains, mais aussi comment les Etats-Unis émettent essentiellement des obligations, tandis que la Chine accumule essentiellement

FIGURE 5.2 – Soldes courants exprimés en pourcentage du PIB de l'eurozone.



Source : FMI.

des réserves de change.

Enfin, la Figure 5.4 confirme les doutes qu'on peut avoir sur les prédictions du modèle de Ramsey. Elle montre une corrélation négative entre croissance de la productivité et entrée de capitaux dans 68 pays émergents entre 1980 et 2000.

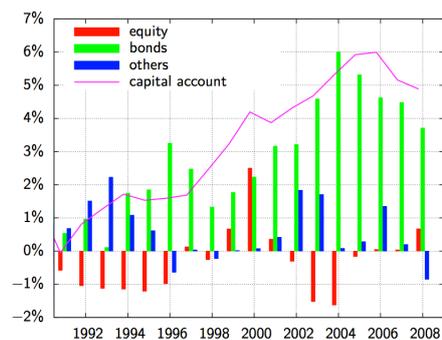
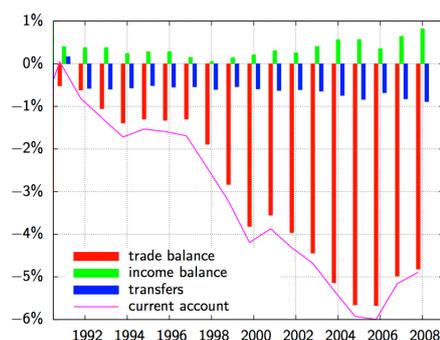
Nous allons donc compléter nos outils d'analyse. Dans la suite de ce chapitre, nous allons mentionner différentes explications possibles pour ces déséquilibres courants. Puis nous présenterons un modèle formalisant l'une d'entre elles (fondée sur la pénurie d'actifs financiers nationaux) dans la fin de ce chapitre. Une autre explication fondée sur l'accumulation de réserves pour motif d'assurance sera étudiée en PC.

5.2.2 Explications possibles des déséquilibres structurels

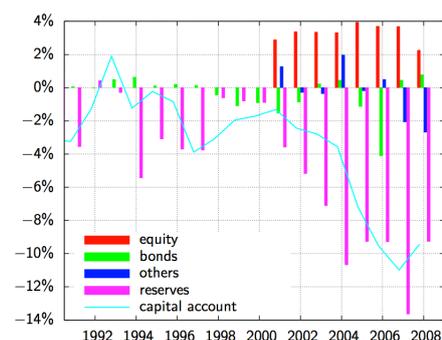
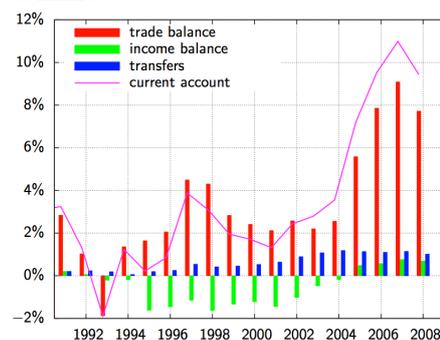
Le modèle néoclassique prédit que les capitaux se déplacent vers les pays ayant des taux d'intérêt autarciques élevés, et que les taux autarciques dépendent principalement de la croissance de la productivité. Les explications actuelles des déséquilibres structurels conservent la première partie de cette prédiction tout en modifiant la seconde.

FIGURE 5.3 – Eléments des comptes courant (gauche) et financier (droite) en % du PIB aux Etats-Unis et en Chine de 1992 à 2008.

Etats-Unis



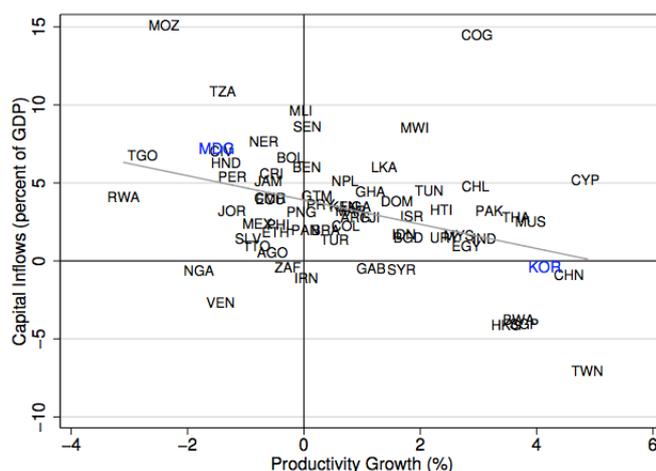
Chine



Source : FMI.

L'ancien gouverneur de la Banque Centrale américaine, l'économiste Ben Bernanke, a formulé dans un discours célèbre (Bernanke 2005) *l'hypothèse d'excès d'épargne* ("savings glut hypothesis"). Selon cette hypothèse, c'est l'existence d'une épargne mondiale abondante, notamment dans plusieurs pays émergents, qui peut expliquer la croissance du déficit courant américain. Cette hypothèse a le mérite d'expliquer l'accroissement substantiel de ce déficit depuis la fin des années 90 et d'être cohérente avec la tendance à la baisse des taux d'intérêt réels (voir Figure 5.5). Elle tranche avec les analyses précédentes du déficit courant américain qui insistait sur des facteurs purement nationaux, notamment le rôle du déficit budgétaire. On pense à l'hypothèse des déficits jumeaux : un déficit budgétaire réduit l'épargne nationale et peut contribuer au déficit courant, d'autant plus facilement aux Etats-Unis avec la forte demande internationale pour les obligations d'Etat

FIGURE 5.4 – Croissance de la productivité moyenne entre 1980 et 2000 et entrées de capitaux dans 68 pays émergents.



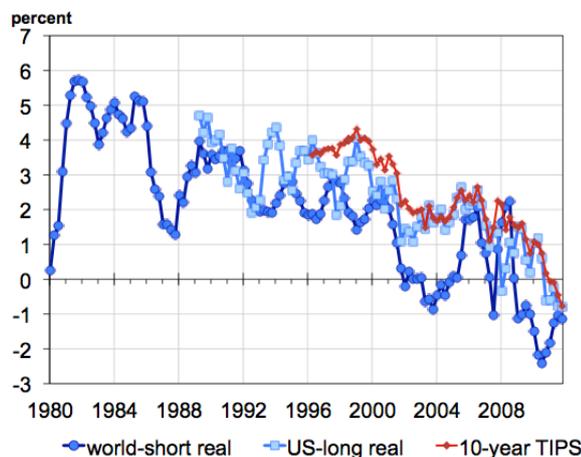
Source : Gourinchas & Jeanne (2013).

américaines. Cependant, cette hypothèse des déficits jumeaux n'est convaincante que jusqu'au milieu des années 90, lorsque le budget fédéral américain devient proche de l'équilibre.

Dans le même discours, Bernanke mentionne plusieurs raisons pouvant expliquer cet excès d'épargne : la hausse de l'épargne des pays asiatiques et exportateurs de pétrole et l'accumulation de réserves par leurs BC ; le vieillissement de la population dans les pays riches en déficit courant, ponctionnant l'épargne de ces pays ; la forte hausse du prix du pétrole pendant les années 2000.

D'autres explications indépendantes de l'hypothèse d'excès d'épargne ont été avancées, liées aux stratégies de croissance par les exportations et de sous-évaluation de la monnaie, ou encore l'anticipation de forts gains de productivité futurs aux Etats-Unis (suivant le mécanisme décrit à la fin du chapitre 4). Cependant ces explications sont moins convaincantes. D'une part, s'il est vrai que certains pays asiatiques ont accumulé des réserves pour maintenir une monnaie sous-évaluée (par un mécanisme simple qui sera expliqué au chapitre 7 consacré aux régimes de change fixe), l'abandon du taux de change fixe par la Chine en 2005 pour un change flottant administré n'a que modérément réduit le déficit courant des Etats-Unis auprès de la Chine. D'autre part, le financement du déficit courant américain principalement par des obligations d'Etat plutôt que des titres émis par les entreprises privées

FIGURE 5.5 – Taux d'intérêt réels mondial et américain.



Source : FMI et OCDE, cités par Gourinchas & Jeanne (2013). Le taux d'intérêt réel désigne le taux d'intérêt de référence (dette publique) corrigé de l'inflation. Les TIPS sont des obligations d'Etat américaines dont le principal est indexé sur l'inflation, donc leur rendement mesure un taux d'intérêt réel.

contredit l'hypothèse de croyance dans de forts gains de productivité futurs.

Nous allons donc passer en revue différentes explications de l'excès d'épargne⁹.

Accumulation de réserves Comme le mentionne l'analyse de Bernanke, une grande partie de l'accumulation d'actifs finançant le déficit courant américain a pris la forme de hausse des avoirs de réserve des Banques Centrales (notamment de la Chine). Dans un grand nombre de modèles macroéconomiques, y compris le modèle de Ramsey, la distinction entre épargne privée et épargne publique a peu de sens, en raison de l'équivalence ricardienne¹⁰.

9. Pour une présentation plus détaillée, voir la section 4 de Gourinchas & Rey (2014) et les références citées dans cette section. On peut noter que la plupart des arguments présentés viennent de travaux récents, et que ce sujet fait partie du programme de recherche actuel en macroéconomie internationale.

10. Selon l'hypothèse d'équivalence ricardienne, toute variation de l'épargne publique est compensée par une variation égale mais de signe opposé de l'épargne privée. Par exemple une hausse de la dette publique entraîne une hausse de l'épargne privée d'un montant équivalent en prévision des hausses d'impôt futures nécessaires pour payer cette dette. En théorie, cette équivalence ricardienne requiert de fortes hypothèses : fiscalité non distortive, marchés financiers complets, horizon temporel des ménages infini. Mais ces hypothèses servent souvent à simplifier les modèles quand la composition de l'épargne n'est pas l'objet d'analyse, comme c'est le cas dans le modèle de Ramsey de base. Voir également le polycopié du cours ECO432, section 3.4.5.

Il faut donc proposer d'autres modèles pour expliquer ce phénomène.

Plusieurs raisons peuvent expliquer l'accumulation de réserves par le secteur public. Premièrement, la crise asiatique de 1997 a causé un traumatisme. D'une part les Banques Centrales des pays touchés par la crise n'ont pas eu la capacité de défendre les cours de leurs monnaies, ni de renverser les départs rapides de capitaux étrangers. D'autre part, les autres pays asiatiques ont pu observer de près à quel point les prêts d'urgence du FMI, quoique utiles pour maintenir le fonctionnement des secteurs bancaires des pays touchés, pouvaient être contraignants sur la politique économique. Deuxièmement, le maintien d'un change fixe par la Chine, comme cité précédemment, a contribué à l'accumulation de réserves. Troisièmement, certains motifs d'épargne privée (cités dans les arguments ci-dessous) ont pu guider l'épargne publique : ainsi certains économistes estiment que la BC chinoise a servi d'intermédiaire financier, guidant l'épargne nationale vers des actifs américains, dans un pays où l'accès des ménages aux marchés financiers étrangers est très restreint. Enfin, comme les agents privés, les BC asiatiques ont pu accumuler des réserves par manque d'actifs financiers nationaux, leur choix se portant effectivement sur les titres les plus liquides, comme des obligations d'Etat américaines.

Démographie Plusieurs pays ayant des excédents courants élevés, comme l'Allemagne, le Japon ou la Chine, ont en commun un vieillissement rapide de leurs populations. Ainsi, le taux de dépendance, mesurant le ratio des plus de 65 ans sur la population des 20-65 ans est-il élevé en Allemagne et au Japon : 33,4 % et 38,4% contre 21% aux Etats-Unis en 2010. De plus, les projections prédisent une forte augmentation de ce taux d'ici 2050 : 62,4% pour l'Allemagne, 76,4% pour le Japon, 45,4% pour la Chine contre 39,5% pour les Etats-Unis. Evidemment, ces évolutions démographiques incitent à accumuler de l'épargne retraite.

En théorie, le rôle de la démographie sur l'épargne est complexe et dépend des détails du changement démographique. Selon le modèle néoclassique (voir PC) une plus forte croissance de la population implique un taux d'intérêt autarcique plus élevé. Mais dès qu'on permet à un modèle d'avoir plusieurs générations imbriquées, comme le modèle de cycle de vie de Modigliani, la hausse de la population jeune et épargnante implique une plus forte épargne et un taux autarcique plus faible. Le modèle de Ferrero (2010) à deux générations prédit qu'un déclin de la mortalité, à croissance de la population donnée, réduit le taux d'intérêt autarcique. Le modèle de Coeudacier, Guibaud & Jin (2012) à trois générations montre qu'une baisse des taux d'intérêt réels mondiaux implique une baisse de l'épargne dans les pays riches, mais une hausse de l'épargne dans les pays émergents. Le mécanisme

passer par l'existence de contraintes de crédit dans les pays émergents, dont les habitants empruntent ne peuvent pas emprunter plus étant jeune, et doivent épargner plus étant dans l'âge moyen.

Épargne de précaution On appelle épargne de précaution une épargne visant à assurer l'individu contre des risques contre lesquels il n'existe pas suffisamment d'assurance - par opposition avec l'épargne visant à transférer du pouvoir d'achat du présent vers le futur. On pense au risque de chômage, de maladie, d'invalidité, éventuellement au risque de faillite pour un entrepreneur. Naturellement, la protection sociale et les assureurs privés couvrent une partie de ces risques : l'ampleur de l'épargne de précaution dépend du degré de couverture qu'offre la protection sociale et du degré de sophistication du marché de l'assurance dans le pays.

Selon cet argument, les pays peu abondants en capital sont aussi des pays dans lesquels il existe peu de moyens pour que les agents s'assurent contre ce type de choc idiosyncratique, expliquant des taux d'épargne plus élevés. On peut aussi voir dans cet argument un avantage comparatif des pays riches dans la production d'actifs financiers peu risqués (puisque les pays riches offrent plus de mutualisation de ces risques). Cet argument est lié à l'argument de sous-développement financier présenté ci-dessous.

Sous-développement financier Selon cet argument, l'excès d'épargne disponible pour financer les déficits courants des pays développés viendrait d'un manque d'actifs financiers liquides dans les pays émergents. La pénurie d'actifs liquides y cause une hausse du prix des actifs et donc une baisse du taux d'intérêt d'équilibre, conduisant à une exportation de capitaux vers les pays développés. Nous allons développer cet argument ci-dessous à l'aide d'un modèle dû à Caballero, Farhi & Gourinchas (2008).

Selon une variante de cet argument, l'existence de contraintes de crédit plus fortes dans les pays émergents devient une source d'avantage comparatif : les pays avec peu de contraintes de crédit se spécialisent dans les secteurs intensifs en capital. Ce mécanisme peut renverser le sens prévu des échanges de capitaux par le modèle de croissance néoclassique.

5.2.3 Un modèle avec marchés financiers incomplets

Nous détaillons ici une des explications candidates dans un modèle qui emprunte à Caballero et al. (2008)¹¹.

11. Pour plus de détails voir la présentation de Gourinchas & Rey (2014), section 4.1.

L'idée de ce modèle est que des pays ayant des marchés financiers peu développés ont un motif supplémentaire pour demander des actifs financiers étrangers, et que cette demande peut renverser le sens des échanges financiers prédit par le modèle néoclassique.

Comme précédemment, la demande d'actifs est motivée par le désir de lisser la consommation face à des flux de revenus volatiles. Pour capturer l'idée d'une capacité d'assurance est limitée tout en gardant un modèle simple, on suppose que le revenu du consommateur est certain jusqu'à la mort, mais que l'incertitude porte sur la date de mort. Par ailleurs, l'individu ne peut pas engager le revenu des générations futures comme collatéral.

Le modèle est en temps continu. On suppose une population de taille 1 dont une fraction $0 < \theta < 1$ meurt à chaque instant et une fraction θ naît à chaque instant. On note $c(s, t)$ la consommation d'un individu né en s ("cohorte s ") à la date $t \geq s$. On note $z(s, t)$ le revenu du travail des individus de cohorte s perçu à la période t , et $w(s, t)$ leur patrimoine financier à la même date. Pour simplifier, on suppose que les agents sont neutres au risque.

Autarcie financière

Considérons un actif financier sans risque au taux r_t et un deuxième actif de type rente viagère. Ce second actif consiste en un échange entre le paiement d'une rente par un assureur à un individu jusqu'à sa mort, contre la cession du capital de l'individu à l'assureur après sa mort. On suppose un marché d'assurance parfaitement concurrentiel, de sorte que la rente est égale à θ (le taux de mortalité) par unité d'épargne. En effet, les assurés sont identiques et la population constante, donc une fraction θ de l'épargne totale revient à l'assureur à chaque période et celui-ci peut offrir jusqu'à θ par unité assurée.

Les consommateurs maximisent

$$E_t \left[\int_t^\infty e^{-\rho(u-t)} \ln(c(s, u)) du \right] = \int_t^\infty e^{-(\rho+\theta)(u-t)} \ln(c(s, u)) du$$

où on a utilisé les hypothèses très commodes de probabilité de mort θ constante à chaque période et de neutralité au risque. Le choix d'utilité logarithmique fait que l'effet substitution et l'effet revenu du taux d'intérêt s'annulent. (Cela simplifie les calculs mais n'est pas nécessaire au modèle.) Remarquons que cette fonction d'utilité représente le cas limite de la fonction d'utilité CRRA utilisée plus haut lorsque $\sigma \rightarrow 1$.

On normalise le prix du bien de consommation à 1. La contrainte de budget s'écrit

$$c(s, t) = z(s, t) + (r_t + \theta)w(s, t) + \frac{dw(s, t)}{dt}$$

Comme dans le modèle de Fisher à deux périodes et comme dans le modèle de Ramsey, la solution du problème du consommateur vérifie l'équation d'Euler

$$\frac{\dot{c}(s, t)}{c(s, t)} = r_t - \rho$$

Avec utilité logarithmique, on sait que la consommation optimale est une fonction linéaire du capital financier et humain (revenu permanent)

$$c(s, t) = (r_t + \theta) (w(s, t) + h(s, t))$$

avec $h(s, t) = \int_t^\infty e^{\int_t^u (r_v + \theta) dv} z(s, u) du$ le capital humain.

Cette équation décrit le choix de consommation optimal pour tous les individus de même cohorte s . En sommant sur l'ensemble des cohortes, on obtient la consommation agrégée

$$\begin{aligned} C_t &= \int_t^\infty c(s, t) \theta e^{-\theta(t-s)} ds = (\rho + \theta) (W_t + H_t) \\ W_t &= \int_t^\infty w(s, t) \theta e^{-\theta(t-s)} ds \\ H_t &= \int_t^\infty h(s, t) \theta e^{-\theta(t-s)} ds \end{aligned}$$

en remarquant que la mesure de chaque cohorte s à la date t s'écrit $\theta e^{-\theta(t-s)} ds$. Comme précisé en introduction, H_t prend en compte le revenu du travail futur des cohortes vivantes, mais pas des cohortes à naître dans le futur. Les cohortes vivantes ne peuvent pas engager ces revenus pour emprunter, contrairement au modèle de Ramsey.

On déduit de la contrainte de budget

$$\frac{dW(s, t)}{dt} = r_t W(s, t) + Z(s, t) - C_t$$

avec $Z(s, t) = \int_t^\infty z(s, t) \theta e^{-\theta(t-s)} ds$

L'évolution de H_t au cours du temps dépend de la distribution du revenu en fonction de l'âge, $\{z(s, t)\}$. Considérons un cas particulier du modèle où

$$z(s, t) = \frac{\phi + \theta}{\theta} e^{-\phi(t-s)} Z_t, \phi \geq 0$$

Cette distribution arbitraire a la propriété suivante : lorsque $\phi = 0$, tous les individus reçoivent le même revenu quel que soit leur âge, tandis que lorsque $\phi \rightarrow \infty$ tout le revenu est perçu par la génération qui vient de naître. L'intérêt

de cette hypothèse est double : ϕ paramètre le degré de désynchronisation entre revenu et consommation (donc le motif de demande d'actifs), et $\frac{dH(s,t)}{dt}$ prend une forme très simple

$$\frac{dH(s,t)}{dt} = (r_t + \theta + \phi)H(s,t) - Z_t$$

On suppose que la production utilise du travail et du capital suivant la fonction donnée par (5.2) :

$$Y_t = (K_t)^\alpha (\xi_t N_t)^\alpha$$

On suppose aussi pour simplifier que le taux de dépréciation du capital est nul.

Finalement on fait une hypothèse importante sur la taille du secteur financier. On suppose qu'elle est égale à δY_t de sorte que $Z_t = (1 - \delta)Y_t$. Le paramètre δ va jouer un rôle important puisque nous cherchons étudier les flux financiers entre deux économies de δ différents.

En autarcie financière, $W_t = K_t$ et $Y_t = r_t K_t + Z_t$. On en déduit facilement $r_t K_t = \delta Y_t$. On peut alors montrer qu'à l'état stationnaire

$$(r^a - \delta(g + \rho + \theta))(r^a + \theta + \phi - g) = (1 - \delta)r^a(\rho + \theta)$$

On note que :

- lorsque $\theta = \phi = 0$ on revient au cas néoclassique $r^a = g + \rho$ (l'élasticité de substitution intertemporelle étant 1)
- lorsque $\phi \rightarrow +\infty$ on obtient $r^a = \delta(g + \rho + \theta)$
- pour des valeurs intermédiaires de ϕ et θ on peut montrer que r^a est compris entre $g + \rho - \phi$ et $g + \rho + \theta$.

Dans une économie avec δ plus faible le taux d'intérêt autarcique est plus faible qu'ailleurs, et plus faible que dans le cas néoclassique. Ainsi un pays avec un fort g mais un faible δ pourrait *prêter* au reste du monde en économie ouverte.

Economie ouverte

On considère maintenant que le pays devient une petite économie ouverte, prêtant ou empruntant au taux d'intérêt mondial r . On se concentre sur le cas $\phi \rightarrow +\infty$, où seule la cohorte qui vient de naître touche un revenu, ce qui implique $r^a = \delta(g + \rho + \theta)$.

Le pays a maintenant accès à des actifs nationaux (K_t) et des actifs étrangers ; $W_t - K_t$ est la position extérieure nette du pays. Comme précédemment, on trouve

$$\frac{K_t}{Y_t} = \frac{\delta}{r}$$

et

$$\frac{W_t}{Y_t} = \frac{1 - \delta}{g + \rho + \theta - r}$$

La première équation décrit l'offre d'actifs nationaux par unité produite, tandis que la deuxième décrit la demande nationale d'actifs par unité produite.

On peut finalement écrire

$$\begin{aligned} \frac{B_t}{Y_t} &\equiv \frac{W_t - K_t}{Y_t} = \frac{\delta(r - r^a)}{r(r^a - \delta r)} \\ \frac{SC_t}{Y_t} &= g \frac{B_t}{Y_t} \end{aligned}$$

Un pays où $g + \rho > r$ mais $\delta(g + \rho + \theta) < r$ connaît une hausse du taux d'intérêt à l'ouverture et devient prêteur net ($B_t > 0$), alors qu'il aurait été emprunteur net dans le modèle de Ramsey.

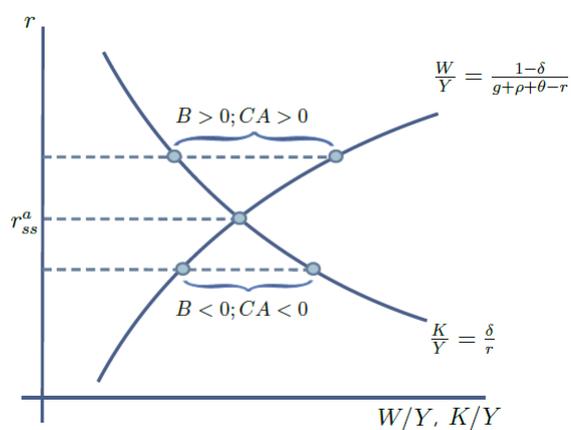
La Figure 5.6 illustre ce résultat à l'aide d'un diagramme de Metzler, qui décrit l'offre et la demande nationales d'actifs financiers en fonction du taux d'intérêt, et ainsi l'excès ou le défaut d'actifs au taux d'intérêt mondial.¹² La courbe croissante représente la demande nationale d'actifs en fonction du taux d'intérêt mondial r (exogène) : plus r est élevé, plus les nationaux épargnent et demandent d'actifs. La courbe décroissante représente l'offre nationale d'actifs en fonction de r : plus r est élevé, plus le pays utilise une technologie peu intensive en capital, et donc plus l'offre nationale d'actifs est faible. Les deux courbes se coupent au point d'ordonnée r^a : si r prend exactement cette valeur, alors $W_t = K_t$ et on revient à l'équilibre autarcique. Lorsque $r > r^a$, le pays tend vers un excédent courant ; lorsque $r < r^a$, le pays tend vers un déficit courant.

Le modèle prédit donc qu'un pays avec un δ faible a un taux d'intérêt autarcique faible, même lorsque sa croissance de la productivité g est élevée, et donc tend vers un déficit courant en économie ouverte.

Comment interpréter le paramètre δ ? Dans le modèle, δ gouverne l'intensité en capital optimale de la production, et donc le stock de capital et l'offre d'actifs nationaux. Des différences de δ entre pays peuvent donc capturer des différences de protection des droits des créanciers, de sophistication des actifs financiers, de corruption, de fiscalité du capital... Un δ faible capture mieux la situation de pays en développement.

12. Le prix des actifs étant lié à l'inverse de r , le diagramme représente donc de façon inhabituelle l'offre par une courbe décroissante et la demande par une courbe croissante.

FIGURE 5.6 – Diagramme de Metzler



Le diagramme de Metzler décrit l'offre (courbe décroissante) et la demande d'actifs (courbe croissante) nationales d'actifs financiers en fonction du taux d'intérêt. On note que l'offre et la demande d'actifs sont les images de la demande et de l'offre de capitaux, respectivement, dont le prix est le taux d'intérêt. Ces courbes se croisent au niveau du taux d'intérêt autarcique. A l'ouverture, le pays tend vers un excédent courant si son taux d'intérêt autarcique est inférieur au taux d'intérêt mondial, et un déficit courant sinon. Source : Gourinchas & Rey (2014, Figure 8)

5.3 Conclusion

Les prédictions du modèle théorique néoclassique sur la direction des flux financiers mondiaux, des pays développés vers les pays en développement, vont à l'encontre des observations empiriques. De plus, le monde a connu dans les trente dernières années des déséquilibres courants persistants, avec des pays comme les Etats-Unis régulièrement en déficit courant, et des pays comme le Japon mais aussi la Chine régulièrement en excédent courant.

Nous avons recensé plusieurs explications possibles pour ce phénomène. Pour la plupart, celles-ci sont liées à l'hypothèse d'excès d'épargne avancée par Ben Bernanke. Ainsi, la pénurie d'actifs financiers ou de moyens d'assurance contre les risques idiosyncratiques, la spécialisation dans les secteurs peu intensifs en capital, le vieillissement prévu de la population et la faiblesse de la protection sociale, l'accumulation de réserves pour se prémunir des retraits de capitaux et crises de change, peuvent expliquer pourquoi la Chine et d'autres pays émergents épargnent nettement plus que les Etats-Unis. Nous avons vu à la fin de ce chapitre (et reverrons en PC) que certaines explications peuvent être intégrées à un cadre d'analyse intertemporel standard.

Ce chapitre conclut notre analyse des balances courantes. Dans les chapitres à venir, nous allons étudier la formation des taux de change et leur rôle dans l'ajustement des soldes courants vers des niveaux compatibles avec les contraintes de budget intertemporelles. Enfin, nous verrons comment cet ajustement peut se faire dans les régimes de change fixe, où le taux de change ne peut plus jouer ce rôle.

Chapitre 6

Les taux de change

Ce chapitre a pour but d'étudier les déterminants du change réel et du change nominal. Dans le chapitre 4, nous avons vu que le taux de change nominal désigne la valeur d'une monnaie exprimée en unités d'une autre monnaie, tandis que le taux de change réel (TCR) exprime la valeur d'un panier de biens dans un pays en unités de paniers de biens d'un autre pays. Nous noterons S_{ijt} le taux de change nominal du pays i et $Q_{ijt} = \frac{S_{ijt}P_{it}}{P_{jt}}$ le taux de change réel du pays i , où P_{jt} désigne l'indice des prix du pays j à la date t exprimé dans la monnaie du pays j . Par nature, ces mesures sont bilatérales, mais on peut également construire des taux de change dits *effectifs* contre un groupe de pays, c'est-à-dire des moyennes des taux bilatéraux pondérées par les flux de commerce de chaque pays. Cette notion de taux effectif nous sera utile pour penser à la détermination du taux de change dans une petite économie ouverte.

6.1 Déterminants du taux de change réel

Le TCR mesure le prix des biens nationaux en unités de biens étrangers. Logiquement, le TCR est affecté par les déterminants de ses deux composantes, l'indice des prix et le change nominal. Cependant, comme le suggère le chapitre 4, le TCR joue un rôle important dans l'ajustement du solde courant, ce qui montre le besoin d'une théorie co-déterminant à la fois le TCR et ses composantes.

6.1.1 L'hypothèse de Parité de Pouvoir d'Achat

Une première hypothèse théorique sur le TCR énonce que le TCR prend une valeur d'équilibre de long terme constante.

On parle de *Parité de Pouvoir d'Achat (PPA) absolue* lorsque le prix des biens est égal dans deux pays, ou de façon équivalente que le taux de change nominal est celui qui donne exactement le même pouvoir d'achat dans les deux pays.

$$Q_{ijt} = 1 \Leftrightarrow S_{ijt} = \frac{P_{jt}}{P_{it}}$$

On parle de *PPA relative* lorsque Q_{ijt} est constant, mais pas nécessairement égal à 1 : les variations du change nominal compensent exactement les variations du différentiel d'inflation.

La théorie de la PPA, due à Cassel (1918), propose que la Parité de Pouvoir d'Achat représente l'état d'équilibre de long terme vers lequel tend l'économie. L'idée est que la PPA est un équilibre de long terme stable, en raison des mécanismes de marché qui corrigent un écart temporaire par une variation du change nominal. Supposons par exemple qu'un pays ait temporairement des prix plus élevés que leur niveau de PPA : alors la demande pour les biens de ce pays diminue, la demande pour la monnaie du pays diminue, la monnaie se déprécie, et le prix exprimé en monnaie étrangère redevient compatible avec la PPA. Une version extrême de cet argument s'applique lorsque tous les biens sont parfaitement échangeables sans aucun coût, comme dans la théorie Heckscher-Ohlin, par exemple. Les prix de chaque bien s'égalisent alors par arbitrage (un phénomène parfois appelé *Loi du prix unique*). Si cette loi s'applique à chaque bien alors les indices des prix sont égaux et on obtient la PPA absolue.

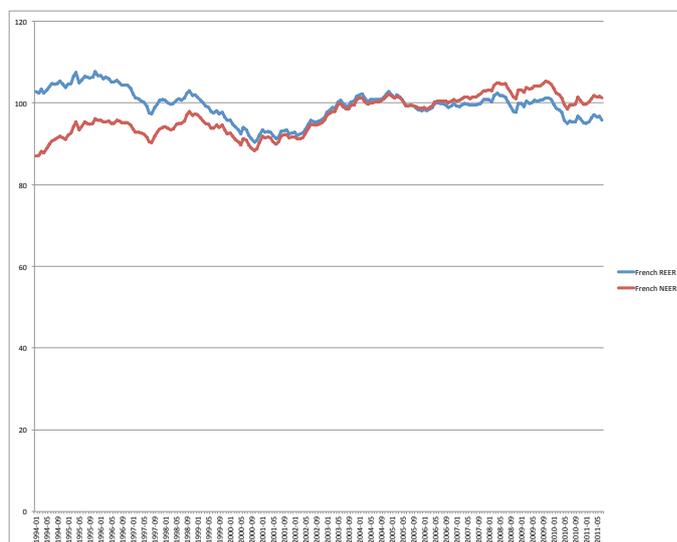
Les hypothèses de PPA absolue et relative sont-elles vérifiées dans la réalité ? Cette question a occupé un grand nombre de travaux en économie et nous allons présenter plusieurs types de tests de cette hypothèses.

Premièrement, on peut simplement regarder l'évolution d'un TCR effectif (TCRE) construit à partir des taux de change nominaux de marché et des indices des prix à la consommation des instituts statistiques. La Figure 6.1 décrit l'évolution du TCRE et du taux de change nominal effectif de la France entre 1994 et 2011, et suggère de fortes variations dans le TCRE, contrairement à l'hypothèse de PPA.

Pour aller plus loin, de nombreux travaux utilisent des séries temporelles de taux de change effectif pour tester l'hypothèse de PPA économétriquement. Pour tester l'hypothèse de PPA, on utilise un *test de racine unitaire*. Plus précisément, on suppose que le processus générateur de la série temporelle prend la forme

$$\ln Q_t - \ln Q_{t-1} = \rho \ln Q_{t-1} + \varepsilon_t$$

FIGURE 6.1 – Taux de change effectifs réel (REER) et nominal (NEER) de la France, 1994-2011.



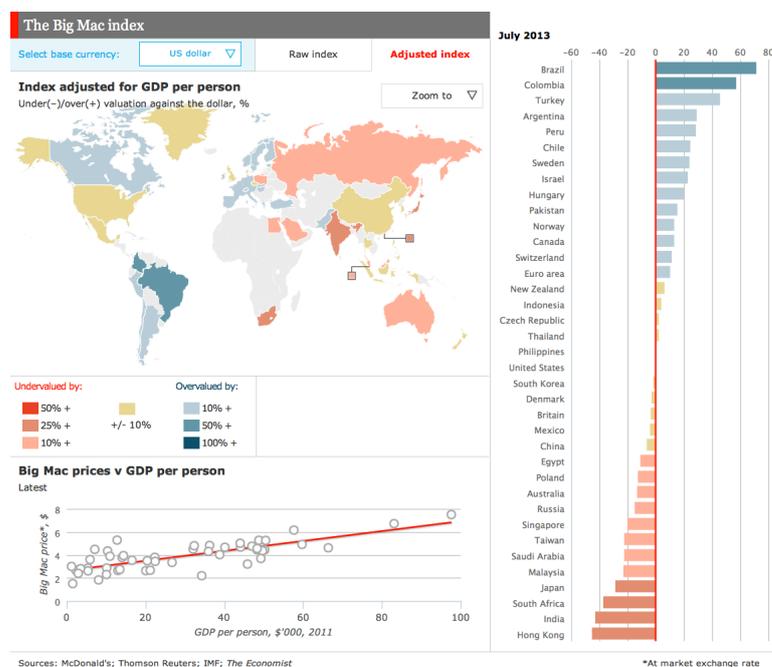
Source : Banque des Règlements Internationaux (BRI/BIS).

où ε_t est un bruit blanc d'espérance nulle et de variance $\sigma^2 > 0$.

Sous l'hypothèse $H_0 : \rho = 0$, la série $\ln Q_t$ a une racine unitaire ($1 - \rho = 1$), ce qui implique que toutes ses déviations sont permanentes, puisqu'elles déplacent l'espérance des Q_t futurs ($E_t Q_{t+1} = Q_t$). Si $H_1 : \rho < 0$ est vérifiée, $\ln Q_t$ revient à sa moyenne et la PPA relative est vérifiée. Rogoff (1996) obtient $\rho \approx -0,15$ pour un TCRE des Etats-Unis contre un groupe de pays développés. On peut déduire de son estimation la vitesse de convergence du TCRE. En effet, l'équation précédente implique que $\ln Q_t = (1 + \rho)^t \ln Q_0$ de sorte que la demi-vie du processus de convergence vers Q_0 a pour durée $T = \frac{-\ln 2}{\ln(1+\rho)}$. Dans le cas de Rogoff cela implique une demi-vie d'environ 4 ans. On conclut que la PPA relative semble vérifiée à long terme, mais pas à court terme.

Un problème pratique lié à l'utilisation des TCRE est que les indices des prix à la consommation n'utilisent pas exactement les mêmes paniers d'un pays à l'autre, et que leur composition peut changer dans le temps. Une première solution consiste à utiliser le travail du International Comparison Program (ICOP), qui mesure les prix de paniers standardisés tous les 5 ans dans de nombreux pays (177 pays en 2011). Cette solution a l'avantage d'être la plus rigoureuse, mais le désavantage d'une fréquence basse. Une deuxième

FIGURE 6.2 – Indice Big Mac ajusté en juillet 2013 (carte), sur- ou sous-évaluation sous-jacente (droite) et corrélation avec le PIB/habitant (bas/gauche).



La carte présente l'indice Big Mac ajusté, c'est-à-dire corrigé de l'influence du PIB/habitant mise en évidence par le nuage de points en bas à gauche. Le diagramme à droite indique si une devise est sur- ou sous-évaluée vis-à-vis du dollar US par rapport au taux de change théorique induit par la PPA en Big Mac. Source : The Economist.

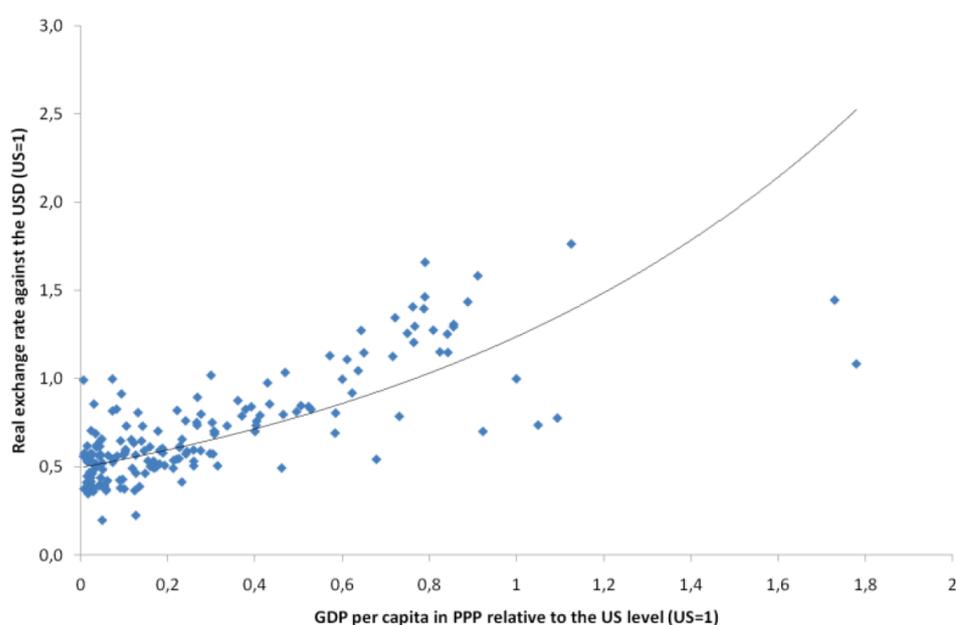
solution, à la suite du journal The Economist, est d'utiliser le prix d'un bien très standardisé utilisant principalement des inputs locaux : le Big Mac de McDonald's¹. Quoique l'indice Big Mac capture d'autres facteurs non liés au TCR comme les politiques de prix des différentes franchises McDonald's, il a le mérite d'être facile à construire, et disponible auprès du journal *The Economist* de façon semestrielle depuis 1986. La Figure 6.2 présente la corrélation de cet indice en juillet 2013 avec le revenu par habitant et sa valeur corrigée de l'influence du PIB/habitant (indice ajusté).

La Figure 6.2 appelle plusieurs observations. Premièrement, elle confirme que la PPA absolue n'est pas vérifiée. Le prix du Big Mac exprimé en dollars

1. Les données brutes de ces deux méthodes peuvent être téléchargées librement sur les sites [ICOP 2011](#) et [Big Mac Index](#) .

US varie fortement d'un pays à l'autre. Naturellement, on ne s'attend pas à ce que le Loi du prix unique s'applique à ce bien non-échangeable, mais on pouvait penser qu'elle s'appliquait à ses inputs. Deuxièmement, il y a une relation systématiquement croissante entre le prix d'un Big Mac et le revenu par habitant du pays, si forte que The Economist calcule un indice Big Mac "ajusté". La Figure 8.6, confirme l'existence d'une corrélation positive entre le PIB par habitant réel calculé en PPA aux Etats-Unis et le TCRE en 2009, calculé à partir d'indices des prix à la consommation.

FIGURE 6.3 – TCRE et PIB par habitant réel en PPA relativement aux Etats-Unis, 2009



Source : IMF, CEPIL.

Ces observations appellent une théorie reliant croissance, revenu par habitant et prix des biens non-échangeables.

6.1.2 TCR, prix des non-échangeables et revenu par habitant

L'existence de biens non-échangeables et de coûts du commerce limite l'application de la Loi du prix unique. Ainsi, une théorie qui expliquerait la corrélation positive entre prix des non-échangeables et PIB/habitant suffirait

à expliquer les déviations systématiques de la PPA illustrées par les figures précédentes.

Balassa (1964) et Samuelson (1964) proposent un argument pour expliquer pourquoi les prix des biens non-échangeables sont systématiquement plus faibles dans les pays pauvres. L'argument prend pour point de départ la forte corrélation positive entre productivité du travail et revenu par habitant. Selon la Loi du prix unique, les prix des biens échangeables doivent être égaux dans tous les pays. De ce fait, les producteurs des pays pauvres, qui emploient plus de travail par unité vendue, doivent payer un salaire plus faible. En raison de la mobilité des travailleurs, les salaires dans les secteurs non-échangeables doivent aussi être plus faibles dans les pays pauvres. Ceci implique des prix des non-échangeables plus faibles et donc des indices de prix plus faibles dans ces pays, d'où un échec de la PPA. Notons que la théorie de Balassa et Samuelson prédit aussi qu'un pays pauvre à forte croissance doit connaître une hausse des salaires, une hausse du prix des non-échangeables, et une appréciation réelle.

On peut formaliser cet argument de façon très simple pour faire apparaître les hypothèses sous-jacentes. Supposons 2 pays, 2 secteurs : échangeable (T) et non-échangeable (N) et un seul facteur, le travail. On note $y^T = a^T L^T$ et $y^N = a^N L^N$ les fonctions de production représentant les technologies des deux secteurs. On suppose que le secteur T est soumis à la loi du prix unique $p^T = \frac{p^{T*}}{S}$ (les variables avec une étoile représentent l'autre pays). Avec concurrence parfaite sur les marchés des biens et du travail et mobilité des travailleurs, les prix sont égaux aux coûts marginaux et les salaires sont égaux dans les deux secteurs. Ceci implique

$$\begin{aligned} p^N &= \frac{a^T}{a^N} p^T \\ p^{N*} &= \frac{a^{T*}}{a^{N*}} p^{T*} \end{aligned}$$

En appliquant la Loi du prix unique on obtient

$$p^T = \frac{p^{T*}}{S} \Rightarrow \frac{p^N}{p^{N*}} = \frac{\frac{a^T}{a^{T*}}}{\frac{a^N}{a^{N*}}}$$

On suppose également que les deux pays ont des préférences identiques représentées par une fonction d'utilité Cobb-Douglas. On en déduit le TCR

$$Q \equiv \frac{P}{P^*} \equiv \frac{(p^T)^\alpha (p^N)^{1-\alpha}}{(p^{T*}/s)^\alpha (p^{N*}/S)^{1-\alpha}} = \left(\frac{\frac{a^T}{a^{T*}}}{\frac{a^N}{a^{N*}}} \right)^{1-\alpha}$$

En notant $\hat{x} \equiv \frac{dx}{x}$ la variation en pourcentage de chaque variable x , on obtient :

$$\hat{Q} = (1 - \alpha) \left[\left(\widehat{a^T} - \widehat{a^{T*}} \right) - \left(\widehat{a^N} - \widehat{a^{N*}} \right) \right]$$

Si le pays étranger est plus riche, au sens où sa productivité du travail vérifie $\left(\widehat{a^T} - \widehat{a^{T*}} \right) > \left(\widehat{a^N} - \widehat{a^{N*}} \right)$, on obtient $Q < 1$ et $\hat{Q} > 0$. Comme il est généralement vrai que la productivité dans les services (une partie importante du secteur N) est moins hétérogène que dans l'industrie et l'agriculture, cette prédiction peut expliquer des différences systématiques de prix entre pays.

Une explication complémentaire à l'argument Balassa-Samuelson a été proposée par Bhagwati, Kravis et Lipsey. Dans un modèle à 2 facteurs, capital et travail, les pays abondants en capital ont une productivité du travail plus élevée. Si les différences de dotations sont suffisamment fortes pour que l'égalisation du prix des facteurs échoue (dotations en dehors du cône de diversification), les salaires seront plus élevés dans ces pays. Or, les non-échangeables étant en grande partie des biens intensifs en travail, il en résulte que leur prix sera plus élevé dans ces pays. Dans la mesure où l'abondance en capital est positivement corrélée avec le revenu par habitant dans les données, cet argument peut aussi expliquer les observations que nous avons faites.

6.1.3 TCR et compétitivité

On peut également voir le TCR comme le reflet de la compétitivité-prix d'un pays. Définissons la compétitivité-prix d'un bien exporté par i comme le prix relatif de ce bien par rapport aux biens concurrents dans le pays de destination j , $\frac{S_{ijt} P_{ijt}^T}{P_{jt}^T}$. De la même manière écrivons $\frac{S_{ijt} P_{ijt}^T}{P_{jt}^T}$ la compétitivité-prix des importations de i en provenance de j .

Pour formaliser le lien entre le TCR et ses mesures de compétitivité, considérons le cas simple de 2 pays (H et F) et 2 biens échangeables, chacun produit dans un seul pays (variétés H et F). Supposons que ces variétés soient imparfaitement substituables selon les fonctions d'utilité Cobb-Douglas

$$\begin{aligned} U^H(x^H, x^F) &= (x^H)^\alpha (x^F)^{1-\alpha} \\ U^F(x^H, x^F) &= (x^H)^{\alpha^*} (x^F)^{1-\alpha^*} \end{aligned}$$

avec $0 < \alpha < 1$, $0 < \alpha^* < 1$ et $\alpha + \alpha^* > 1$, cette dernière hypothèse capturant une préférence pour les biens nationaux dans chaque pays (qui peut être une forme réduite des coûts du commerce capturés dans les estimations d'équations de gravité, par exemple). Notons S le taux de change nominal.

En utilisant les indices de prix dits idéaux $P^H = (p^H)^\alpha (Sp^F)^{1-\alpha}$ et $P^F = (Sp^F)^{\alpha^*} (p^H)^{1-\alpha^*}$,² on trouve qu'à l'optimum du consommateur le TCR s'écrit

$$Q^H = \left(\frac{p_H}{Sp_F^*} \right)^{\alpha + \alpha^* - 1}$$

Le TCR est ainsi une fonction croissante des *termes de l'échange*, c'est-à-dire du ratio entre prix des exportations et prix des importations.

Cette vision du TCR n'est pas incompatible avec la vision précédente liée au prix relatif des non-échangeables. Pour le voir, considérons un modèle à 4 biens, incluant à la fois une variété nationale imparfaitement substituable d'échangeables et un bien non-échangeable dans chaque pays.

On suppose les fonctions d'utilité suivantes :

$$\begin{aligned} U^H(x^H, x^F, x^N) &= ((x^T)^\alpha (x^F)^{1-\alpha})^\gamma (x^N)^{1-\gamma} \\ U^F(x^H, x^F, x^{N*}) &= ((x^F)^{\alpha^*} (x^H)^{1-\alpha^*})^\gamma (x^{N*})^{1-\gamma} \end{aligned}$$

avec $0 < \alpha < 1$, $0 < \alpha^* < 1$, $0 < \gamma < 1$ et $0 < \alpha + \alpha^* > 1$.

Par le même type de calcul que précédemment, on trouve le TCR

$$Q^H = \frac{\left(\frac{p^N}{p^H} \right)^\gamma}{\left(\frac{p^{N*}}{Sp^F} \right)^{\gamma^*}} \left(\frac{p_H}{Sp_F} \right)^{\alpha + \alpha^* - 1}$$

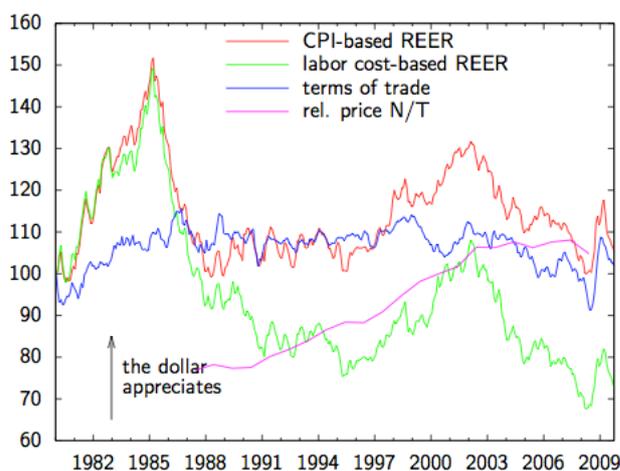
On trouve ainsi que le TCR varie positivement à la fois avec le prix relatif national des non-échangeables et avec les termes de l'échange.

La Figure 6.4 présente l'évolution du TCRE des Etats-Unis de 1980 à 2010, parallèlement à l'évolution du prix des non-échangeables (courbe rose) et des termes de l'échange (courbe bleue). Deux mesures du TCRE sont proposées, une fondée sur les indices des prix à la consommation, l'autre sur les coûts salariaux. La Figure suggère que le TCRE suit les mouvements des termes de l'échange ainsi que du prix relatif des non-échangeables, du moins en tendance.

On retient donc que les deux visions de la déviation à la PPA, liées à l'existence de non-échangeables et à l'imparfaite substituabilité des échangeables importés, contribuent à expliquer l'évolution des taux de change réels.

2. On peut vérifier que ces indices de prix sont égaux à la dépense minimale nécessaire pour obtenir une unité d'utilité avec les fonctions d'utilité Cobb-Douglas supposées plus haut.

FIGURE 6.4 – TCR des Etats-Unis, prix relatif des non-échangeables et termes de l'échange, 1982-2009



Les courbes représentent un TCRE calculé à l'aide d'indice des prix à la consommation, un TCRE calculé à l'aide de données sur le coût du travail, un ratio des termes de l'échange et un prix relatif des biens non-échangeables. Source : FMI, CEPII.

6.1.4 Discrimination géographique par les prix (Pricing to Market)

Pour être complet, on mentionnera brièvement une dernière cause importante d'échec de la PPA : la discrimination géographique par les prix (*pricing-to-market*)³.

Nous savons déjà que lorsque l'arbitrage d'un pays à l'autre est coûteux, la Loi du prix unique ne s'applique pas. Plusieurs raisons peuvent expliquer pourquoi les producteurs choisissent de vendre le même bien à différents prix en fonction du pays.

- en concurrence imparfaite, les producteurs ont intérêt à fixer des prix plus élevés dans les pays où l'élasticité-prix de la demande est plus faible
- il existe des rigidités de prix dans la monnaie du pays importateur : coûts de catalogues, contrats de long terme, indexation des prix...
- une partie des coûts de l'exportateur, notamment la distribution, sont payés dans la monnaie du pays importateur

3. Pour un traitement plus approfondi, on peut se référer à la revue de littérature de Burstein & Gopinath (2014).

- dans le cas du commerce intra-firme, le choix du prix de transfert interne peut varier d'un pays à l'autre pour des raisons légitimes (politique interne d'incitations) ou moins légitimes (optimisation fiscale)
- de nombreux exportateurs vendent des versions de qualité différente du même bien, avec une qualité systématiquement supérieure dans les pays riches

Cette discrimination par les prix a deux implications notables. Premièrement, un mouvement de taux de change ne sera pas entièrement transmis dans le prix en monnaie étrangère (*imperfect pass-through*). Ceci doit être pris en compte dans la modélisation des effets d'une dévaluation, par exemple. Deuxièmement, un même mouvement de taux de change bilatéral ne se traduira pas de la même manière dans tous les marchés.

6.2 Les déterminants du taux de change nominal

Dans cette section, on s'intéresse aux déterminants du taux de change nominal (parfois abrégé en taux de change, quand la confusion avec le TCR n'est pas possible).

Comme vu dans le chapitre introductif, les taux de change sont largement négociés de gré à gré sur des marchés mondiaux liquides, surtout sur certaines paires de devises importantes, à des faibles coûts de transaction. Ils sont donc soumis à l'attention des intervenants de marché, et sujets à de nombreuses opérations spéculatives au jour le jour. Ces observations guident notre analyse des taux de change nominaux de deux façons : premièrement, une théorie du change nominal doit intégrer que les instruments de change (taux de change spot, forward, swap) sont valorisés selon une logique de non-arbitrage ; deuxièmement, les anticipations des marchés financiers vont affecter ces taux.

6.2.1 Les parités de taux d'intérêt

Les taux de change comptant et à terme font l'objet de conditions de non-arbitrage simples appelées Parité de Taux de Taux d'Interêt Couverte ("Covered Interest-Rate Parity") et Parité de de Taux d'Interêt Non Couverte ("Uncovered Interest-Rate Parity"). Nous utiliserons les sigles anglais CIP et UIP ci-après.

CIP La CIP gouverne le prix d'un taux de change à terme. Dans un contrat à terme de taux de change, les parties s'entendent simplement sur le volume

et le prix d'une transaction à une date future donnée : par exemple, échanger 1 million d'euros contre des dollars US dans 6 mois à 1,3 dollars par euro. Le contrat "couvre" donc les parties contre le risque de change.

La CIP énonce que deux stratégies d'investissement apportent les mêmes gains : d'une part, prêter en monnaie nationale pendant t périodes ; d'autre part, prêter en monnaie étrangère pendant t périodes en convertissant les gains à un taux de change à terme convenu aujourd'hui. Si les gains de la deuxième stratégie étaient supérieurs, par exemple, on pourrait emprunter localement pour prêter à l'étranger et faire un profit positif sans risque (profit d'arbitrage). Sur des marchés liquides, ce genre d'arbitrage est impossible. CIP définit ainsi la valeur d'équilibre du taux de change à terme.

Formellement on peut écrire la CIP comme

$$(1+i)^t = (1+i^*)^t \frac{S_0}{F_t} \Leftrightarrow \left(\frac{1+i}{1+i^*} \right)^t = \frac{S_0}{F_t}$$

où i et i^* désignent les taux d'intérêt national et étranger, S_0 désigne le taux de change comptant aujourd'hui et F_t le taux de change à terme dans t périodes.

Cette formule est souvent exprimée en logs pour $t = 1$ (on utilise le fait que i est proche de zéro) :

$$i - i^* = s_0 - f$$

avec s_0 le log du taux comptant et f le log du taux à terme.

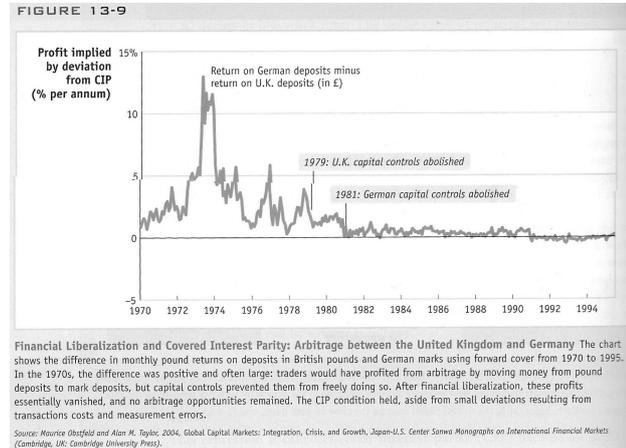
Quoique cette formule ne prenne pas en compte les coûts de transactions ou le risque de défaut, la CIP est relativement bien vérifiée dans les données, comme le suggère la Figure 6.5.

UIP Dans le même esprit, la parité UIP s'applique au taux de change comptant futur anticipé. La parité est dite non couverte puisque l'investisseur prend un risque de change : le taux de change futur n'est pas garanti par contrat et son anticipation peut se révéler fausse. Ainsi UIP spécifie que la stratégie dite de "carry trade" (emprunter en monnaie nationale puis prêter en monnaie étrangère en convertissant au taux de change futur comptant espéré) doit rapporter un profit nul. On remarque que UIP repose sur l'hypothèse de neutralité au risque.

On écrit cette condition

$$(1+i)^t = (1+i^*)^t \frac{S_0}{S_t^e} \Leftrightarrow \left(\frac{1+i}{1+i^*} \right)^t = \frac{S_0}{S_t^e}$$

FIGURE 6.5 – Test empirique de la parité de taux d'intérêt couverte (CIP) sur le cours GBP/DEM, 1970-1996



Le graphique représente les gains à emprunter en livres sterling (GBP) pour investir en deutsche marks (DEM) de 1970 à 1996. La libéralisation financière en Allemagne et au Royaume-Uni s'accompagne d'une chute de ces gains vers des niveaux négligeables.

ou en logs, pour $t = 1$,

$$i - i^* = s - s^e$$

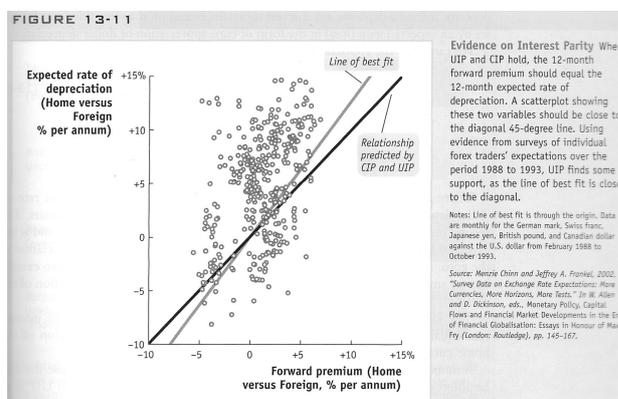
avec S_t^e et s^e le niveau et le log du taux de change comptant futur espéré.

UIP est moins validée empiriquement que CIP. Les données sur les anticipations de taux de change futur sont rares, mais une étude de Chinn & Frankel (2002) utilise une enquête sur les anticipations des intervenants sur le marché des changes sur le cours du dollars US contre 5 grandes monnaies entre 1988 et 1993. Ces auteurs testent plus précisément l'hypothèse que le taux à terme est égal au taux comptant espéré (impliquée à la fois par UIP et CIP), ou

$$\underbrace{\frac{F_t}{S_t} - 1}_{\text{'forward premium'}} = \underbrace{\frac{S_t^e}{S_t} - 1}_{\text{'expected rate of depreciation'}}$$

Les résultats sont décrits par la Figure 6.6. La Figure montre des déviations faibles entre les taux anticipés et à terme observés, d'une part, et la relation d'égalité théorique impliquée par CIP et UIP, d'autre part. Comme la dérivation de UIP et CIP néglige les coûts de transaction et différents risques, on peut considérer que ces observations soutiennent les hypothèses UIP et CIP.

FIGURE 6.6 – Test empirique des parités de taux d'intérêt UIP et CIP sur données d'enquêtes aux Etats-Unis de 1988 à 1993



Le graphique représente la déviation du taux de change anticipé impliqué par UIP au taux de change à terme impliqué par CIP. Sous l'hypothèse que UIP et CIP sont toutes deux valides, cette déviation est nulle et les deux droites sont confondues. Les données viennent d'enquêtes mesurant les anticipations mensuelles des intervenants de marché sur le cours du dollar US contre 5 grandes monnaies (CAD, DEM, CHF, GBP, JPY) de février 1988 à octobre 1993.

D'autres études testant UIP donnent des résultats moins favorables. Ces études utilisent des données historiques sur le taux de change et en calculant l'espérance de taux de change futur à l'aide d'un modèle économétrique de type

$$s_t - s_{t-1} = a + b(i_t - i_t^*) + u_t$$

Tandis que UIP implique $a = 0, b = 1$, on trouve généralement $a \neq 0, b < 0$ à court terme, et $b > 0$ seulement à long terme.

L'échec de ce test peut venir des hypothèses fortes de neutralité au risque et d'anticipations rationnelles. Ainsi le développement de "carry trades" dans les années 2000, notamment au Japon contre le dollar australien, va à l'encontre de UIP. Ces stratégies spéculatives consistent à emprunter dans la monnaie d'un pays ayant des taux d'intérêt faibles et d'investir à moyen ou long terme ("carry") dans des obligations d'un autre pays où les taux sont plus élevés. Par exemple, au début des années 2000 de nombreux investisseurs japonais, y compris individuels, ont commencé à emprunter en yen à des taux proches de zéro pour investir en Australie à des taux nettement plus élevés. En principe, si UIP est valide, le yen doit se déprécier vis-à-vis du dollar australien jusqu'à ce que cette stratégie n'apporte aucun profit. En pratique, cette stratégie s'est avérée rentable jusqu'en 2008, mais après

la crise financière de 2008 certains investisseurs ont subi une lourde perte. Cet épisode illustre l'importance du risque dans les choix de portefeuille, que nous avons ignoré pour simplifier notre dérivation de la relation UIP⁴.

Malgré ces limites, UIP reste une référence théorique utile lorsqu'on considère des changements n'affectant pas l'aversion au risque ou les anticipations.

6.2.2 Modèles monétaires du change nominal

Jusqu'ici, nous avons vu deux familles de théories du taux de change. La théorie de la PPA prédit la valeur du taux de change en fonction des prix. La théorie UIP prédit les taux de change en fonction des taux d'intérêt. On voudrait pouvoir intégrer ces prédictions théoriques dans un modèle macroéconomique dans lequel prix et taux d'intérêt sont endogènes.

On voudrait également pouvoir rendre compte de la volatilité considérable du change nominal. A titre d'illustration, la Figure 6.7 compare le taux de change nominal entre le deutsche mark et le dollar US observé et le ratio des indices des prix, donc la valeur du change prédite par la PPA, de 1970 à 1994, à la fois en niveau et en variation. La partie droite de la Figure montre bien la volatilité nettement plus élevée du change nominal par rapport à la valeur impliquée par la PPA⁵.

Nous allons donc présenter plusieurs modèles monétaires du taux de change nominal. Ces modèles endogénéisent les taux d'intérêt et les prix, et -pour certains- parviennent à expliquer la volatilité excessive du change nominal. Il existe dans la littérature plusieurs modèles candidats. Nous choisissons ici de présenter 3 modèles qui sont à la fois influents et suffisamment proches les uns des autres pour les présenter comme des variantes d'un modèle canonique.

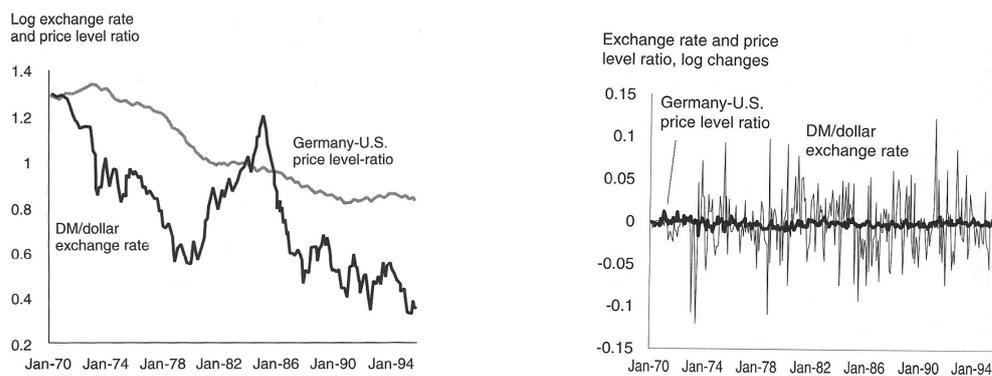
Modèle canonique

On considère le modèle suivant en temps continu d'une *petite économie ouverte* :

4. Pour plus de détails sur le "carry trade", voir Krugman et al. (2012, chapitre 14).

5. Cette volatilité a fait l'objet de nombreux travaux empiriques depuis les années 70 (grossièrement depuis l'abandon du système de Bretton Woods). Certains ont trouvé que les fluctuations à court terme du change nominal suivent une marche aléatoire, autrement dit sont imprévisibles. D'autres ont observé des transitions de régimes de change flexible vers des régimes de change fixe s'accompagnant d'une volatilité du change supérieure à la volatilité des autres fondamentaux économiques (production, emploi, inflation...). La plupart des explications proposées pour expliquer cet excès de volatilité implique des rigidités nominales (surtout des prix) et des stratégies de pricing-to-market.

FIGURE 6.7 – Taux de change nominal DEM/USD et ratio des indices des prix allemand et américain : en niveaux (logs, gauche) et en variation (log-différences, droite).



Données mensuelles. Source : Obstfeld & Rogoff (1996)

$$y_t = \delta(p_t^* - p_t - s_t) + g_t - \sigma i_t \quad (\text{IS})$$

$$m_t - p_t = \psi y_t - \phi i_t \text{ ou } i_t = i_t^* + \lambda(p_t - u_t) \quad (\text{LM ou MP})$$

$$\dot{p}_t = \theta(y_t - \bar{y}_t) \quad (\text{PC})$$

$$i_t = i_t^* - E\dot{s}_t \quad (\text{UIP})$$

avec $\delta > 0$, $\sigma > 0$, $\psi > 0$, $\phi > 0$, $\lambda > 0$, $\theta > 0$. Toutes les variables hors i et i^* sont en logs. y_t désigne (le log de) la production, p_t et p_t^* les niveaux des prix national et étranger, s_t le taux de change nominal au certain, g_t les dépenses publiques, i_t et i_t^* les taux d'intérêt nominaux national et étranger, m_t la quantité de monnaie en circulation, $u_t > 0$ un choc de politique monétaire. p_t^* et i_t^* sont traitées comme des constantes dans cette petite économie ouverte.

Le modèle décrit par ces équations rappelle fortement le modèle OA-DA (ou ASAD en anglais), un des modèles traditionnels pour expliquer les fluctuations macroéconomiques de court terme. On peut rappeler que ce modèle prédit comment la production et les prix fluctuent à court terme sous l'effet de facteurs d'offre et de demande⁶. Les deux premières équations représentent le bloc demande du modèle, qui implique la courbe AD. La première équation est une courbe IS augmentée : à un niveau de prix donné, la production est

6. Pour une présentation complète du modèle, on peut se référer au chapitre 5 du polycopié du cours ECO432.

déterminée par la demande pour les biens nationaux (consommation et investissement), qui croît sous l'effet d'une hausse des dépenses publiques, d'une baisse du taux d'intérêt, ou d'une baisse du TCR. La deuxième équation, dans sa première variante, est la courbe LM : l'offre de monnaie contrôlée par la banque centrale doit être égale à la demande de monnaie qui croît avec la production, pour motif de transaction, et décroît avec le taux d'intérêt, le coût d'opportunité de détenir de la monnaie. Une variante de cette deuxième équation, MP, décrit également le fonctionnement du marché monétaire, mais en donnant un rôle central au taux d'intérêt, que les banques centrales contrôlent mieux que l'offre de monnaie. Enfin, la troisième équation résume le bloc offre du modèle (courbe de Phillips, étroitement liée à la courbe AS). La production atteint son niveau de prix flexibles \bar{y} lorsque les producteurs anticipent correctement le niveau des prix, mais dépasse temporairement ce niveau quand des chocs d'offre ou des erreurs d'anticipation font que les producteurs sous-estiment le niveau des prix, ou tombe en deçà de ce niveau quand ils surestiment le niveau des prix. Finalement, l'équilibre est défini comme le couple $\{y_t, p_t\}$ à l'intersection des courbes AS et AD.

Nous allons décliner ce modèle ASAD augmenté de UIP en trois versions : prix flexibles (Mussa 1976), prix fixes (modèle Mundell-Fleming, déjà vu en ECO432) et prix rigides (Dornbusch 1976). Le tableau 6.2.2 synthétise les différences entre ces trois versions.

	Prix	Anticipations	Courbe de Phillips	Clôture du modèle
Mussa (1976)	Flexibles	$E\dot{s}_t = \dot{s}_t$	$\theta = +\infty$	y exogène, p endogène
Mundell-Fleming	Fixes	$E\dot{s} = 0$	$\theta = 0$	y endogène, p exogène
Dornbusch (1976)	Rigides	$E\dot{s}_t = \dot{s}_t$	$0 < \theta < +\infty$	y endogène, p rigide

Version à prix flexibles

Dans cette variante (Mussa 1976), les prix sont parfaitement flexibles et s'ajustent immédiatement à tout type de choc. Le niveau de production \bar{y}_t n'est donc pas affecté par les variables nominales, et on peut le concevoir comme le niveau de "long terme" qui serait prédit par un modèle de croissance (dans le modèle ASAD, cela correspond à une courbe AS de long terme verticale et une courbe de Phillips verticale). Les agents ont des anticipations rationnelles et le taux de change nominal, comme un prix d'actif financier, incorpore toutes les anticipations des agents sur les taux d'intérêt futurs.

La variante à prix flexibles s'écrit :

$$\bar{y}_t = y_t = \delta(p_t^* - p_t - s_t) + g_t - \sigma i_t \quad (\text{IS,PC})$$

$$i_t = i_t^* + \lambda(p_t - u_t) \quad (\text{MP})$$

$$\dot{i}_t = i_t^* - \dot{s}_t \quad (\text{UIP})$$

La première équation, qui résume l'équilibre sur le marché des biens, fixe le log du TCR, égal à $\frac{g_t - \bar{y}_t - \sigma i_t}{\delta}$.

On déduit des trois équations une équation différentielle linéaire du premier ordre qui détermine le taux de change nominal :

$$\dot{s}_t = \frac{\lambda\delta}{\delta + \lambda\sigma} (s_t - s_t^f)$$

avec $s_t^f = \frac{g_t - \bar{y}_t - \sigma i_t^*}{\delta} + p_t^* - u_t$ la valeur dite *fondamentale* du taux de change.

La solution à cette équation différentielle prend la forme générale

$$s_t = B e^{\mu t} + \mu \int_t^{+\infty} e^{-\mu(\tau-t)} s_\tau^f d\tau$$

avec $\mu \equiv \frac{\lambda\delta}{\delta + \lambda\sigma}$ et B une constante arbitraire.

Comme $\mu > 0$, on remarque qu'il existe un continuum de solutions explosives (pour $B \neq 0$) et une seule solution non-explosive dite fondamentale. Dans ce qui suit, on fixera arbitrairement $B = 0$ ⁷. Le taux de change s'écrit donc

$$s_t = \mu \int_t^{+\infty} e^{-\mu(\tau-t)} s_\tau^f d\tau$$

On remarque que le taux de change s'écrit comme une somme pondérée de toutes les valeurs fondamentales futures anticipées, *exactement comme un prix d'actif financier*. Supposons que les agents anticipaient un taux de change stationnaire, mais que soudain ils apprennent qu'un choc permanent (et un seul) va se produire et change la valeur fondamentale future. Au moment d'apprendre la nouvelle, le terme de droite change : le taux de change varie de façon discrète. En revanche, en l'absence de seconde nouvelle, le taux de change varie de façon continue jusqu'à ce que le choc se produise, puis reste constant à cette nouvelle valeur.

7. On notera cependant que le phénomène de "bulle" capturé par la solution explosive (le taux de change augmente parce qu'on anticipe qu'il augmente dans le futur) peut en principe expliquer une partie de la volatilité du change nominal. Dans un modèle plus complet on pourrait endogénéiser les croyances des agents sur le type d'équilibre, fondamental ou explosif, et inclure les deux types de dynamique.

Les valeurs d'état stationnaire de i_t , p_t , s_t s'écrivent :

$$\begin{aligned} i_t &= i_t^* \\ p_t &= u_t \\ s_t &= s_t^f = \frac{g_t - \bar{y}_t - \sigma i_t^*}{\delta} + p_t^* - u_t \\ q_t &= \frac{g_t - \bar{y}_t - \sigma i_t^*}{\delta} \end{aligned}$$

avec q_t le (logarithme du) TCR.

Le modèle nous fournit trois enseignements. Premièrement, le taux de change se comporte comme un prix d'actif financier : il incorpore toutes les anticipations sur les chocs futurs qui pourront affecter les fondamentaux, et il est affecté par ces chocs futurs dès le moment où les agents apprennent leur existence. Il peut également être sujet à des bulles, bien que nous n'ayons pas développé cette partie du modèle. Deuxièmement, les chocs qui peuvent affecter la valeur fondamentale du taux de change sont à la fois des chocs monétaires et réels. Ainsi, un choc réel négatif sur g (austérité budgétaire) ou positif sur \bar{y} (gains de productivité) cause une dépréciation réelle et nominale ; un choc positif sur u (politique monétaire nationale expansionniste) cause une dépréciation nominale, et un choc négatif cause une appréciation. Ainsi, en signalant une nouvelle politique monétaire plus restrictive, une Banque Centrale peut obtenir une appréciation du change, comme a pu le faire la Banque Centrale russe fin 2014. Enfin, une politique monétaire expansionniste à l'étranger (baisse de i^*) cause une appréciation nominale, une politique restrictive cause une dépréciation nominale. Ceci implique que l'absence de coordination des politiques monétaires entre pays rend le change plus volatile : lorsque les Banques Centrales adoptent la même politique monétaire, les effets sur le change nominal s'annulent. Ce modèle apporte donc un argument en faveur de la coordination des politiques monétaires.

Version à prix fixes : le modèle Mundell-Fleming

Le modèle Mundell-Fleming est une extension de ISLM en économie ouverte. Comme dans ISLM, on s'intéresse au très court terme dans lequel les prix sont fixes et ce sont les quantités produites qui s'ajustent à tout choc de demande, notamment ceux provenant de la politique monétaire et de la politique budgétaire. L'analyse exclut les fluctuations ayant pour origine le côté offre de l'économie. Par rapport à ISLM, qui prédit comment la quantité produite et le taux d'intérêt d'équilibre fluctuent avec des chocs de demande, le modèle de Mundell-Fleming prédit également l'évolution du change nominal.

Le modèle s'écrit :

$$y_t = \delta(p_t^* - p_t - s_t) + g_t - \sigma i_t \quad (\text{IS})$$

$$m_t - p_t = \psi y_t - \phi i_t \quad (\text{LM})$$

$$i_t = i_t^* \quad (\text{UIP})$$

On obtient facilement $s_t = p_t^* - p_t + \frac{1}{\delta} \left(g_t - \sigma i_t^* - \frac{m_t - p_t + \phi i_t^*}{\psi} \right)$. Dans l'espace $\{y, i\}$ IS est croissante et LM décroissante, comme dans le modèle ISLM en économie fermée, mais le taux d'intérêt est fixé à i^* , ce qui fixe s . Dans l'espace $\{y, s\}$, la courbe issue de LM est *verticale* et la courbe issue de IS est décroissante, ce qui fixe i .

Ce modèle très simple offre pourtant des prédictions différentes de ISLM en économie fermée. Une politique budgétaire de relance (hausse de g) n'a plus aucun effet sur la production : le prix et le taux d'intérêt étant fixes, la production est contrainte par la quantité de monnaie disponible. Même si une relance stimule la demande nationale, la monnaie s'apprécie pour que la demande étrangère diminue et devienne compatible avec la production. Donc une relance budgétaire perd son efficacité en économie ouverte. Une politique monétaire expansionniste (hausse de m) a au contraire et comme dans ISLM un effet positif sur la production, et comme dans la variante précédente cause une dépréciation.

On note que le modèle Mundell-Fleming est aussi utilisé pour étudier les politiques monétaire et budgétaire en régime de change fixe. s_t est remplacé par sa valeur en change fixe, tandis que m_t devient endogène, en raison des mouvements de réserves de change (voir chapitre 7, consacré aux régimes de change). Les conclusions sont opposées au cas de change flexible : la politique budgétaire redevient efficace, car l'appréciation consécutive à une relance n'a plus lieu. En revanche, la politique monétaire perd son effet sur la production. Au contraire du régime de change flexible, le régime à change fixe est plus stable en présence de chocs monétaires et plus instable en présence de chocs réels. On reviendra sur ces prédictions dans le prochain chapitre.

Version à prix rigides et excès de volatilité du change

Les deux variantes précédentes nous ont appris que le taux de change fluctue avec les changements de politique monétaire présents et anticipés. Les taux de change devraient donc varier en proportion de ces changements. Cependant, comme mentionné plus haut, les changements de politique monétaire ne suffisent pas à expliquer l'ampleur de la volatilité du change nominal.

Cette dernière variante (Dornbusch 1976) propose une explication de la volatilité du change nominal. Lorsque les prix sont rigides, des chocs qui

se traduiraient par des changements de prix dans la première variante se répercutent sur le taux de change. Cependant, au cours du temps, les prix s'ajustent et le taux de change reprend sa valeur de long terme. Le taux de change présente donc un *excès de volatilité* ("overshooting" en anglais) par rapport aux prédictions du modèle à prix flexibles : à court terme, le change répond de façon excessive aux chocs monétaires.

Par exemple, supposons un choc monétaire positif. A long terme, les prix doivent augmenter proportionnellement à ce choc pour atteindre le nouvel équilibre. Comme précédemment, le TCR d'équilibre, inchangé, exige une dépréciation de même proportion que le choc monétaire. Mais si les prix sont fixes à court terme, le choc de politique monétaire aura un effet sur le taux d'intérêt, qui baissera immédiatement, et par UIP ceci implique une dépréciation immédiate. On en déduit que le taux de change se dépréciera de façon transitoire au-delà du niveau exigé par le nouvel équilibre de long-terme, et restera inférieur à cette valeur le temps que les prix augmentent jusqu'à leur nouvelle valeur d'équilibre.

Ce modèle s'écrit

$$y_t = \delta(p_t^* - p_t - s_t) + g_t - \sigma i_t \quad (\text{IS})$$

$$i_t = i_t^* + \lambda(p_t - u_t) \quad (\text{MP})$$

$$\dot{p}_t = \theta(y_t - \bar{y}_t) \quad (\text{PC})$$

$$i_t = i_t^* - \dot{s}_t \quad (\text{UIP})$$

A l'état stationnaire du modèle, on obtient

$$i_t = i_t^*$$

$$y_t = \bar{y}_t$$

$$\bar{y}_t = \delta(p_t^* - p_t - s_t) + g_t - \sigma i_t^*$$

$$p_t = u_t$$

Les deux variables principales du modèle ont pour valeur à l'état stationnaire

$$\bar{p} = u_t$$

$$\bar{s} = s_t^f = \frac{g_t - \bar{y}_t - \sigma i_t^*}{\delta} + p_t^* - u_t$$

Les termes de droite ne varient dans le temps qu'en présence de chocs, et restent constants sinon, d'où l'absence d'indices de temps dans les termes de gauche. On remarque que le taux de change d'état stationnaire a la même valeur dans cette variante et dans la première variante à prix flexibles, sans

surprise puisque dans cette variante les rigidités nominales ont disparu à long terme.

On peut étudier la dynamique du modèle en substituant dans les équations (PC) et (UIP) :

$$\begin{aligned}\dot{p}_t &= \theta(y_t - \bar{y}_t) = \theta\delta(\bar{s} - s_t) + (\theta\delta + \lambda\theta\sigma)(\bar{p} - p_t) \\ \dot{s}_t &= \lambda(\bar{p} - p_t)\end{aligned}$$

ce qui peut se réécrire

$$\begin{pmatrix} \dot{s}_t \\ \dot{p}_t \end{pmatrix} = \begin{pmatrix} 0 & -\lambda \\ -\theta\delta & -(\theta\delta + \lambda\theta\sigma) \end{pmatrix} \begin{pmatrix} s_t - \bar{s} \\ p_t - \bar{p} \end{pmatrix}$$

ou sous forme matricielle

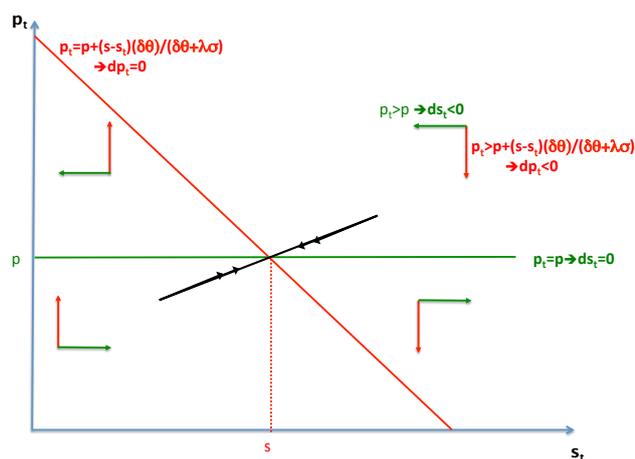
$$\dot{Z}_t = AZ_t$$

On peut vérifier que la matrice A a pour déterminant $-\lambda\theta\delta < 0$ et admet 2 valeurs propres de signe opposé, ce qui implique un seul sentier de convergence vers l'état stationnaire. La Figure 6.8 résume la dynamique du modèle dans un diagramme de phase. Le diagramme résume l'évolution de $\dot{Z}(t)$ au cours du temps. Les deux droites décrivent les valeurs de s_t et p_t compatibles avec l'une des deux variables à sa valeur d'état stationnaire. La solution d'état stationnaire se trouve à l'intersection de ces deux droites. Les flèches décrivent l'évolution de s_t et p_t à partir des points n'appartenant pas à ces deux droites. Enfin, on a tracé le sentier de convergence unique vers l'état stationnaire. Les autres trajectoires, divergentes, sont omises pour alléger le diagramme.

Comme dans les variantes précédentes, on cherche à connaître l'effet de chocs monétaires et de chocs réels sur le taux de change. Une politique monétaire expansionniste $du_t > 0$ a pour effet immédiat de réduire le taux d'intérêt national et de déprécier la monnaie. A plus long terme, les prix s'ajustent vers un niveau d'équilibre supérieur, le taux d'intérêt national se rapproche du taux étranger et la monnaie s'apprécie pendant cette transition jusqu'à atteindre un niveau d'équilibre plus faible. La Figure 6.9 résume la dynamique de ce modèle identique au précédent sauf pour la valeur de u_t . On remarque l'ajustement instantané du taux de change sur le nouveau sentier de convergence, en raison des anticipations rationnelles, tandis que les prix sont encore fixes. On remarque également que cette dépréciation est excessive par rapport au nouveau niveau d'état stationnaire du taux de change, et l'appréciation graduelle pendant la convergence vers l'état stationnaire.

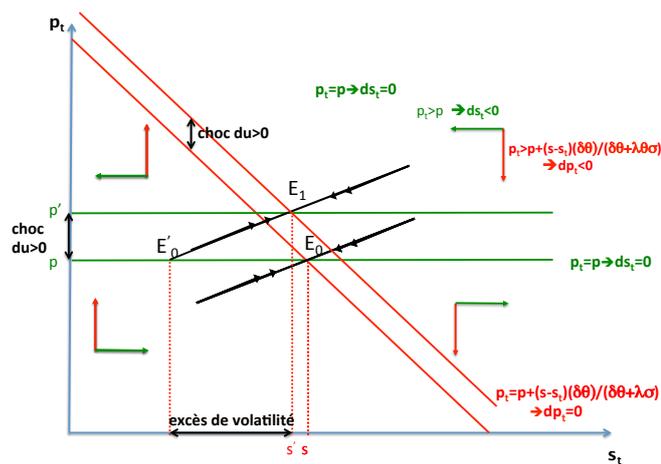
Pour résumer la variation des prix, du change et du taux d'intérêt au cours du temps suite au choc de politique monétaire, la Figure 6.10 présente un

FIGURE 6.8 – Dynamique du taux de change nominal et des prix dans le modèle de Dornbusch



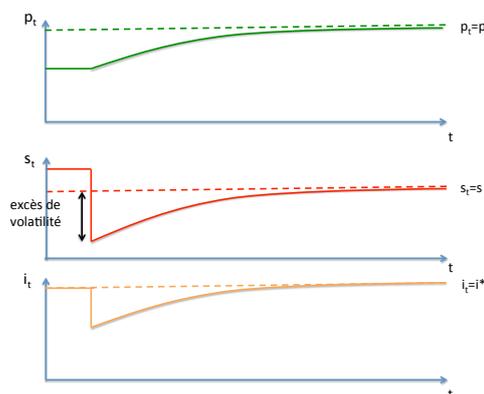
La notation est légèrement différente du texte : s et p désignent les valeurs d'état stationnaire de s_t et p_t respectivement.

FIGURE 6.9 – Dynamique du taux de change nominal et des prix après un choc monétaire dans le modèle de Dornbusch



La notation est légèrement différente du texte : s et p désignent les valeurs d'état stationnaire de s_t et p_t respectivement.

FIGURE 6.10 – Réponse impulsionnelle des prix, du taux de change et du taux d'intérêt après un choc monétaire dans le modèle de Dornbusch



La notation est légèrement différente du texte : s et p désignent les valeurs d'état stationnaire de s_t et p_t respectivement.

diagramme de réponse impulsionnelle. Ce diagramme illustre bien la lenteur de la réponse des prix au choc monétaire relativement à la réponse du taux de change nominal et du taux d'intérêt.

Finalement, une politique budgétaire de relance $dg_t > 0$ a pour effet immédiat une appréciation réelle, et comme les prix sont rigides, une appréciation nominale simultanée. Le changement de g_t affectant s_t et \bar{s} de la même façon, le taux de change nominal atteint immédiatement sa nouvelle valeur d'équilibre sans excès de volatilité. La production ne varie pas car l'appréciation compense exactement le choc de demande positif, comme dans le modèle Mundell-Fleming.

Discussion

Les 3 modèles monétaires du change nominal nous ont appris quelques leçons importantes. Premièrement, le taux de change nominal se comporte comme un prix d'actif financier. Il incorpore les anticipations sur des événements futurs (typiquement la politique monétaire), varie avec des changements d'anticipations et parfois avec des bulles. Ceci explique une partie de sa volatilité. Deuxièmement, l'impact de chocs réels et monétaires sur le change nominal dépend du type de modèle. Dans un modèle à prix flexibles, qu'on peut interpréter comme la situation de long terme, des chocs réels négatifs sur

g (austérité budgétaire) ou positifs sur \bar{y} (gains de productivité) causent une dépréciation ; une politique monétaire nationale expansionniste cause une dépréciation, tandis que la même politique à l'étranger cause une appréciation. Dans un modèle à prix fixes, qu'on peut interpréter comme le très court terme, une politique de relance cause une appréciation mais sans effet sur la production, tandis qu'une politique monétaire expansionniste cause une dépréciation nominale et réelle et stimule la production. Finalement, dans un modèle à prix rigides, une politique monétaire expansionniste cause une dépréciation excessive dans l'immédiat, ce qui stimule temporairement la production, puis une appréciation lente vers un taux de change plus faible et un retour vers la production de long terme. Une politique de relance budgétaire cause une appréciation nominale et réelle immédiates, compensant le choc de demande positif et rétablissant la production à son niveau de long terme, comme dans le modèle à prix fixes.

Par rapport aux autres variantes le modèle à prix rigides apporte deux intuitions importantes supplémentaires. Quand les prix sont momentanément fixes, les variations de change nominal causent les variations de change réel, du moins pendant la transition. Pendant cette transition les variations de change nominal peuvent être excessives, puisque le change nominal opère l'ajustement que d'autres variables plus rigides ne peuvent pas opérer.

Ces modèles doivent être compris comme une introduction à l'étude du change nominal. S'ils contiennent des messages importants, ils présentent aussi de nombreuses limites. Ils négligent l'approche intertemporelle vue dans les chapitres précédents et ne traitent pas les effets des variations de change sur les déséquilibres courants (les TCR sont fixes). Les hypothèses sur les anticipations sont très fortes. Les actifs financiers nationaux et étrangers sont traités comme des substituts parfaits (seul le taux d'intérêt compte), ce qui ne permet pas de traiter l'impact des variations de change sur les flux d'actifs financiers. Les modèles ne formalisent pas la rigidité des prix, et notamment le choix de la monnaie dans laquelle les prix sont libellés. Finalement, la performance empirique du modèle de Dornbusch est mitigée. On observe notamment que le renversement de la variation de change s'opère bien plus tard que prédit par le modèle. Cependant, cette observation pourrait être réconciliée avec le modèle dans une extension où les agents découvrent lentement la persistance du choc de politique monétaire (Gourinchas & Tornell 2004).

6.3 Conclusions

Dans ce chapitre, nous avons étudié les déterminants du taux de change réel (TCR) et du taux de change nominal. Nous avons vu que malgré leurs

noms similaires ces deux variables économiques jouent des rôles assez différents. Le TCR, qui mesure le prix relatif des biens nationaux vis-à-vis des biens étrangers, garantit l'équilibre sur les marchés des biens nationaux et étrangers. Le taux de change nominal est un prix d'actif financier. Il est gouverné par des relations de non-arbitrage et incorpore toutes les anticipations des agents, en particulier celles sur la politique monétaire future. Il est parfois sujet à des phénomènes de bulles.

Nous avons vu que le change nominal est bien plus volatile que sa valeur théorique de PPA. Autrement dit, les théories du change nominal fondées sur le change réel ne sont pas suffisantes pour la prévision. En revanche, les modèles monétaires du taux de change nominal révèlent plusieurs mécanismes pouvant expliquer une partie de cette volatilité : comme prix d'actif, le taux de change nominal varie au gré des changements d'anticipation voire des bulles ; comme prix instantanément flexible, le taux de change nominal peut fluctuer pour se substituer à la fluctuation du prix des biens, contraints par plusieurs sources de rigidité.

Finalement, on peut se demander comment ces deux objets expliqués par des théories différentes peuvent être cohérents l'un avec l'autre. La réponse dépend du modèle utilisé. Dans le modèle à prix flexibles, les prix s'ajustent pour atteindre le TCR garantissant l'équilibre sur le marché des biens ; étant donné ce TCR, la politique monétaire affecte le taux de change nominal. Le taux de change nominal dépend donc de facteurs monétaires, mais aussi de facteurs réels à travers le TCR. Dans le modèle à prix rigides, le TCR dévie du niveau qui équilibrerait les marchés des biens. La politique monétaire détermine le change nominal comme précédemment, mais c'est le change nominal qui détermine le change réel jusqu'à ce que les prix s'ajustent. Comme les prix sont lents à s'ajuster, le taux de change est excessivement volatile pendant la transition.

Chapitre 7

Régimes de change fixe et flexible

Jusqu'ici, nous avons considéré que les taux de change étaient déterminés par des échanges libres sur les marchés des changes internationaux. Ceci est largement vrai pour les monnaies de grandes économies développées (USD/EUR, USD/JPY, EUR/GBP...). Cependant, il existe un grand nombre d'autres arrangements dans lesquels le taux de change ne varie pas librement. Ces arrangements méritent d'être étudiés non seulement parce qu'ils couvrent une partie significative de l'économie mondiale (du taux de change chinois administré par la Banque Centrale à l'union monétaire de la zone euro), mais aussi parce que les crises de régime de change fixe sont instructives pour les pays en change flexible.

Dans ce chapitre nous présenterons les différents arrangements apparentés aux régimes de change fixe. Puis nous décrirons comment l'ajustement du solde courant diffère du cas du change flexible. Enfin nous étudierons différents types de crises de change, c'est-à-dire de situations dans lesquelles le cours du change dévie brutalement du cours fixe désiré. Notre discussion illustrera une proposition due à Mundell, et popularisée en Europe par Padoa-Schioppa, appelée "triangle d'impossibilité" : on ne peut avoir simultanément indépendance de la politique monétaire, liberté des mouvements de capitaux et taux de change fixe.

7.1 Typologie des régimes de change

Commençons par décrire brièvement les différents types de régimes de change existants.

Jusqu'ici nous avons décrit un régime de change flexible, dans lequel les devises sont librement échangées sur les marchés de change mondiaux. La Banque Centrale (qui a le monopole d'émission monétaire dans la devise du

pays) peut choisir de garder des réserves de devises étrangères, notamment pour assurer la liquidité des marchés en cas de défaillance passagère de certains participants. Mais elle n'a pas de cible explicite de taux de change et n'intervient pas sur les marchés de change pour rétablir un cours cible. Cependant, comme nous l'avons vu, la Banque Centrale affecte le change nominal quand elle détermine ses instruments de taux d'intérêt et indirectement la masse monétaire.

A l'autre extrême, il existe des devises qui ne sont simplement pas convertibles. Seuls les résidents sont autorisés à utiliser la monnaie nationale, et ils ne peuvent en principe pas l'échanger contre d'autres devises¹. En pratique, il peut bien sûr exister un marché noir, qui peut être alimenté en partie par des activités illicites. Un marché noir fonctionnera assez similaire à un système de change flexible, mais dans un contexte bien moins concurrentiel (on s'attend à des marges d'intermédiation importantes autour des valeurs prédites par les parités).

Entre ces deux extrêmes se trouvent plusieurs variantes de systèmes de change fixe. La variante la plus contraignante est l'union monétaire, dans laquelle il n'existe qu'une seule monnaie dans les pays participants. C'est le cas de la zone euro, mais aussi de la zone franc CFA, de la zone rand autour de l'Afrique du Sud... Les statuts de l'union monétaire définissent les modalités d'émission monétaire. Dans le cas de la zone euro, la BCE gouverne l'émission de monnaie totale dans la zone euro, bien que cette émission soit décentralisée dans l'Eurosystème. Historiquement, certaines unions monétaires *de facto* se sont produites, lorsqu'un pays a spontanément adopté la monnaie d'un autre pays. Dans ce cas, l'union est asymétrique, au sens où un pays abandonne totalement à l'autre pays le pouvoir d'émission monétaire. Ceci a été le cas notamment de l'Equateur, qui a adopté le dollar US, ou du Kosovo, qui a adopté l'euro.

D'autres régimes de change fixe moins contraignants que l'union monétaire consistent simplement à ce que la Banque Centrale intervienne sur les marchés des changes pour maintenir un cours fixé à l'avance. A partir du milieu du XIXe siècle a émergé le système dit d'étalon-or, dans lequel chaque

1. Il existe un grand nombre de devises de pays émergents officiellement non-convertibles mais pour lesquelles des échanges sont possibles à l'aide d'instruments à terme libellés en dollars ('non-deliverable forwards'). Pour certaines devises, comme par exemple le renminbi chinois, la roupie indienne, ou encore le real brésilien, la liquidité de ces marchés à terme rend la devise *de facto* convertible. Dans d'autres pays émergents, souvent plus petits comme par exemple le Bangladesh, ces marchés n'existent pas et la devise est effectivement non-convertible. On peut également citer l'exemple de Cuba, où jusqu'à récemment coexistaient une devise convertible indexée au dollar US et une devise non convertible.

monnaie avait un cours indexé sur le cours de l'or. Les Banques Centrales utilisaient leurs réserves d'or pour maintenir ce cours, en achetant leur propre monnaie contre de l'or pour éviter une dépréciation, et en vendant leur monnaie pour éviter une appréciation. Après la Seconde Guerre Mondiale et les accords de Bretton Woods, ce système a été remplacé dans les pays occidentaux par un système plus flexible où les cours des monnaies étaient indexés de façon rigide sur le cours du dollar, la Fed gardant la possibilité de modifier le cours du dollar en or. Enfin, depuis l'abandon de ce système au début des années 70, les Banques Centrales des pays en change fixe administrent des réserves de devises étrangères qu'elles utilisent pour maintenir les cours, parfois en conjonction avec des restrictions sur l'utilisation de devises par les ménages et les entreprises. Enfin, on peut mentionner l'existence de "currency boards", autorités indépendantes gérant des réserves en devises étrangères égales à la totalité de la masse monétaire nationale, comme dans le cas de l'Argentine dans les années 90 (on discutera dans les sections suivantes de l'utilité de ce dispositif).

Il existe des variantes du système de change fixe moins contraignantes. Certains pays choisissent une trajectoire de cours prévisible, plutôt qu'un cours fixe (Mexique au début des années 90). D'autres choisissent un cours-cible avec des bandes de fluctuations (Système Monétaire Européen avant l'euro, Danemark aujourd'hui). D'autres enfin choisissent d'administrer le cours par des interventions discrétionnaires sur le marché des changes, comme la Chine depuis 2005.

Malgré leur diversité, les arrangements de change fixe ont plusieurs points communs. Ils reposent sur la capacité des Banques Centrales à influencer les cours grâce à leurs réserves, et sur la confiance des agents dans cette capacité. Ils éliminent ou réduisent la possibilité d'ajustement du solde courant par le change nominal, ce qui fait l'objet de la section suivante.

7.2 Ajustement en change fixe et flexible

Comme nous l'avons vu dans les chapitres précédents, l'ajustement du solde courant dépend du taux de change réel, qui lui-même dépend du taux de change nominal. A partir du moment où les prix ne sont pas parfaitement flexibles, l'ajustement en change fixe ne se fait plus de la même façon et des déséquilibres sont possibles.

7.2.1 Ajustement en change flexible

Dans le modèle à deux biens et deux périodes du chapitre 4, un ajustement du solde courant cohérent avec la contrainte de budget intertemporelle nécessitait un changement du TCR : par exemple un pays en déficit courant avait besoin d'une dépréciation réelle (baisse du prix relatif des non-échangeables) permettant un excédent courant futur. Dans ce modèle, la balance commerciale était déterminée par la demande nationale d'échangeables, elle-même issue de l'utilité Cobb-Douglas, et par l'offre exogène. Nous cherchons ici à étudier sous quelles conditions ce lien entre TCR et solde courant reste valide dans des modèles plus généraux.

Intuitivement, comme le TCR est un prix relatif, une dépréciation réelle accroît la quantité de biens nationaux demandée et réduit la quantité de biens étrangers demandée. Mais cette dépréciation réelle causant une baisse du prix des exportations relativement au prix des importations, l'effet sur les exportations nettes (exprimées en valeur) est a priori ambigu. Il dépend des élasticités-prix des importations et des exportations.

Pour comprendre comment le TCR et le solde courant varient, on peut étudier un exemple simple sans biens non-échangeables et avec une demande isoélastique pour les biens nationaux et étrangers. De plus, on néglige les flux de revenus primaires de et vers l'étranger, de sorte que le solde courant est égal à la balance commerciale.

Soit

$$\begin{aligned} X &= \left(\frac{SP_X}{P^*}\right)^{-\sigma_X}, & P_X &= P \\ M &= \left(\frac{P_M}{P}\right)^{-\sigma_M}, & P_M &= \frac{P^*}{S} \end{aligned}$$

Alors $\frac{dX}{X} = -\sigma_X \frac{dQ}{Q}$ et $\frac{dM}{M} = \sigma_M \frac{dQ}{Q}$ avec $Q = \frac{SP}{P^*}$.

La balance commerciale, exprimée en devise étrangère, s'écrit :

$$B = SP_X X - P_M M$$

On différencie totalement cette expression au voisinage d'une balance commerciale équilibrée ($B = 0$) :

$$\frac{dB}{p_X X} = -(\sigma_X + \sigma_M - 1) \left(\frac{dP}{P} - \frac{dP^*}{P^*} + \frac{dS}{S} \right)$$

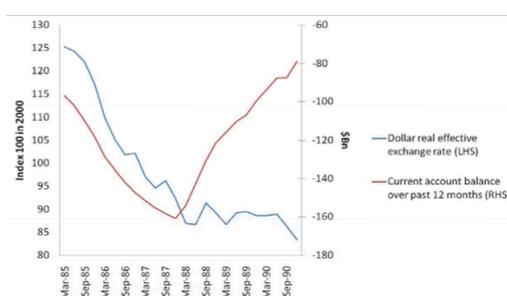
Par conséquent, une dépréciation réelle réduit le solde courant si et seulement si la condition de Marshall-Lerner

$$\sigma_X + \sigma_M > 1 \quad (\text{Marshall-Lerner})$$

est vérifiée.

La condition de Marshall-Lerner spécifie que la demande pour les biens nationaux et étrangers doit être suffisamment élastique pour l'effet de la dépréciation sur les quantités domine l'effet mécanique sur les prix. Les données suggèrent que cet effet domine à moyen terme, mais on observe souvent que l'effet mécanique sur les prix domine à court terme : ainsi, après une dépréciation réelle, le solde courant suit une *courbe en J*, comme l'illustre la Figure 7.1 dans le cas des Etats-Unis de 1985 à 1990².

FIGURE 7.1 – Courbe en J : TCRE et solde courant des Etats-Unis, 1985-1990



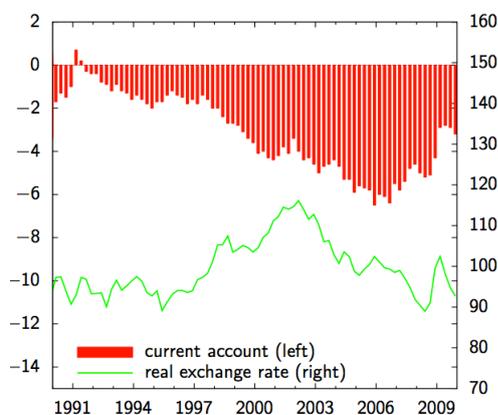
Données mensuelles. TCRE : indice 100 en 2000. Solde courant des 12 derniers mois en milliards de dollars US. Source : BRI.

Sur longue période on observe souvent un comportement du TCR et du solde courant plutôt cohérent avec la condition de Marshall-Lerner. Les Figures 7.2 et 7.3 montrent l'évolution du solde courant et du TCR dans le cas des Etats-Unis et du Royaume-Uni. On remarque immédiatement que la corrélation entre le TCR et du solde courant est ambiguë. Mais dans le cas des Etats-Unis, on observe une forte détérioration du solde courant au moment d'une appréciation réelle (fin des années 90) et un fort ajustement du solde courant au moment d'une forte dépréciation réelle en 2007-2008. Dans le cas du Royaume-Uni, on constate également une détérioration de la balance commerciale pendant la période d'appréciation réelle (fin des années 90 et début des années 2000) et un ajustement après la forte dépréciation réelle de 2008-2009.

La Figure 7.4 représente les mêmes variables dans les cas plus spectaculaires de crises de change (abandon du change fixe) dans plusieurs pays émergents. Le point commun entre ces 4 pays est d'avoir connu des crises de change causant des dévaluations spectaculaires : en 1995 pour le Mexique, en

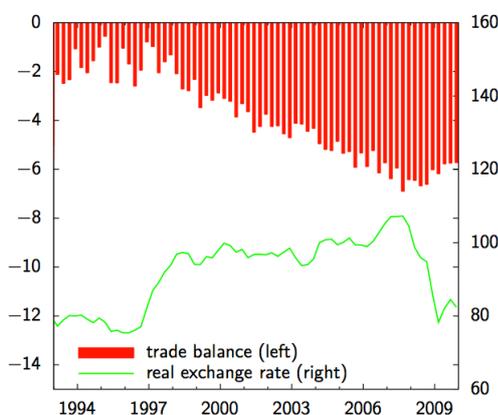
2. Ces observations concordent avec les estimations d'élasticités-prix à court et long terme dans 14 pays développés présentées par Krugman et al. (2012, chapitre 17, tableau 17A2-1, p462).

FIGURE 7.2 – TCRE et solde courant des Etats-Unis, 1991-2010



Solde courant exprimé en % du PIB nominal. Source : FERI, FMI.

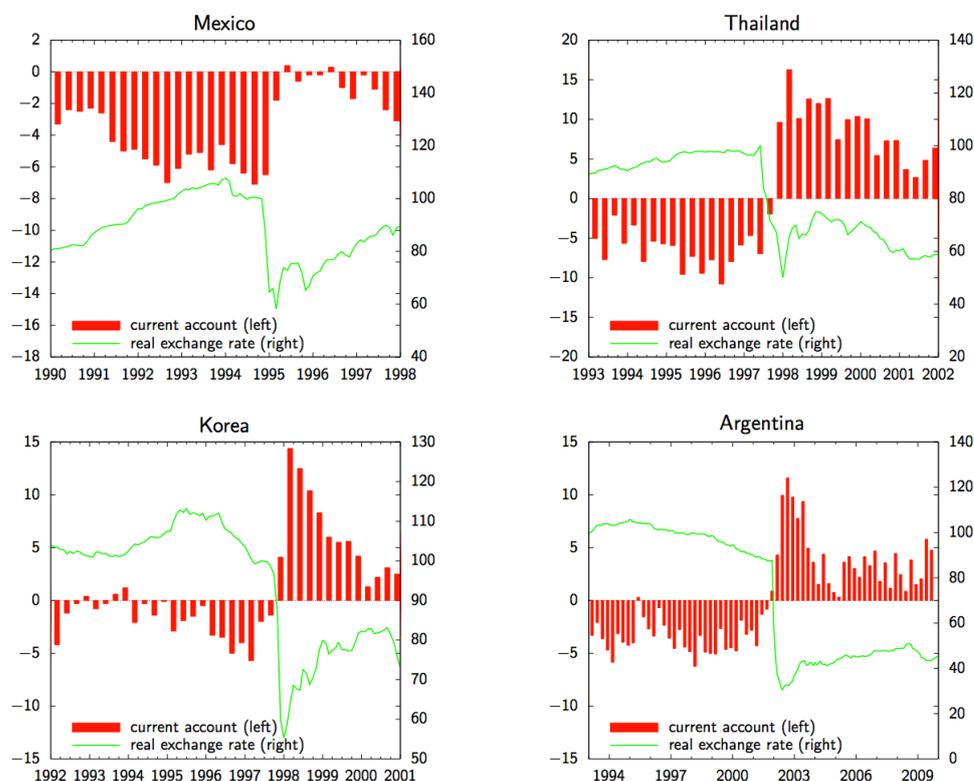
FIGURE 7.3 – TCRE et balance commerciale du Royaume-Uni, 1991-2010



Balance commerciale exprimée en % du PIB nominal. Source : FERI, FMI.

1997 pour la Corée et la Thaïlande, en 2001-2002 pour l'Argentine. Dans les 4 cas, on observe une détérioration du solde courant avant la crise, puis une nette amélioration après la crise. Ces graphiques suggèrent que l'abandon du régime de change fixe a rendu possible un ajustement qui n'était possible qu'au prix de mouvements de réserve non soutenables sur le long terme.

FIGURE 7.4 – TCR et soldes courants dans des pays émergents touchés par des crises de change



TCR bilatéral par rapport aux Etats-Unis. Solde courant exprimé en % du PIB nominal. Source : FERI, FMI.

7.2.2 Ajustement en change fixe

L'ajustement du solde courant en change fixe pose plus de difficultés. Comme on peut facilement le comprendre, tout dépend du degré de flexibilité des prix, dont dépend directement la flexibilité du TCR.

L'analyse de l'ajustement à long terme (prix flexibles) est héritée de l'analyse du système d'étalon-or par le philosophe écossais David Hume au milieu du XVIIIe siècle. A l'époque de Hume, la livre sterling était convertible en or à un taux officiel. Une hausse des réserves d'or permettait de créer plus de monnaie dans le pays. Un solde courant positif impliquait un mouvement d'or des pays créditeurs vers le Royaume-Uni. Contre l'école de pensée mercantiliste, qui préconise de toujours maintenir un excédent courant pour accumuler de l'or (au besoin en subventionnant les exportations), Hume décrit

un mécanisme d'ajustement interdisant des excédents permanents.

Imaginons le scénario suivant ³ : supposons que le stock d'or du Royaume-Uni soit multiplié par 5, et donc (par approximation) que la masse monétaire soit multipliée par 5. Selon Hume, cette forte expansion monétaire a pour effet une forte inflation du prix des biens et des salaires. Sous l'effet de cette inflation les exportations britanniques deviennent hors de prix tandis que les importations du reste du monde deviennent tellement bon marché que même un gouvernement protectionniste ne parviendrait pas à les bloquer. Le déficit courant causé par l'expansion monétaire initiale aurait ainsi pour effet un paiement en or du Royaume-Uni vers le reste du monde, jusqu'à retourner à la situation initiale.

On voit que Hume fait implicitement référence au concept moderne de TCR. Une façon plus moderne de présenter l'argument de Hume est la suivante : après une appréciation réelle, la balance commerciale doit se détériorer. Ceci crée un excès de demande pour la monnaie étrangère, et le cours de la monnaie doit être maintenu par la vente de réserves de change (plutôt que de réserves d'or). En réduisant ses réserves de change, une Banque Centrale réduit la monnaie en circulation qui en est la contrepartie, ce qui comme une réduction des réserves d'or diminue la masse monétaire. A long terme il s'ensuit une déflation, qui restaure le TCR à sa valeur initiale, et le déficit courant est éliminé.

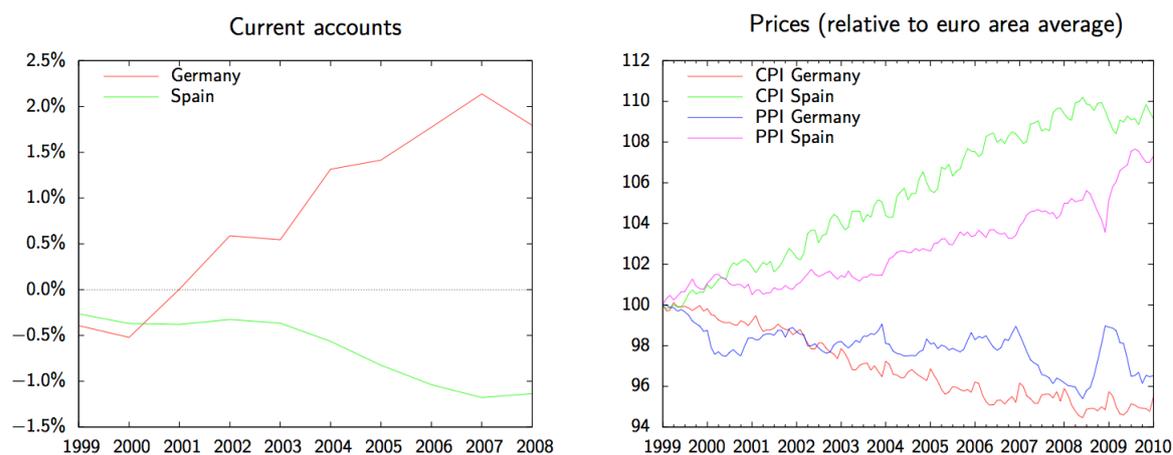
Le mécanisme d'ajustement décrit par Hume permet d'expliquer un certain nombre d'épisodes de déflation pendant le système de l'étalon-or (fin du XIXe et début du XXe siècles). Il fait également écho à la situation de pays en change fixe au XXe siècle, comme nous le verrons plus bas, et aux débats de politique économique dans la zone euro au XXIe siècle. Dans une union monétaire, l'émission monétaire se fait par une autorité unique, de sorte que les transferts de réserves ou d'or n'ont pas de sens. Cependant, l'idée qu'une déflation (et typiquement une baisse de salaire associée) permet d'ajuster le TCR et résorber les déficits courants, plus connue sous le nom de "dévaluation interne" dans les débats sur la zone euro, fait écho au mécanisme de Hume. On remarque enfin que ce mécanisme justifie l'hypothèse de PPA relative.

La Figure 7.5 illustre le lien entre solde courant et TCR dans la zone euro, en comparant l'Allemagne et l'Espagne pendant les années 2000. On constate une divergence des prix allemands et espagnols, tant pour les biens de consommation que pour les biens intermédiaires, et donc une baisse du TCR de l'Allemagne et une hausse du TCR de l'Espagne vis-à-vis du reste de la zone euro. Dans le même temps, on constate des déséquilibres croissants,

3. On paraphrase ici Hume (1752, essai "On the Balance of Trade").

avec d'une part des excédents courants allemands croissants et des déficits courants espagnols croissants, tous deux d'ampleur considérable. Sur les dernières années de l'échantillon, on constate une convergence des prix et une atténuation des déséquilibres. Avec une forte baisse des salaires, partiellement due à la hausse du chômage consécutive à la crise de 2008-2009 et à un effort de modération salariale, l'Espagne a amorcé une "dévaluation interne". En l'absence de transferts budgétaires intra-européens, de mobilité substantielle des travailleurs ou de fiscalité distorsive en faveur des biens espagnols, cette baisse des prix espagnole semble être la seule issue pour ajuster le solde courant. Côté allemand, l'introduction d'un salaire minimum inter-branches national début 2015 devrait faciliter une appréciation réelle permettant de résorber une partie des excédents courants.

FIGURE 7.5 – Soldes courants (gauche) et TCR (droite) vis-à-vis de la zone euro en Allemagne et en Espagne, 1999-2008



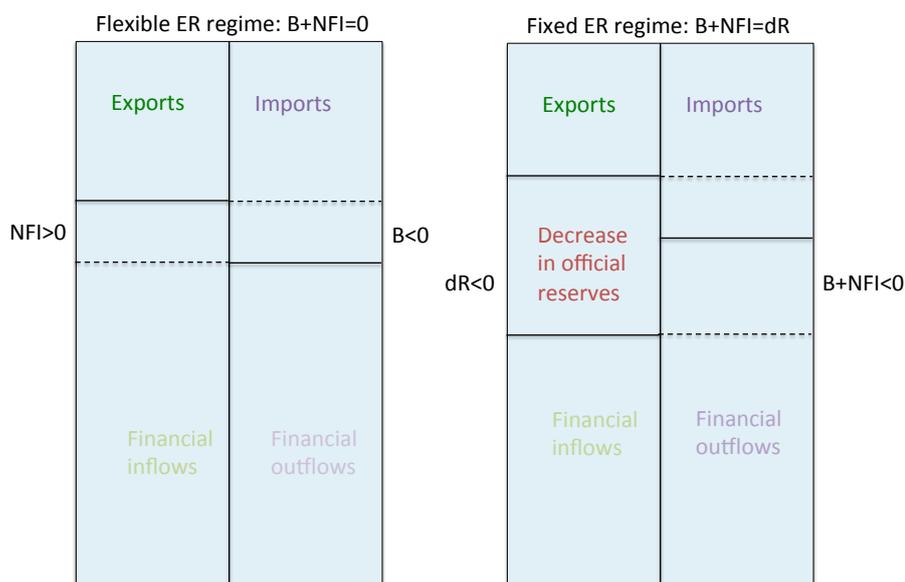
TCR vis-à-vis des prix moyens de la zone euro. Soldes courants en % du PIB nominal de la zone euro. Source : FERI, FMI.

7.2.3 Une comparaison de l'ajustement en change fixe et flexible

La Figure 7.6 décrit les changements de Balance des Paiements en régimes de change fixe et flexible. Dans les cas de change flexible, le solde courant a pour contrepartie exacte les flux financiers nets sortants. Dans le cas de change fixe, le solde courant a pour contrepartie exacte le solde financier,

c'est-à-dire les flux financiers nets sortants plus les variations de réserves de change.

FIGURE 7.6 – Ajustement en change fixes et flexible.



NFI désigne les flux financiers nets entrants. B désigne le solde courant. R désigne les réserves de change.

Nous pouvons conclure notre discussion en la résumant à l'aide d'un petit modèle à prix flexibles très simple, avec la notation habituelle où les minuscules désignent des variables en log :

$$p = m$$

$$m = r$$

$$b = \beta(p^* - p - s), \beta > 0$$

avec r le log des réserves de change.

Sous un régime de change fixe le taux de change nominal ne varie pas et les soldes courants ont pour contrepartie des mouvements de réserve $dr =$

$b, s = 0$

$$dp = dm = dr = b = \beta(p^* - p)$$

La PPA est restaurée lorsque les prix se sont ajustés : $p = p^*$.

Sous un régime de change flexible le taux de change nominal varie librement et $dr = 0$ donc

$$dp = dm = dr = 0 \Rightarrow s = p^* - p$$

L'ajustement se fait instantanément.

7.3 Crises de change

Finalement, nous étudions dans cette section les crises de régime de change fixe, ou crises de change. Par crise de change on entend une déviation brutale du taux de change de son cours officiel, la plupart du temps une dévaluation, mettant de facto fin au régime de change fixe. Loin d'être une curiosité théorique, ce type d'événement s'est produit relativement fréquemment. Dans les 30 dernières années, on peut citer les exemples notables du Système Monétaire Européen en 1992 (notamment la Finlande, l'Espagne, la Suède, l'Italie, le Portugal et le Royaume-Uni) ; le Mexique en 1994 ; la Thaïlande, l'Indonésie, la Corée du Sud et d'autres pays asiatiques en 1997, suivis par la Russie et le Brésil en 1998 ; l'Argentine en 2001 ; l'Islande, la Hongrie, et le Pakistan en 2008.

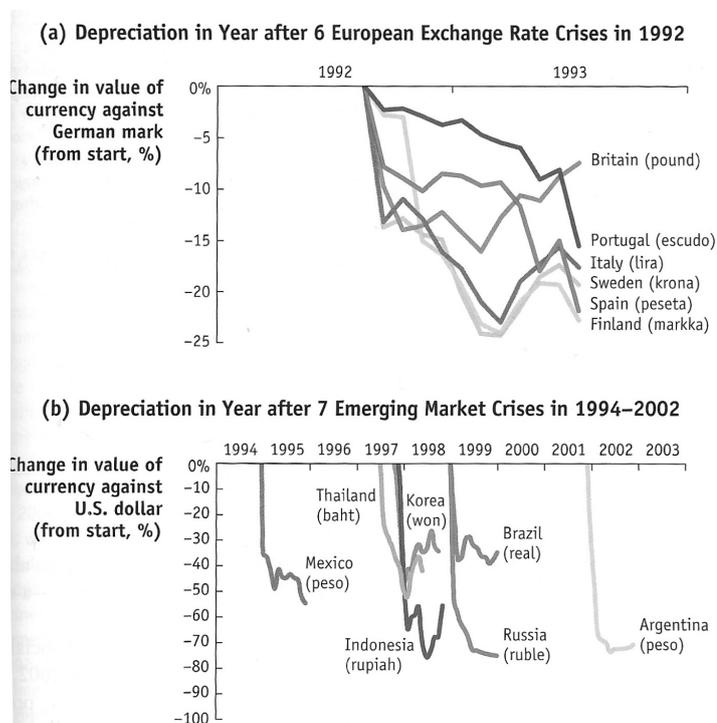
La Figure 7.7 illustre l'amplitude des dévaluations lors des crises de change en Europe en 1992 et dans les pays émergents en 1997-1998.

Nous allons étudier successivement 3 catégories de crises de change, parfois appelées "générations" de crises : crises de Balance des Paiements, liées à des attaques spéculatives et une pénurie de réserves de change ; crises autoréalisatrices, consistant en l'abandon volontaire du système de change fixe lorsqu'il perd sa crédibilité auprès des agents ; crises jumelles (crise bancaire et crise de change, ou crise de dette souveraine et crise de change).

Nous allons voir que ces 3 générations de crises ont pour point commun qu'elles illustrent le "triangle d'impossibilité" de Mundell,) savoir la proposition qu'on ne peut avoir simultanément indépendance de la politique monétaire, liberté des mouvements de capitaux et taux de change fixe, mais seulement deux de ces trois institutions. Pour comprendre l'intuition derrière cette proposition, rappelons la définition du TCR :

$$TCR_{ijt} = \frac{S_{ijt}P_{it}}{P_{jt}} \quad (\text{TCR})$$

FIGURE 7.7 – Evolution du change lors de crises de change en Europe (1992) et en Asie du Sud-Est (1997-1998).



Source : Feenstra & Taylor (2011)

Dans une petite économie ouverte i , qui traite P_{jt} comme donnée, cette équation fixe trois variables les unes par rapport aux autres : le taux de change nominal S_{ijt} , le niveau des prix P_{it} et le taux de change réel TCR_{ijt} . Intuitivement, un régime de change fixe détermine S_{ijt} ; une politique monétaire indépendante détermine P_{it} (toutes les Banques Centrales ont au moins un objectif d'inflation) ; enfin, l'ouverture financière permet au solde courant d'être non nul, ce qui selon la logique du chapitre 4 fixe TCR_{ijt} .

7.3.1 Crises de change de première génération

Nous commençons par étudier les crises de change de première génération, dites crises de Balance des Paiements, suivant l'article fondateur de Krugman (1979a). De façon informelle, ces crises se produisent lorsqu'une expansion monétaire nationale crée de l'inflation, causant une pression à la dépréciation, et que la Banque Centrale en défendant le cours officiel se rapproche

de l'épuisement de ses réserves de change. Si les agents ont des anticipations rationnelles, ils anticiperont l'effondrement du régime *avant* l'épuisement des réserves, et spéculeront contre la Banque Centrale en vendant massivement leurs avoirs en monnaie nationale contre des actifs étrangers, ce qui provoquera la crise anticipée. La fuite des capitaux provoquée par cette attaque spéculative donne le nom de crise de Balance des Paiements. On remarque que la logique de cette crise rappelle fortement les ruées sur les banques ("bank runs"), c'est-à-dire des retraits massifs et simultanés de dépôts bancaires causant la faillite de banques n'ayant en réserve qu'une fraction des dépôts.

Pour bien comprendre une crise de première génération, il est utile de présenter schématiquement le bilan d'une Banque Centrale. Comme pour toute banque, la comptabilité de la Banque Centrale (BC) suit les principes de la comptabilité en partie double. L'actif du bilan présente l'emploi des fonds de la BC, tandis que le passif présente la source de ces fonds.

Commençons par présenter le passif. Concrètement, chaque banque possède un compte à la BC et ces opérations se traduisent par des écritures sur ces comptes. La plupart des BC exigent des banques qu'elles y ouvrent un compte et déposent des réserves obligatoires sur ces comptes. Ces dépôts représentent un premier élément de passif. La monnaie fiduciaire en circulation (billets et pièces) constitue un deuxième élément de passif. En effet, historiquement, un billet servait de titre de créance à un déposant désireux de ne pas garder sur soi de l'or. Bien que les billets servent aujourd'hui principalement de moyen de paiement entre agents économiques et ne soient plus convertibles, ils restent une créance sur la Banque Centrale. Finalement, les BC ont des actionnaires (souvent, d'autres banques) et leurs capitaux propres figurent au passif.

Du côté de l'actif, on sait que les Banques Centrales créent de la monnaie en prêtant aux banques (opérations de refinancement) ou en leur achetant et vendant des actifs financiers (opérations d'open market). Les créances qu'elles ont sur ces banques et les titres qu'elles possèdent représentent donc un élément d'actif. On sait également qu'une grande partie de la masse monétaire est créée par les banques, de sorte qu'en contrôlant les conditions de leur refinancement la BC influence la masse monétaire totale. Enfin, les réserves de change représentent un élément d'actif. Quand la BC vend des devises nationales contre des devises étrangères, elle crédite le compte du vendeur si c'est une banque ou émet des billets, ce qui accroît son passif, en échange d'une quantité de devises étrangères à l'actif.

Le Tableau 7.1 résume cette présentation schématique du bilan d'une Banque Centrale.

On retient de cette présentation que la croissance du crédit dans une

TABLE 7.1 – Bilan simplifié d'une Banque Centrale

Actif		Passif	
Réserves de change	200	Capitaux propres	100
Titres	400	Monnaie en circulation	850
Prêts aux banques	400	Dépôts des banques	50

Note : les chiffres indiqués sont fictifs et non représentatifs du bilan d'une Banque Centrale.

économie nécessite une création de monnaie, comme on le savait, ou une réduction des réserves de change. De la même façon, une vente de réserves de change réduit la masse monétaire en circulation. Dans l'exemple du Tableau 7.1, si la BC vend 100 euros de réserves contre de la monnaie nationale, le poste Réserves de change deviendra égal à 100, tandis que la Monnaie en Circulation sera égale à 750 (les 100 euros achetés n'étant plus en circulation). Pour cette raison, les interventions des BC sur le marchés des changes sont souvent *stérilisées*, c'est-à-dire qu'elles sont accompagnées d'une autre opération visant à rendre l'ensemble neutre vis-à-vis de la masse monétaire. Pour continuer l'exemple précédent, une opération d'open market nationale, augmentant les titres à l'actif de 100 et créditant les comptes des banques au passif de 100, suffirait à stériliser l'intervention sur le marché des changes.

Ce type d'intervention stérilisée est au coeur du modèle de crise de première génération présenté ci-dessous, et partiellement inspirée de Krugman (1979a). Comme précédemment, toutes les variables sont en minuscules et notées en minuscules, sauf les taux d'intérêt nominaux i et i^* . On se place dans une petite économie ouverte traitant i^* et p^* comme exogènes. On suppose pour simplifier que y est exogène. Le modèle s'écrit :

$$\begin{aligned}
 m_t - p_t &= \phi y - \beta i_t, \phi > 0, \beta > 0 \\
 p_t &= p^* - s_t \\
 i_t &= i^* - \dot{s}_t \\
 m_t &= \theta r_t + (1 - \theta)d_t \\
 d_t &= d_0 + \mu t
 \end{aligned}$$

Les trois premières équations décrivent la courbe LM (équilibre sur le marché monétaire), la relation de PPA et la relation UIP vues au chapitre 6. La quatrième équation décrit le bilan de la BC : la croissance des réserves de change R_t et du crédit D_t a pour contrepartie la croissance de la masse monétaire. Enfin, la dernière équation décrit l'évolution (supposée exogène) du crédit à un taux de croissance fixe μ . Cette hypothèse forte a pour but de simplifier

l'analyse. Elle capture une croissance de l'endettement au moins temporairement non soutenable, que ce soit l'endettement public, par manque de discipline budgétaire (comme dans le modèle de Krugman), ou l'endettement bancaire, suite à une bulle financière ou immobilière, etc.

Considérons l'équilibre de cette économie en régime de change fixe $s_t = \bar{s}$. UIP et PPA impliquent la constance de i et p , donc LM implique la constance de m . Ceci implique que toute intervention de la BC sur le marché des changes devra être stérilisée (le fait que cette contrainte découle de PPA peut capturer l'obligation de maintenir un solde courant équilibré sous la pression de créanciers étrangers).

On obtient :

$$\begin{aligned} m &= m_0 = (1 - \theta)d_0 + \theta r_0 \\ \bar{s} &= -m + \phi y - \beta i^* + p^* \\ r_t &= \frac{1}{\theta} (m_0 - (1 - \theta)d_t) = r_0 - \frac{1 - \theta}{\theta} \mu t \end{aligned}$$

où r_0 désigne le niveau initial exogène de réserves.

On voit que les réserves de change sont tendanciellement décroissantes. Par construction du modèle, la croissance régulière du crédit oblige la BC à continuellement vendre ses réserves de change. On peut prévoir que ces réserves atteindront le niveau minimal r_{min} après T périodes,⁴ où

$$T = \frac{\theta}{1 - \theta} \frac{1}{\mu} (r_0 - r_{min})$$

Des agents à anticipations rationnelles savent qu'à la période T le régime de change fixe devrait théoriquement s'effondrer, c'est-à-dire que la BC ne pourra plus acheter sa monnaie au cours officiel. Cependant, comme dans un scénario de ruée sur la banque, chaque agent souhaiterait échanger ses avoirs en monnaie nationale *avant* l'effondrement du régime, pour ne pas subir la dévaluation consécutive à l'effondrement. De tels agents peuvent ainsi déduire du modèle la valeur théorique de la monnaie en changes flexibles, et vendre avant que cette valeur ait atteint le cours officiel. C'est à ce moment que tous les agents souhaiteront vendre leurs avoirs : comme les réserves ne seront pas encore épuisées, on parle d'attaque *spéculative*. Cependant, avec anticipations rationnelles ce type d'attaque ne peut que réussir.

4. Une valeur intuitive de r_{min} est zéro. Mais il est également possible qu'un niveau minimal de réserves soit indispensable pour le bon fonctionnement du commerce, du tourisme... Enfin, les BC prévoient de s'assurer les uns les autres contre le type d'attaque de ce modèle, de sorte qu'un r_{min} négatif (mais fini) est également plausible.

Pour calculer à quel moment l'attaque se produira, il est utile de résoudre le modèle en change flexible pour trouver la valeur *théorique* non officielle du change. Comme dans les modèles monétaires du chapitre 6, on obtient une équation différentielle en s qui admet une solution fondamentale (on ignore les solutions de bulle). Ici la solution fondamentale est égale à

$$s_t^f(r) = \bar{s} - \theta(r - r_0) - (1 - \theta)\mu t - \beta(1 - \theta)\mu$$

en utilisant la valeur \bar{s} solution du modèle en change fixe. Si la BC décidait arbitrairement d'abandonner le régime de change fixe à une date t , les agents anticiperaient un nouveau cours $s_t^f(r)$ immédiatement après l'annonce du changement de régime.

Sans surprise, on remarque que $s_t^f(r)$ décroît avec t . En particulier, $s_t^f(r_{min})$ atteint le niveau du cours officiel en change fixe à la période T' donnée par

$$s_{T'}^f(r) = \bar{s} \Leftrightarrow T' = \frac{\theta}{1 - \theta} \frac{r_0 - r_{min}}{\mu} - \beta$$

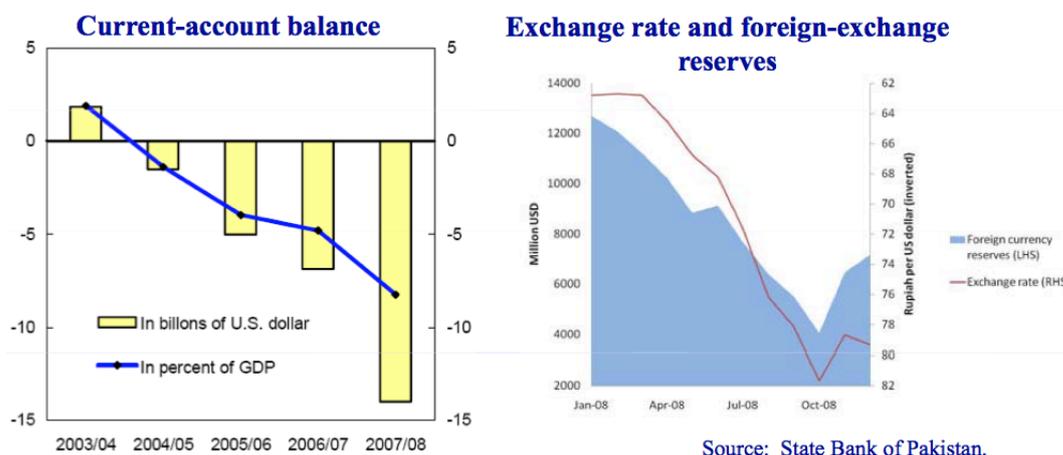
$\beta > 0$ implique que $T' < T$. On peut aussi vérifier que $r_{T'} = r_{min} + \frac{1-\theta}{\theta}\beta\mu$ est supérieur à r_{min} . Autrement dit, avant même que les réserves soient épuisées, les agents savent qu'après la période T' le cours théorique sera inférieur au cours officiel. Tout abandon du régime de change fixe après T' causerait une perte à tous les détenteurs de monnaie nationale. Ils vendront donc leurs avoirs à T' de façon simultanée⁵. A T' la Banque Centrale ne pourra maintenir le cours officiel face à cette vente massive de monnaie nationale, et sera contrainte d'abandonner le régime de change fixe.

La Figure 7.8 décrit une crise de change de première génération avec une attaque spéculative contre la roupie pakistanaise en 2008. On note l'accumulation de déficits courants jusqu'à des niveaux très élevés (près de 10% du PIB pakistanais). On observe également l'abandon du change fixe alors que les réserves de change chutent rapidement, mais avant leur épuisement complet.

En conclusion, cette approche suggère que pour prévenir une crise de change, les Banques Centrales devraient observer en permanence non seulement l'évolution de leurs réserves, mais plus généralement toutes les variables prédisant l'évolution du crédit futur.

5. Pour être plus précis, il existe un équilibre symétrique du jeu entre spéculateurs dans lequel tous les spéculateurs vendent à T' . Il peut en principe exister des équilibres asymétriques, mais cette discussion dépasse l'objet de notre analyse. Le message important du modèle est de rationaliser des attaques spéculatives non coordonnées par une multitude d'agents.

FIGURE 7.8 – Crise de change pakistanaise : solde courant de 2003 à 2008 (gauche), taux de change et réserves de change en 2008 (droite)



7.3.2 Crises de change de deuxième génération

On peut critiquer cette première approche par le rôle mécanique qu'elle donne à la Banque Centrale. Dans le modèle vu précédemment, la Banque Centrale vend continuellement des réserves en sachant pertinemment que la crise approche. Même dans un cadre théorique aussi simple, la BC pourrait augmenter le taux d'intérêt pour augmenter la valeur théorique du change. Certes, cette politique réduirait la masse monétaire et pourrait causer une déflation, mais l'abandon d'un régime de change fixe implique également tout un ensemble de coûts juridiques voire politiques. On souhaiterait avoir un modèle qui permette de comparer ces coûts et prenne en compte les choix de la BC. Par ailleurs, l'abandon du SME en 1992 suite aux craintes d'attaque spéculative contre le franc suggère que les BC parfois refusent délibérément de continuer à défendre un cours. Cet exemple est particulièrement intéressant car la France en 1992 ne montrait pas de signe de forte croissance du crédit ou de tensions inflationnistes (moins même que l'Allemagne qui venait de se réunifier).

Pour toutes ces raisons, on appelle crise de change de deuxième génération une crise autoréalisatrice, c'est-à-dire une crise provoquée par un changement d'anticipation des agents : lorsque les agents anticipent que la BC n'a plus intérêt à défendre le cours officiel, il devient coûteux pour la BC de défendre le

cours et elle choisit en effet d'abandonner le régime de change fixe. Ce type de phénomène peut être rationalisé dans un modèle à *équilibres multiples*, c'est-à-dire un modèle dans lequel plusieurs ensembles distincts de valeurs d'équilibre et d'anticipations sont solution du modèle.

Un exemple numérique simple permet de mieux comprendre pourquoi une hausse du taux d'intérêt ne suffirait pas à éviter une crise de change, comme suggéré plus haut. Rappelons la formule de UIP sur n périodes

$$i_t = i_t^* - \frac{1}{n}(s_{t,t+n}^e - s_t)$$

Si on définit une période comme un mois, et que les agents anticipent une baisse de la valeur théorique du change de 10% par mois, alors le taux d'intérêt doit augmenter de $\frac{0.1}{1/12} = 120\%$. Une telle augmentation pourrait avoir des effets désastreux sur la production et l'emploi. Une BC typique ne défendra donc pas un cours officiel à n'importe quel coût, et un changement d'anticipation peut la conduire à abandonner le régime de change fixe. Dans le cas du SME, le début de récession et le fort taux de chômage français ont pu jouer un rôle important sur ces anticipations⁶.

Nous allons illustrer cette idée dans un modèle dit de "clause de sortie" inspiré de Obstfeld (1991)⁷. Dans ce modèle, des rigidités nominales créent un dilemme inflation-chômage. Le régime de change optimal dépend des anticipations des agents : une dévaluation a un effet stimulant sur la production (et l'emploi) si elle est imprévue, mais pas si elle est anticipée. Le maintien du régime de change fixe limite l'inflation s'il est prévu, mais déprime la production et réduit les prix si une dévaluation est anticipée. Le coût social à maintenir le régime de change fixe est donc plus élevé quand les agents anticipent une dévaluation. Quand ce coût est suffisamment fort pour rendre le maintien du change fixe sous-optimal, une crise de change autoréalisatrice se produit.

Soit le modèle suivant d'une petite économie ouverte :

$$\begin{aligned} \dot{p}_t &= -\dot{s}_t \\ \dot{s}_t &= \arg \min_{\dot{s}_t} \{ (y_t - y^*)^2 + \theta(-\dot{s}_t)^2 + c(\dot{s}_t) \} \\ y_t &= y' + \eta(\dot{s}_t^e - \dot{s}_t) - \chi_t, \eta > 0 \end{aligned}$$

6. Pour plus de détails, voir Jeanne (1996) et la PC associée à ce chapitre.

7. Le nom se réfère à certains accords de change fixe prévoyant des clauses de sortie en cas de situation exceptionnelle : le modèle cherche à prédire dans quel cas un pays choisirait de sortir de l'accord.

avec $\theta > 0$ et

$$c(\dot{s}_t) = \begin{cases} 0 & \text{if } \dot{s}_t = 0 \\ c^d & \text{if } \dot{s}_t < 0 \\ c^a & \text{if } \dot{s}_t > 0 \end{cases}$$

La première équation représente l'hypothèse de PPA. La deuxième équation décrit le choix du planificateur social (Banque Centrale ou gouvernement) de conserver le régime de change fixe ($\dot{s}_t = 0$) ou de l'abandonner ($\dot{s}_t \neq 0$). Le planificateur social concilie trois objectifs : la stabilité de la production vis-à-vis de son niveau de long terme y^* ; la stabilité des prix, l'inflation étant égale à $-\dot{s}_t$; et la stabilité politique, $c(\dot{s}_t)$ représentant un coût politique égal à zéro quand le régime est conservé, et positif ($c^a > 0, c^d > 0$) quand il est abandonné⁸. Finalement, la troisième équation décrit la courbe de Phillips accélératrice d'un modèle ASAD : lorsque l'inflation est supérieure à l'inflation anticipée, la production augmente temporairement au-delà de son niveau de long terme (par exemple parce que la baisse des salaires réels stimule l'emploi et la production, ce qui fonde la relation croissante entre prix et production de la courbe AS). χ_t représente un choc d'offre aléatoire, réduisant la production à un niveau de prix donné. $\chi_t > 0$ peut capturer un choc sur le prix de l'énergie, sur le coût du travail, une catastrophe naturelle, voire un choc sur la confiance des investisseurs étrangers causant un retrait des capitaux.

L'hypothèse de coût politique $c(\dot{s}_t)$ a plusieurs justifications. Si le planificateur social agit dans l'intérêt de l'agent représentatif, différents groupes de pression peuvent influencer le taux de change : les exportateurs penchent pour une dévaluation, les consommateurs et les importateurs préfèrent une appréciation, le status quo politique a donc une valeur. En pratique, un épisode de crise de change peut endommager une carrière politique. Selon Feenstra & Taylor (2011), dans 58% des crises de changes le gouverneur de la Banque Centrale et/ou le ministre des Finances quittent leur poste dans l'année, contre une probabilité de 36% hors crise. Certaines crises de change spectaculaires ont contribué à la démission d'un chef d'Etat (de la Rúa en Argentine en 2001), voire au renversement d'un dictateur (Suharto en Indonésie en 1997). Enfin, le coût politique peut aussi capturer le coût d'un changement institutionnel rendu nécessaire par l'abandon du change fixe, ou des compensations aux pays partenaires prévues dans la clause de sortie d'un accord de change fixe.

8. La forme de la fonction objectif est familière en macroéconomie monétaire : elle correspond à une approximation de second ordre de la fonction d'utilité de l'agent représentatif, sous certaines hypothèses sur ses préférences (aversion au risque constante notamment). Voir le polycopié du cours ECO432.

Le taux de change optimal qui minimiserait les deux premiers termes de la fonction objectif est donné par

$$\begin{aligned} \min_{\dot{s}_t} \{ & (y' + \eta(\dot{s}_t^e - \dot{s}_t) - \chi_t - y^*)^2 + \theta(-\dot{s}_t)^2 \} \\ \Rightarrow \dot{s}_t = & \frac{\eta}{\eta^2 + \theta} (y' - \chi_t + \eta\dot{s}_t^e - y^*) \end{aligned}$$

A ce taux de change correspond la production

$$y_t = \frac{\theta}{\theta + \eta^2} y' + \frac{\eta}{\theta + \eta^2} (\theta\dot{s}_t^e + \eta(\chi_t + y^*))$$

et donc la fonction objectif prend pour valeur

$$L^{float} = \frac{\theta}{\theta + \eta^2} (y' + \eta\dot{s}_t^e - \chi_t - y^*)^2 + c^a \dot{\mathbf{s}}_t > \mathbf{0} + c^d \dot{\mathbf{s}}_t < \mathbf{0}$$

Sous le régime de taux de change fixe

$$y_t = y' + \eta\dot{s}_t^e - \chi_t$$

et la fonction objectif prend pour valeur

$$L^{fix} = (y' + \eta\dot{s}_t^e - \chi_t - y^*)^2$$

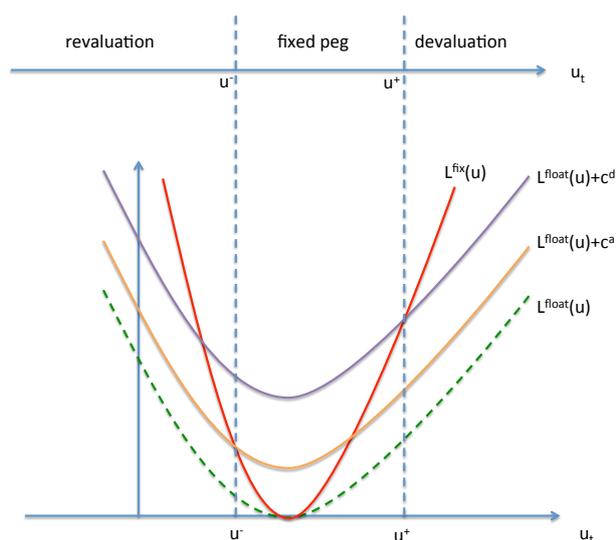
Si les coûts c^a et c^d tendaient vers zéro, le planificateur abandonnerait toujours le régime de change fixe : le système de change flexible lui permet de concilier stabilisation de la production et de l'inflation. Avec un coût politique non négligeable, la comparaison entre L^{float} et L^{fix} dépend de la valeur du choc χ_t . Il existe en général deux valeurs du choc χ^+ et χ^- solutions de $L^{fix}(\chi^-) = L^{float}(\chi^-) + c^a$ and $L^{fix}(\chi^+) = L^{float}(\chi^+) + c^d$:

$$\begin{aligned} \chi^+ &= y' + \eta\dot{s}_t^e - y^* + \frac{1}{\eta} \sqrt{(\theta + \eta^2)c^d} \\ \chi^- &= y' + \eta\dot{s}_t^e - y^* - \frac{1}{\eta} \sqrt{(\theta + \eta^2)c^a} \end{aligned}$$

Le choix du planificateur contingent à la réalisation du choc est résumé par la Figure 7.9.

On constate que les deux seuils χ^+ et χ^- sont croissants en \dot{s}_t^e . Supposons que le régime de change fixe soit initialement crédible : $\chi^- < \chi < \chi^+$, $\dot{s}_t^e = 0$. Toute nouvelle qui conduirait les agents à anticiper une dévaluation déplacerait χ^+ vers la gauche et pourrait rendre une dévaluation optimale, *sans aucun changement des fondamentaux économiques*. Dans cette configuration,

FIGURE 7.9 – Chocs d’offre et régime de change optimal quand $c^a < c^d$ dans le modèle de crise de change de deuxième génération.



La notation est légèrement différente du texte. Le choc d’offre est ici noté u au lieu de χ .

tant le maintien que l’abandon du régime de change fixe sont des situations d’équilibre, et une crise de change autoréalisatrice est possible.

Que peut faire la politique économique face à la possibilité de crises autoréalisatrices ? S’il est très difficile d’agir pendant une telle crise, il existe des moyens de prévention. Ces moyens consistent à augmenter les coûts politiques d’une sortie du régime de change fixe : on voit dans le modèle qu’une hausse de c^a et c^d élargit l’intervalle de paramètres pour lequel le régime de change fixe est crédible. Une Banque Centrale peut par exemple refuser d’abandonner son cours officiel pour se créer une réputation de forte aversion à l’inflation. La perte de cette réputation constituerait ainsi un coût politique pour la Banque Centrale. Une autre option suivant la même idée est la création d’un "currency board". Un currency board est une agence ayant l’obligation de rendre la monnaie intégralement convertible à une monnaie de référence, c’est-à-dire de conserver en réserves de change l’équivalent de 100% de la masse monétaire. Ce mécanisme institutionnellement coûteux est une forme d’engagement politique fort envers un change fixe. On remarque que ces deux options politiques limitent considérablement la capacité de la BC

du pays à créer de la monnaie, évoquant à nouveau le triangle d'impossibilité de Mundell, et qu'elles peuvent avoir un coût élevé en terme de chômage au moins initialement.

7.3.3 Crises de change de troisième génération

Les crises de change dites de troisième génération sont des crises jumelles : elles combinent à la fois crise de change et crise bancaire ou crise de dette. On peut citer les exemples récents de l'Islande en 2008, des pays du Sud-Est asiatique (Thaïlande, Indonésie notamment) en 1997, des pays scandinaves (Suède et Norvège) au début des années 1990.

Le point commun de ces crises est qu'un changement brutal de taux de change affecte le risque de défaut des banques et/ou des gouvernements. Dans un monde avec libre mouvement des capitaux, il n'y a pas de raison que l'actif et le passif soit libellés dans la même monnaie ou aient la même maturité. Ainsi, une banque aura typiquement des actifs principalement en monnaie nationale et à long terme (prêts aux ménages et aux entreprises nationales), tandis qu'elle peut se refinancer à court terme en monnaie étrangère pour obtenir les taux les plus avantageux. De la même façon un gouvernement ne peut taxer que des actifs nationaux tandis qu'il peut se financer à l'étranger. Avec une telle asymétrie de devise et de maturité dans le bilan, une dépréciation subite accroît le risque de défaut des banques et des gouvernements. Par ailleurs, une BC faisant face à ce problème a tendance à émettre de la monnaie pour refinancer le secteur bancaire, ce qui renforce la dépréciation. Cet argument vaut particulièrement pour les pays qui ne parviennent pas à émettre de la dette sur les marchés financiers internationaux dans leur propre devise, comme la plupart des pays émergents.

La Figure 7.10 présente le ratio dette extérieure de court terme sur réserves de change de plusieurs pays émergents. Ce ratio mesure la capacité d'un pays à obtenir les devises nécessaires pour rembourser toute sa dette extérieure de court terme, si les prêteurs étrangers cessaient brutalement de refinancer le pays. Cette mesure capture le degré d'illiquidité des actifs d'un pays et sa fragilité financière. On constate une certaine corrélation entre ce ratio et l'occurrence de crises jumelles (Corée, Indonésie et Thaïlande en 1997, Mexique en 1994).

Le cas de l'Islande en 2008 illustre bien l'idée de crise jumelle. Après plusieurs années d'afflux massifs de capitaux, notamment britanniques et néerlandais, l'Islande connaît une crise jumelle du change et du secteur bancaire. La Figure 7.11 décrit l'inverse du ratio précédent et la valeur de la dette extérieure avant la crise (graphique de gauche) ainsi que le risque de défaut du gouvernement et des 3 principales banques, mesuré par le prix

FIGURE 7.10 – Ratio dette de court terme étrangère sur réserves de change dans 10 pays émergents, 1990-1997.

Table 1 SHORT-TERM FOREIGN DEBT/INTERNATIONAL RESERVES					
Date ^a	Ratio				
	Indonesia	Korea	Malaysia	Philippines	Thailand
1990	2.21	1.06	0.22	3.18	0.59
1994	1.73	1.61	0.25	0.41	0.99
1997	1.70	2.06	0.61	0.85	1.45
Date ^a	Ratio				
	Argentina	Brazil	Chile	Mexico	Peru
1990	2.09	2.63	0.89	2.24	3.87
1994	1.33	0.70	0.51	1.72	0.38
1997	1.21	0.79	0.45	1.19	0.50

Source: BIS, IMF.
^aJune.

Le ratio dette de court terme étrangère sur réserves de change mesure la capacité d'un pays à éviter un défaut au cas où les prêteurs internationaux cesseraient soudainement de le refinancer. Un ratio supérieur à 1 indique une forte asymétrie entre actif et passif et une fragilité financière. Source : Chang & Velasco (1999).

d'un contrat d'assurance de dette de type Credit Default Swap (CDS), avant et pendant la crise. La Figure 7.12 décrit l'ampleur de la dépréciation de la couronne islandaise vis-à-vis de l'euro, qui s'accélère début 2008, c'est-à-dire au moment où les réserves de change ne représentent plus que 7% de la dette de court terme privée et la hausse du prix des CDS des principales banque islandaises s'accélère. Après une dépréciation d'environ 35% entre janvier et octobre 2008 et deux des trois banques du pays en quasi-faillite, la Banque Centrale est intervenue pour suspendre la convertibilité. Le mois suivant le parlement a voté l'introduction de restrictions aux mouvements de capitaux et l'obligation aux exportateurs islandais de conserver leurs devises étrangères localement. En 2009 l'Islande a même porté candidature pour accéder à l'Union Européenne et à la zone euro, mais les négociations ont été mises en sommeil en attendant un référendum sur la question. Une fois de plus, ces mesures rappellent le triangle d'impossibilité.

7.3.4 Conclusion sur les crises de change

En résumé, nous avons vu trois types de crises des régimes de change fixe. Premièrement, lorsque la croissance du crédit rend les interventions sté-

FIGURE 7.11 – Illiquidité croissante de l'Islande de 2000 à 2007, à gauche ; Valeur des CDS sur la dette des 3 principales banques et du gouvernement islandais, en points de base, de janvier à décembre 2008, à droite.

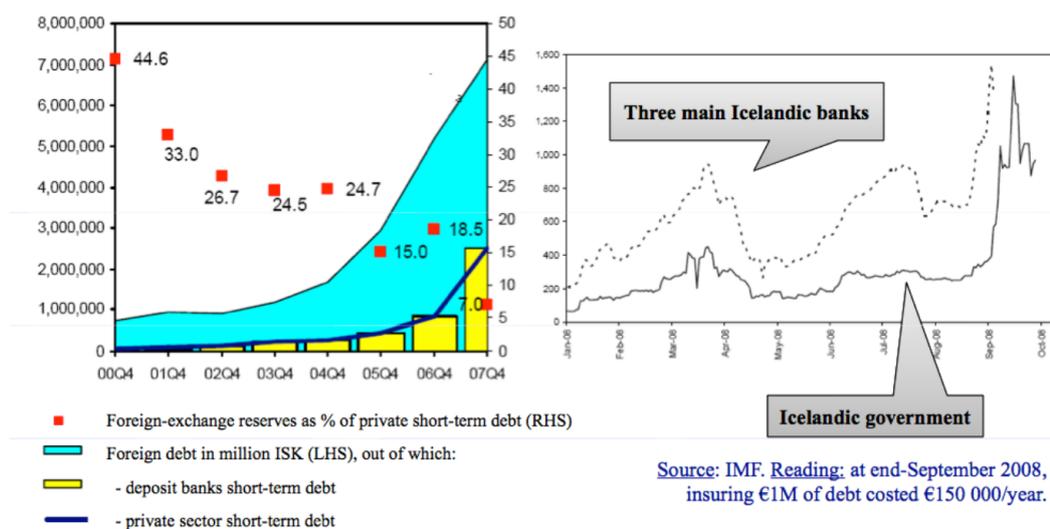


FIGURE 7.12 – Taux de change nominal EUR/ISK, 2007-2008.



rilisées de la BC de plus en plus difficiles, et que les réserves de change s'épuisent, une attaque spéculative est rationnelle. Deuxièmement, quand les agents commencent à anticiper une dévaluation, maintenir le change fixe est de plus en plus coûteux en terme de chômage, et il devient dans l'intérêt de la BC d'abandonner la défense du cours. Finalement, lorsque le secteur bancaire ou le gouvernement sont fortement endettés en devises étrangères à court terme, la BC perd sa capacité à défendre un change fixe, la priorité devenant la lutte contre un défaut des banques ou du gouvernement dont le coût serait considérable.

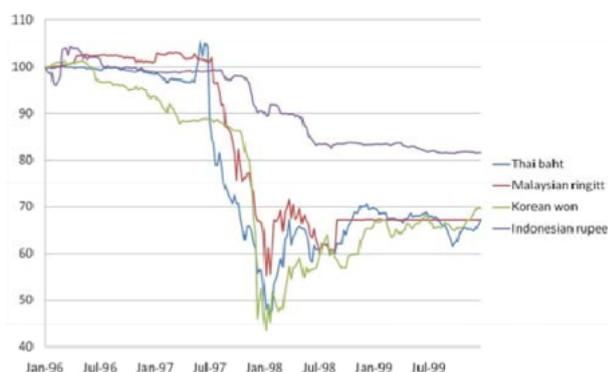
Les modèles étudiés suggèrent des recommandations de politique économique face à chacun des trois types de crises. Pour éviter une crise de 1e génération, le modèle suggère une surveillance de la croissance du crédit et de la dette publique, l'accumulation de réserves et éventuellement le recours aux prêts conditionnels du FMI. Les modèles de 2e génération suggèrent une surveillance de fondamentaux économiques élargis à l'activité réelle (chômage, capacité d'utilisation) et des mécanismes institutionnels augmentant le coût social d'une dévaluation (réputation, currency board). Enfin, face pour éviter une crise de troisième génération, il convient de réduire l'asymétrie de devises ou de maturités, par exemple en encourageant les banques à renégocier leurs maturités auprès des créanciers ou un engagement de refinancement. La BC peut également surveiller la liquidité des banques en intervenant sur le marché monétaire si nécessaire et éventuellement proposer des swaps avec d'autres BC sur une partie des réserves.

L'existence de crises de change illustre la proposition du triangle d'impossibilité de Mundell, selon laquelle seules deux des trois institutions suivantes sont compatibles : un change fixe, une politique monétaire indépendante et la libre circulation des capitaux. Par exemple, en cas d'attaque spéculative imminente un pays peut être tenté de bloquer les flux de capitaux, comme l'Argentine en 2001 peu avant la crise ; ou pour prévenir une telle crise un pays peut accumuler de fortes réserves pour dissuader une attaque, comme la Chine depuis le début des années 2000). Les crises de deuxième génération illustrent bien comme un régime de change fixe peut faire dévier un pays de sa politique monétaire optimale, quand par exemple la BC cherche à se construire une réputation, ou avec un currency board qui rend très coûteuse toute croissance de la masse monétaire.

On peut finalement mentionner le phénomène de contagion dans les crises de change et plus généralement les crises financières. Dans le cas des crises de change, on peut citer l'exemple des crises de change quasi-simultanées en Thaïlande, Corée, Indonésie et Malaisie en 1997 (voir Figure 7.13) et de crises l'année suivante en Russie et au Brésil. En 2008, la crise de change de la Hongrie a été suivie de crises de change en Ukraine et en Lettonie.

Ces phénomènes de contagion ne sont pas encore bien compris et font l'objet de recherches. Si empiriquement il est souvent difficile de distinguer causes communes et contagion dans les données, plusieurs hypothèses théoriques ont été avancées. Ces hypothèses impliquent le canal du commerce et le canal financier. Premièrement, un pays atteint par une crise de change réduit fortement ses importations provenant des partenaires commerciaux, et réduit leur compétitivité-prix sur les marchés tiers. Deuxièmement, une crise de change affecte directement les détenteurs de monnaie nationale dans d'autres pays, et indirectement toutes les contreparties des échanges financiers impliquant les banques nationales atteintes par la crise. Par ailleurs, dans le cas des pays émergents, les grandes institutions financières groupent souvent leurs investissements dans les pays émergents dans les mêmes fonds, de sorte qu'en cas de perte dans un pays émergent ces fonds vendent leurs participations dans d'autres pays émergents, transmettant ainsi une baisse du prix des actifs.

FIGURE 7.13 – Taux de change de plusieurs monnaies asiatiques en dollars US, 1996-1999.



Source : Reuters.

7.4 Conclusion

Dans ce chapitre, nous avons étudié les différences entre régimes de change fixe et régimes de change flexible. Nous avons vu que l'ajustement en change fixe est plus lent et repose sur des transferts de réserves de change, ouvrant la possibilité d'attaques spéculatives. Nous avons également étudié différents types de crises de change. Ces crises ont illustré l'impossibilité d'avoir simultanément un change fixe, une politique monétaire indépendante et une

libre circulation des capitaux. Enfin, comme évoqué dans le chapitre 6 avec le modèle Mundell-Fleming, en présence de rigidités nominales les mécanismes d'ajustement habituels échouent, et des déséquilibres courants persistants peuvent apparaître. D'autres combinaisons de politiques économiques doivent être utilisées pour stabiliser l'économie et réduire ces déséquilibres, particulièrement dans une union monétaire comme la zone euro.

Chapitre 8

Politique économique dans la zone euro

FIGURE 8.1 – Taux de change euro-dollar depuis la création de l'UME



Source : BCE. Taux de change défini en termes de dollars par euro

Bibliographie

- Arkolakis, C., Costinot, A. & Rodriguez-Clare, A. (2012), ‘New Trade Models, Same Old Gains?’, *American Economic Review* **102**(1), 94–130.
- Aviat, A. & Coeurdacier, N. (2007), ‘The geography of trade in goods and asset holdings’, *Journal of International Economics* **71**(1), 22–51.
- Balassa, B. (1963), ‘An Empirical Demonstration of Classical Comparative Cost Theory’, *Review of Economics and Statistics* **45**, 231–238.
- Balassa, B. (1964), ‘The Purchasing Parity Doctrine : A Reappraisal’, *Journal of Political Economy* **72**, 584–596.
- Baldwin, R. (1971), ‘Determinants of the Commodity Structure of U.S. Trade’, *American Economic Review* **61**, 126–145.
- Bernanke, B. S. (2005), The Global Savings Glut and the US Current Account Deficit, Sandridge lecture, Virginia Association of Economics, Richmond, Virginia.
- Blanchard, O. & Fischer, S. (1989), *Lectures on Macroeconomics*, MIT Press.
- Bowen, H. P., Leamer, E. E. & Sveikauskas, L. (1987), ‘Multicountry, Multifactor Tests of the Factor Abundance Theory’, *American Economic Review* **77**(5), 791–809.
- Buch, C. M. (2005), ‘Distance and International Banking’, *Review of International Economics* **13**(4), 787–804.
- Burstein, A. & Gopinath, G. (2014), International Prices and Exchange Rates, in G. Gopinath, E. Helpman & K. Rogoff, eds, ‘Handbook of International Economics (Vol. 4)’, Elsevier.
- Caballero, R. J., Farhi, E. & Gourinchas, P.-O. (2008), ‘An Equilibrium Model of Global Imbalances and Low Interest Rates’, *American Economic Review* **98**(1), 358–393.
- Cassel, G. (1918), ‘Abnormal Deviations in International Exchanges’, *Economic Journal* **28**, 413–415.
- Chang, R. & Velasco, A. (1999), Liquidity Crises in Emerging Markets : Theory and Policy, NBER Working Paper 7272.

- Chinn, M. D. & Frankel, J. (2002), Survey Data on Exchange Rate Expectations : More Currencies, More Horizons, More Tests, *in* D. Dickinson & W. Allen, eds, 'Monetary Policy, Capital Flows and Financial Market Developments in the Era of Financial Globalization : Essays in Honour of Max Fry', Routledge.
- Chinn, M. D. & Ito, H. (2006), 'What matters for financial development ? Capital controls, institutions, and interactions', *Journal of Development Economics* **81**(1), 163–192.
- Coeudacier, N., Guibaud, S. & Jin, K. (2012), Credit Constraints and Growth in a Global Economy, CEPR Discussion Paper 9109.
- Deardorff, A. (2003), Time and Trade : The Role of Time in Determining the Structure and Effects of International Trade, with an Application to Japan, *in* R. M. Stern, ed., 'Analytical Studies in U.S.-Japan International Economic Relations', Edward Elgar Publishing Inc.
- Disdier, A.-C. & Head, K. (2008), 'The Puzzling Persistence of the Distance Effect on Bilateral Trade', *The Review of Economics and Statistics* **90**(1), 37–48.
- Dixit, A. K. & Stiglitz, J. E. (1977), 'Monopolistic Competition and Optimum Product Diversity', *American Economic Review* **67**(3), 297–308.
- Dornbusch, R. (1976), 'Expectations and Exchange Rate Dynamics', *Journal of Political Economy* **84**(6), 1161–1176.
- Dornbusch, R., Fischer, S. & Samuelson, P. A. (1977), 'Comparative Advantage, Trade, and Payments in a Ricardian Model with a Continuum of Goods', *American Economic Review* **67**(5), 823–839.
- Eaton, J. & Kortum, S. (2002), 'Technology, Geography, and Trade', *Econometrica* **70**(5), 1741–1779.
- Feenstra, R. & Taylor, A. (2011), *International Economics (2nd Edition)*, Worth Publishers.
- Ferrero, A. (2010), 'A Structural Decomposition of the US Trade Balance : Productivity, Demographics and Fiscal Policy', *Journal of Monetary Economics* **57**(5), 478–490.
- Fisher, I. (1930), *The Theory of Interest*, Macmillan.
- Fontagné, L., Freudenberg, M. & Gaulier, G. (2006), 'A Systematic Decomposition of World Trade into Horizontal and Vertical IIT', *Review of World Economics (Weltwirtschaftliches Archiv)* **142**(3), 459–475.
- Gaulier, G., Lemoine, F. & Unal-Kesenci, D. (2006), 'Chine : le prix de la compétitivité', *La Lettre du CEPII* **254**.

- Gourinchas, P.-O. & Jeanne, O. (2013), 'Capital Flows to Developing Countries : The Allocation Puzzle', *Review of Economic Studies* **80**(4), 1484–1515.
- Gourinchas, P.-O. & Rey, H. (2014), External Adjustment, Global Imbalances and Valuation Effects, in G. Gopinath, E. Helpman & K. Rogoff, eds, 'Handbook of International Economics (Vol. 4)', Elsevier.
- Gourinchas, P.-O. & Tornell, A. (2004), 'Exchange Rate Puzzles and Distorted Beliefs', *Journal of International Economics* **64**(2), 303–333.
- Head, K. & Mayer, T. (2000), 'Non-Europe : The magnitude and causes of market fragmentation in the EU', *Review of World Economics (Weltwirtschaftliches Archiv)* **136**(2), 284–314.
- Head, K. & Mayer, T. (2014), Gravity Equations : Workhorse, Toolkit, and Cookbook, in G. Gopinath, E. Helpman & K. Rogoff, eds, 'Handbook of International Economics (Vol. 4)', Elsevier, chapter 3.
- Head, K., Mayer, T. & Ries, J. (2010), 'The erosion of colonial trade linkages after independence', *Journal of International Economics* **81**(1), 1–14.
- Heckscher, E. (1919), 'The Effect of Foreign Trade on the Distribution of Income', *Ekonomisk Tidskrift* pp. 497–512.
- Helpman, E., Melitz, M. & Rubinstein, Y. (2008), 'Estimating Trade Flows : Trading Partners and Trading Volumes', *The Quarterly Journal of Economics* **123**(2), 441–487.
- Hume, D. (1752), *Political Discourses*, R. Fleming.
- Hummels, D. (1999), Toward a Geography of Trade Costs, GTAP Working Papers 1162, Center for Global Trade Analysis, Department of Agricultural Economics, Purdue University.
- Hummels, D. (2001), Time as a Trade Barrier, GTAP Working Papers 1152, Center for Global Trade Analysis, Department of Agricultural Economics, Purdue University.
- Hummels, D. (2007), 'Transportation Costs and International Trade in the Second Era of Globalization', *Journal of Economic Perspectives* **21**(3), 131–154.
- Jeanne, O. (1996), 'Les modèles de crises de change : un essai de synthèse en relation avec la crise du franc de 1992-1993', *Economie et Prévision* **2-3**, 147–162.
- Krugman, P. (1980), 'Scale Economies, Product Differentiation, and the Pattern of Trade', *American Economic Review* **70**(5), 950–59.
- Krugman, P., Obstfeld, M. & Melitz, M. (2012), *International Economics Theory and Policy (9th Edition)*, Pearson.

- Krugman, P. R. (1979a), 'A Model of Balance-of-Payments Crises', *Journal of Money, Credit and Banking* **11**(3), 311–325.
- Krugman, P. R. (1979b), 'Increasing returns, monopolistic competition, and international trade', *Journal of International Economics* **9**(4), 469–479.
- Lane, P. R. & Milesi-Ferretti, G. M. (2007), 'The external wealth of nations mark II : Revised and extended estimates of foreign assets and liabilities, 1970-2004', *Journal of International Economics* **73**(2), 223–250.
- Lawrence, R. (1996), *Single World, Divided Nations ? : International Trade and OECD Labor Markets*, OECD Development Centre.
- Leamer, E. E. & Levinsohn, J. (1995), International Trade Theory : The Evidence, in 'Handbook of International Economics (Vol. 3)', Elsevier, pp. 1339–1394.
- Lemieux, T. (2008), 'The changing nature of wage inequality', *Journal of Population Economics* **21**(1), 21–48.
- Leontief, W. (1953), 'Domestic Production and Foreign Trade : The American Capital Position Re-examined', *Proceedings of the American Philosophical Society* **97**, 331–349.
- López-Mejía, A. (1999), 'Flux massifs de capitaux : Leurs causes, leurs conséquences et la manière de les gérer', *Finances et Développement* .
- MacDougall, G. (1951), 'British and American Exports : A Study Suggested by the Theory of Comparative Costs. Part I', *The Economic Journal* **61**, 697–724.
- McCallum, J. (1995), 'National Borders Matter : Canada-U.S. Regional Trade Patterns', *American Economic Review* **85**(3), 615–23.
- Melitz, M. J. (2003), 'The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity', *Econometrica* **71**(6), 1695–1725.
- Mody, A., Razin, A. & Sadka, E. (2002), The Role of Information in Driving FDI : Theory and Evidence, CEPR Discussion Paper 3619.
- Mussa, M. (1976), 'The Exchange Rate, the Balance of Payments and Monetary and Fiscal Policy under a Regime of Controlled Floating', *Scandinavian Journal of Economics* **78**(2), 229–248.
- Obstfeld, M. (1991), Destabilizing Effects of Exchange Rate Escape Clauses, NBER Working Paper 3603.
- Obstfeld, M. & Rogoff, K. (1996), *Foundations of International Macroeconomics*, MIT Press.
- Ohlin, B. (1933), *Interregional and International Trade*, Cambridge : Harvard University Press.

- Piketty, T. (2013), *Le Capital au XXIème siècle*, Le Seuil.
- Piketty, T. & Saez, E. (2006), 'The Evolution of Top Incomes : A Historical and International Perspective', *American Economic Review* **96**(2), 200–205.
- Portes, R. & Rey, H. (2005), 'The determinants of cross-border equity flows', *Journal of International Economics* **65**(2), 269–296.
- Rauch, J. E. (2001), 'Business and Social Networks in International Trade', *Journal of Economic Literature* **39**(4), 1177–1203.
- Rogoff, K. (1996), 'The Purchasing Power Parity Puzzle', *Journal of Economic Literature* **34**, 647–668.
- Romalis, J. (2004), 'Factor Proportions and the Structure of Commodity Trade', *American Economic Review* **94**(1), 67–97.
- Rose, A. K. (2000), 'One money, one market : the effect of common currencies on trade', *Economic Policy* **15**(30), 7–46.
- Ruoen, R. & Manying, B. (2002), 'China's Manufacturing Industry in an International Perspective : A China-Germany Comparison', *Economie Internationale* **92**, 103–130.
- Samuelson, P. (1947), *Foundations of Economic Analysis*, Cambridge : Harvard University Press.
- Samuelson, P. (1964), 'Theoretical Notes on Trade Problems', *Review of Economics and Statistics* **46**, 145–154.
- Stern, R. (1962), 'British and American Productivity and Comparative Costs in International Trade', *Oxford Economic Paper* **14**, 275–296.
- Tinbergen, J. (1962), *Shaping the World Economy : Suggestions for an International Economic Policy*, Twentieth Century Fund, New-York.
- Trefler, D. (1995), 'The Case of the Missing Trade and Other Mysteries', *American Economic Review* **85**(5), 1029–46.
- Vanek, J. (1968), 'The Factor Proportions Theory : The n-factor case', *Kyklos* **21**(4), 749–756.

**Ecole Polytechnique, 2ème Année, Eco-434 Economie
Internationale
PC 1 - L'avantage comparatif**

Exercice 1 : Avantage absolu et avantage comparatif

La table ci-dessous compare la productivité du travail aux Etats-Unis et en Chine, dans la production de trois types de biens, les vêtements, le textile et le blé. La productivité du travail dans les secteurs de l'habillement et du textile est mesurée par les ventes réalisées par employé. Dans le secteur agricole, elle est mesurée en nombre de tonnes par mois. Discuter la structure des avantages absolus et des avantages comparatifs. A quels flux de commerce peut-on s'attendre ?

	Etats-Unis	Chine
	Ventes/Employé	Ventes/Employé
Vêtements	\$ 92,000	\$ 13,500
Textile	\$ 140,000	\$ 9,000
	Tonnes/Mois	Tonnes/Mois
Blé	27.5	0.1

Exercice 2 : Le modèle ricardien

Le monde est constitué de deux pays, le "Nord" et le "Sud" (N et S). Dans chaque pays, deux biens ("1" et "2") sont produits à partir de travail. Le bien 1 est utilisé comme numéraire et on note p le prix relatif du bien 2 en termes de bien 1.

Soit y_{ij} la production du bien i dans le pays j et y_j la production totale du pays j , évaluée en unités du bien 1. Les fonctions de production sont linéaires. Il faut a_{ij} unités de travail pour produire une unité du bien i dans le pays j :

$$L_{ij} = a_{ij}y_{ij}$$

avec L_{ij} la quantité de travail employée dans le secteur i du pays j . Dans l'application numérique, on utilisera les coefficients techniques suivants : $a_{1N} = 2$, $a_{2N} = 4$, $a_{1S} = 3$ et $a_{2S} = 12$.

Chaque pays est doté en travail en quantités fixes, respectivement $L_N = 4000$ et $L_S = 9000$ dans le Nord et dans le Sud. Les deux pays partagent les mêmes préférences. Les fonctions de consommation sectorielles s'écrivent : $c_{1j} = y_j/2$ and $c_{2j} = y_j/2p$ où c_{ij} est la quantité consommée du bien i dans le pays j .

1. On appelle “Frontière des Possibilités de Production” (FPP) du pays j l’ensemble des productions sectorielles (y_{2j}, y_{1j}) qui sont atteignables étant donnée la dotation en travail L_j . Tracer la PPF dans le plan (y_{2j}, y_{1j}) . Déterminer le prix relatif p et les quantités produites et consommées de chaque bien à l’équilibre autarcique.

2. Discuter la structure des avantages comparatifs. Les deux pays signent un accord de libre-échange. Quels sont les nouveaux prix d’équilibre ?

3. Déterminer les productions sectorielles, les consommations sectorielles ainsi que la structure du commerce (exportations et importations) à l’équilibre de libre-échange. Faire une représentation graphique et l’utiliser pour discuter la nature des gains à l’échange dans ce modèles.

Exercice 3 : Deux facteurs de production

On considère une petite économie ouverte dans laquelle les entreprises produisent dans un environnement parfaitement concurrentiel. Il y a deux industries, l’aéronautique (A) et le textile (T), et deux facteurs de production, le capital et le travail, disponibles en quantités K et L . Dans chaque secteur, la fonction de production est de type Cobb-Douglas :

$$Y_A = L_A^{\alpha_A} K_A^{1-\alpha_A} \text{ et } Y_T = L_T^{\alpha_T} K_T^{1-\alpha_T}, \text{ avec } L = L_A + L_T, K = K_A + K_T, \text{ et } 0 < \alpha_i < 1, i = A/T$$

On s’intéresse à l’équilibre de court terme dans lequel l’allocation du capital entre secteurs est fixe (à court terme, il est difficile de réalouer les machines utilisées dans le secteur de l’aéronautique pour augmenter la production de textile). Sans perte de généralité, on suppose : $K_A = K_T = 1$. Au contraire, le travail peut être réaloué librement entre secteurs. Par conséquent, les salaires d’équilibre sont égalisés entre secteurs : $w_A = w_T = w$. Soit $p = P_T/P_A$ le prix relatif du textile en termes d’avions.

1) L’autarcie

Le pays est fermé au commerce international.

a. Calculer le taux marginal de transformation (TMT), i.e. la production d’avions à laquelle il faut renoncer pour produire une unité supplémentaire de textile. Comment le TMT varie-t-il avec les productions sectorielles (Y_A et Y_T) ? On suppose $0 < \alpha_A < \alpha_T < 1$, c’est-à-dire que le secteur du textile est relativement intensif en travail. Tracer la frontière des possibilités de production (FPP) dans le plan (Y_T, Y_A) .

b. Montrer que l’équilibre concurrentiel s’établit au point où la pente de la FPP est égale à $-p$.

c. Choisir un point de la FPP et tracer une courbe d’isovaleur, i.e. l’ensemble des combinaisons de productions sectorielles (Y_T, Y_A) qui conduisent à la même production totale en valeur.

d. On suppose une fonction d'utilité de type Cobb-Douglas : $U(C_A, C_T) = C_A^\beta C_T^{1-\beta}$ où C_i est la consommation de bien i ($i = A, T$). Ecrire la condition du premier ordre correspondant au programme de maximisation de l'utilité sous contrainte budgétaire. Déterminer l'équilibre général du modèle autarcique. Le représenter sur le graphique. On notera \hat{p} le prix relatif d'équilibre à l'autarcie.

2) L'économie ouverte

L'économie s'ouvre au commerce international. Le capital reste immobile, à la fois entre secteurs et au niveau international. Sous l'hypothèse d'une économie suffisamment petite, le prix relatif $p^* = P_T/P_A$ déterminé sur les marchés internationaux s'impose au pays. On suppose que celui-ci est différent du prix autarcique \hat{p} .

a. Ecrire les conditions du premier ordre correspondants au comportement des entreprises et du ménage représentatif. En déduire une condition d'équilibre de la balance courante.

b. Représenter graphiquement l'équilibre sous l'hypothèse $p^* < \hat{p}$. Commentaire.

**Ecole Polytechnique, 2ème Année, Eco-434 Economie
Internationale**

PC 2 - Le modèle Hecksher-Ohlin-Samuelson

On considère une petite économie ouverte qui produit deux biens ($i = 1, 2$) à partir de deux facteurs de production (le travail et le capital). Le capital comme le travail sont supposés mobiles entre secteurs et immobiles internationalement. Le pays est doté d'une combinaison de travail et de capital (les dotations étant notées L et K). La technologie est résumée par les fonctions de production suivantes (respectivement pour les biens 1 et 2) :

$$Y_1 = f_1(K_1, L_1) = K_1^\alpha L_1^{1-\alpha} \quad Y_2 = f_2(K_2, L_2) = K_2^\beta L_2^{1-\beta} \quad 0 < \alpha < \beta < 1$$

Dans ce qui suit, on note $k_i = K_i/L_i$ le ratio capital sur travail utilisé à l'équilibre dans la production du bien i ($i = 1, 2$), et $k = K/L$ le ratio capital sur travail disponible dans l'économie dans son ensemble. Soient w et r les coûts unitaires du travail et du capital, $p = p_1/p_2$ le prix relatif du bien 1 en termes de bien 2 et $\omega = w/r$ la compensation relative offerte au facteur travail. Les entreprises évoluent dans un environnement parfaitement concurrentiel sur les marchés de biens ainsi que sur les marchés de facteurs.

1) Déterminer le niveau optimal des ratios sectoriels k_1 et k_2 , en fonction du coût relatif du travail ω . Discuter le sens de la relation entre k_i et ω . Montrer que, à l'équilibre de la firme représentative, $\kappa = k_2/k_1$ dépend uniquement de α et β et que $\kappa > 1$. Tracer les fonctions $k_1^{-1}(\omega)$ et $k_2^{-1}(\omega)$ dans le plan (k_i, ω) . Montrer que, dès lors que ω se stabilise en dehors de l'intervalle $[\omega_{min}, \omega_{max}]$, l'économie se spécialise complètement dans la production d'un unique bien. Dans ce qui suit, on suppose que : $\omega_{min} < \omega < \omega_{max}$.

2) Le théorème de Stolper-Samuelson : Les fonctions de production se réécrivent $Y_1 = k_1^\alpha L_1$ et $Y_2 = k_2^\beta L_2$. Calculer le taux marginal de transformation ($- dY_2/dY_1 | \omega = cst$) en fonction de ω . Quelle est la relation entre p et ω à l'équilibre ? En déduire qu'une augmentation du prix relatif du bien intensif en travail s'accompagne d'une hausse de la rémunération relative du travail.

3) Le théorème de Rybczynski : En notant que $k_1 L_1 = K - k_2(L - L_1)$ et $k_2 L_2 = K - k_1(L - L_2)$, écrire L_1/L et L_2/L en fonction de k , k_1 et k_2 .

Etudier l'effet d'une variation de L sur l'emploi et la production de chaque secteur, à rémunération relative du travail (ω) et à dotation en capital (K) données. En déduire que la production du bien intensif en travail augmente

lorsque la force de travail s'élargit / que la production du bien intensif en capital se réduit lorsque la force de travail s'élargit.

4) Montrer qu'on peut écrire la production sectorielle relative : $\frac{Y_1}{Y_2} = -\frac{k_1^\alpha}{k_2^\beta} \left(\frac{k-k_2}{k-k_1} \right)$.

On suppose que les fonctions d'utilité sont telles que le budget des ménages se répartit équitablement entre les deux secteurs ($C_2 = pC_1$). Quel est le prix relatif qui s'établit à l'équilibre autarcique, selon la valeur de ω ? Comment ce prix varie-t-il en fonction des dotations relatives en facteurs k ?

5) On suppose maintenant que le monde est constitué de deux pays. Les dotations en capital et en travail dans le pays domestique et à l'étranger sont telles que $k < k^*$, où k^* correspond au ratio capital sur travail du pays étranger. Comment le prix relatif du bien 1 évolue-t-il dans les deux pays, après l'ouverture au commerce? Implications pour les prix relatifs du travail (ω et ω^*)? Discussion.

**Ecole Polytechnique, 2ème Année, Eco-434 Economie
Internationale
PC 3 - Concurrence imparfaite**

Intégration européenne et “Effet-frontière”

Dans cet exercice, nous utilisons un cadre empirique pour étudier l’effet sur le commerce de l’intégration européenne, notamment de l’Acte Unique Européen entré en vigueur en juillet 1987. Pour cela, nous utilisons le modèle de commerce en concurrence imparfaite pour dériver théoriquement une équation de gravité. Ce cadre théorique est ensuite appliqué à des données de commerce bilatéral intra-européen pour estimer l’effet de l’intégration européenne.

1ère partie : Dérivation de l’équation de gravité

On considère un modèle à C pays, indexés $i \in \{1..C\}$. Le revenu nominal du pays j est noté Y_j . Ce revenu est utilisé pour consommer les différentes variétés d’un bien différencié produites au niveau mondial. Dans chaque pays, N_i variétés du bien sont produites et vendues sur les marchés internationaux, chaque variété étant indexée $k(i)$. Les préférences du consommateur représentatif dans le pays j sont les suivantes :

$$U_j = \left[\sum_{i=1}^C \sum_{k(i)=1}^{N_i} c_{k(i)j}^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}$$

où $\sigma > 1$ est l’élasticité de substitution, supposée constante, entre les différentes variétés du bien. $c_{k(i)j}$ est la consommation par le consommateur j de la variété $k(i)$ produite dans le pays i .

Dans ce qui suit, on note $p_{k(i)j}$ le prix de vente dans le pays j d’une variété k produite dans le pays i . A l’équilibre, ce prix inclut le coût de transport, supposé de type *iceberg* dans ce qui suit : Pour vendre une unité de bien dans le pays j , une firme localisée dans le pays i doit produire $\tau_{ij} \geq 1$ unités de biens, une partie étant “absorbée” pendant le transport.

1) La fonction de demande : Montrer que la demande adressée par le consommateur représentatif du pays j au producteur de la variété $k(i)$ s’écrit :

$$c_{k(i)j} = \left(\frac{p_{k(i)j}}{P_j} \right)^{-\sigma} \frac{Y_j}{P_j}$$

où

$$P_j = \left[\sum_{i=1}^C \sum_{k(i)=1}^{N_i} p_{k(i)j}^{1-\sigma} \right]^{\frac{1}{1-\sigma}}$$

est l'indice de prix idéal dans le pays j .

2) Comportement des entreprises : On suppose que, pour produire, les entreprises du pays i doivent faire face à un coût fixe F et un coût variable a_i , les deux éléments de coûts étant payés en unités de travail au salaire w_i . Dériver le prix optimal de vente choisi par le producteur de la variété $k(i)$ pour le marché j ($p_{k(i)j}$).

3) Commerce bilatéral : En déduire que le commerce bilatéral (en valeur) entre deux pays i et j (X_{ij}) peut s'écrire sous une forme gravitaire de la forme :

$$\ln X_{ij} = C + \mathbf{X}_i + \mathbf{M}_j + \alpha \ln \tau_{ij}$$

où C est une constante, \mathbf{X}_i et \mathbf{M}_j sont des fonctions qui ne dépendent que des variables spécifiques à l'exportateur i et à l'importateur j , respectivement, et α est un coefficient qui dépend des paramètres structurels (ces différents éléments étant à définir explicitement en fonction des différentes variables et des paramètres du modèle). Commentaire sur l'effet des coûts à l'échange sur le commerce international.

2ème partie : Effet-frontière et intégration européenne

Le texte suivant est issu d'un article scientifique qui cherche à évaluer l'effet de l'Acte Unique Européen sur le degré d'intégration économique des économies européennes. Dans ce qui suit, nous utilisons les résultats de cette étude pour discuter comment le cadre théorique de la 1ère partie peut-être utilisé empiriquement pour mesurer l'impact de l'intégration européenne.

“Dans quelle mesure l'Acte unique a-t-il renforcé l'intégration du marché européen unique ? Pour répondre à cette question, il nous faut d'abord définir une mesure de l'intégration commerciale d'un ensemble économique. Notre mesure est inspirée de la définition proposée par la commission européenne dans un rapport récent, “Le Marché intérieur. Dix ans sans frontières” (Commission européenne, 2003). Pour la commission, le but du Marché unique est de supprimer toute signification économique aux frontières nationales des Etats membres. Le (manque d')intégration se mesure donc à travers l'impact de ces frontières et l'évolution de cet impact depuis le passage au Marché unique.

[...] Notre mesure utilise les flux d'échange entre pays de l'Union européenne et à l'intérieur de ces pays pour évaluer l'effet global des frontières

nationales sur les échanges à l'intérieur de l'UE. Nous évaluons ainsi de façon détaillée le niveau d'intégration du marché européen et son évolution par secteur.

[...] [Méthodologie :] *En commerce international, un marché est défini comme unique si les frontières nationales n'ont aucun impact ni sur l'endroit où les consommateurs choisissent d'acheter leur produit ni sur la destination vers laquelle les producteurs écoulent leur production. Afin de définir l'impact résiduel des frontières et donc le degré de fragmentation des marchés, il nous faut disposer d'une référence permettant de dire ce que le commerce "devrait être". A partir de cette norme, il est possible de mesurer dans quelle mesure les frontières nationales réduisent les volumes échangés. L'équation de gravité constitue une norme naturelle, en raison de son importante capacité de prédiction sur le plan empirique. McCallum (1995) fut le premier à utiliser ce cadre, en l'appliquant au cas nord-américain, à savoir les échanges des provinces canadiennes entre elles comparés aux flux entre provinces et Etats américains. [...] L'idée est d'insérer les flux intranationaux dans une équation de gravité à côté de tous les flux internationaux afin d'identifier l'impact des frontières en mesurant le surplus de commerce intranational.*

[...] *Cette méthode inverse donc la logique des équations de gravité traditionnelles. La plupart du temps, l'équation de gravité est utilisée pour mesurer dans quelle mesure le commerce entre pays membres d'accords régionaux (l'Union européenne par exemple) est supérieur au commerce dans le reste du monde. La méthode des effets frontières considère comme norme de référence l'espace le plus intégré possible : la nation.*"

4) L'effet-frontière : A partir du texte, expliquer en quoi l'estimation de l'équation de gravité introduite ci-dessus, appliquée à des données relatives au commerce intranational et au commerce international (ie sur X_{ii} et X_{ij} , $i \neq j$), peut permettre de mesurer le degré d'intégration d'une zone économique. Comment spécifier les coûts à l'échange τ_{ij} pour mesurer un "effet-frontière" ?

5) Mesure de l'ampleur des restrictions aux échanges : Soit $\phi_{ij} = \tau_{ij}^{1-\sigma}$ une mesure de "liberté" du commerce entre i et j , inversement corrélée à l'ampleur des frictions à l'échange international. On suppose pour cette question que les coûts de transport sont symétriques ($\tau_{ij} = \tau_{ji}$) et que les coûts à l'échange intranational sont nuls ($\tau_{ii} = 1$). Montrer que, sous ces hypothèses et dans le modèle de la partie 1, la "liberté" du commerce peut être exprimée comme une fonction des flux bilatéraux de commerce entre deux pays et de la consommation de biens produits domestiquement (i.e. exprimer ϕ_{ij} en fonction de X_{ij} , X_{ji} , X_{ii} et X_{jj}).

Le tableau 8.1 reporte l'ampleur de l'indicateur de liberté du commerce international (ϕ_{ij}) obtenu avec cette méthode. Les données utilisées couvrent

TABLE 8.1 – “Liberté” du commerce (ϕ_{ij}), par secteur et année (médiane des coefficients bilatéraux obtenus sur les 4 plus grands pays de l’Union Européenne, Allemagne, France, Grande-Bretagne et Italie)

Industrie	1980	1985	1990	1995	2000
Chimie industrielle	6,55	8	13,3	16,4	16,55
Matériel de transport	5,2	7,9	10,3	14,1	10,75
Instruments de mesure	11,85	14	39,7	13,25	6,5
Équipements électriques	2,95	3,85	5,45	9,25	8,6
Cuir	4,65	5,65	7,7	9,1	10,3
Machines (hors électriques)	5,65	8,7	9,15	8,45	11,75
Textiles	4	4,1	6,35	7,4	5,7
Caoutchouc	3,45	4,15	6,2	7,3	8,85
Verre	3,55	4,5	6,45	6,9	5,3
Métaux non ferreux	5,75	4,15	5,85	6,85	10,6
Sidérurgie	2,25	2,95	5,4	6,25	5,65
Habillement	5,65	5,5	5,3	5,25	5,4
Chimie (autre)	2,55	3,4	5	5,25	8
Faïence	2,8	3	3,9	4,4	3,4
Industrie du papier	1,75	2,5	2,9	4,35	4,7
Habillement	2,05	2,2	4,4	3,6	4
Boissons	1,85	2,4	2,75	3	3,05
Produits métalliques	2,35	2,75	3,1	2,85	2,3
Plastiques	2,05	1,95	2,45	2,3	2,3
Produits alimentaires	1,05	1,2	1,7	1,7	1,9
Meubles	1,35	1,55	2,25	1,45	1,2
Produits minéraux non métalliques	1,3	1,45	1,8	1,3	1,2
Bois (hors meubles)	0,85	1,05	1,4	1,2	1,3
Imprimerie Édition	0,75	0,75	0,85	0,7	0,8
Raffinage de pétrole	1,35	1,45	1	0,65	0,85
Tabac	0,35	0,65	0,7	0,6	0,6

La liberté du commerce ϕ_{ij} est mesurée à partir de données sur les flux bilatéraux réciproques (X_{ij} et X_{ji}) et les flux intranationaux (X_{ii} et X_{jj}) pour les différents couples $i * j$ impliquant un des quatre pays listés ci-dessus. Les nombres reportés dans la table correspondent à la médiane entre couples $i * j$, pour chaque secteur et chaque année. On normalise la “liberté” du commerce intranational à 100 (ie $\phi_{ii} = 100$) de sorte que les nombres reportés pour ϕ_{ij} sont exprimés en pourcentage du niveau maximum atteignable.

différentes industries et sont relatives aux échanges bilatéraux et à la consommation sectorielle de produits domestiques pour les 4 plus grands pays de l'Union Européenne (Allemagne, France, Grande-Bretagne et Italie). Pour chaque couple de pays, on calcule une valeur de ϕ_{ij} à partir des échanges bilatéraux réciproques (X_{ij} et X_{ji}) et des échanges intranationaux (X_{ii} et X_{jj}) en utilisant la formule dérivée ci-dessus. Les chiffres reportés dans la table correspondent au résultat médian obtenu sur les différentes paires ainsi formées. Les résultats sont multipliés par 100 de façon à ce que l'indicateur de liberté soit exprimé en pourcentage du niveau maximum de liberté du commerce (i.e. $\phi_{ii} = 100$ correspondant au commerce intra-national de ces différents pays).

Commenter les résultats. Vous pouvez pour cela utiliser comme point de comparaison les résultats relatifs au commerce bilatéral entre les provinces canadiennes et les Etats américains obtenus à la suite de l'article de McCallum (1995) et qui font état d'un indice $\phi_{US,CAN}$ entre 17 et 20. Wolf (2000) étudie quant-à-lui le commerce entre Etats américains, en utilisant comme référence la valeur des échanges à l'intérieur d'un état américain. En moyenne, il obtient une valeur égale à 33.

6) Estimation de l'effet-frontière : On suppose maintenant que les coûts à l'échange peuvent être représentés par la fonction suivante :

$$\tau_{ij} = dist_{ij}^{\delta} \eta^{B_{ij}}$$

où $dist_{ij}$ est la distance entre le pays i et le pays j et B_{ij} une indicatrice égale à 1 si $i \neq j$. Les coûts à l'échange domestique ne sont plus nécessairement nuls ($\tau_{jj} \neq 1$) et $dist_{jj}$ mesure la distance "interne" du pays j qui permet de rendre compte de la taille de l'économie nationale.

Ecrire le ratio du commerce à la consommation domestique (X_{ij}/X_{jj}) obtenu sous cette hypothèse et en déduire une version modifiée de l'équation de gravité discutée dans la question 3). A quelle valeur de η correspond une situation d'intégration parfaite des économies nationales entre elles ?

On suppose dans ce qui suit que l'économètre dispose de données sur le commerce intra- et international et qu'il est également capable d'observer le nombre de variétés produites dans une économie, le niveau des salaires et de la productivité du travail dans chaque pays, ainsi que les distances bilatérales et internes. Expliquer comment l'estimation de l'équation de gravité modifiée permet d'estimer l'"effet-frontière" décrit dans le texte. En quoi cette stratégie empirique donne-t-elle une idée plus précise de l'"effet-frontière" que la mesure de ϕ_{ij} utilisée dans la question 5) ? Discuter.

Le tableau 8.2 reporte la valeur estimée de cet effet-frontière, pour les membres de l'Union Européenne, pour différentes périodes. Afin de maintenir un échantillon constant de pays, l'analyse est restreinte aux 9 premiers

membres de l'Union Européenne, participant déjà à l'expérience d'intégration européenne au début de la période, en 1976. Les données sont divisées en trois sous-périodes : 1976-1986 ie la période pré-Acte Unique Européen, 1987-1993 ie la période après l'Acte Unique Européen mais avant la mise en place du Marché Unique pendant laquelle les dispositions du Marché Unique ont progressivement été mises en place et 1994-2000 ie la période post-Marché Unique. Les nombres reportés correspondent à l'exponentielle de moins la constante estimée. Commenter les résultats.

TABLE 8.2 – Valeurs estimées de l'effet-frontière (données relatives aux 9 premiers membres de l'Union Européenne)

Industrie	Effet frontière			Δ entre périodes (%)	
	76-86 (1)	87-93 (2)	94-00 (3)	2/1	3/2
Meubles	50,9	110,5	21,9	117	- 80,2
Chaussures	61,8	26,8	5,7	- 56,6	- 78,6
Cuir	8,2	4,8	1,5	- 41,5	- 69,1
Caoutchouc	18,7	7,8	2,6	- 58,2	- 67,1
Machines électriques	24,7	19,6	8,2	- 20,8	- 58,2
Sidérurgie	18,3	16,2	9,5	- 11,2	- 41,4
Chimie (autre)	13,1	13,4	8,2	2,2	- 38,7
Boissons	31,8	22	14,5	- 30,6	- 34,1
Habillement	28,5	18,1	13,2	- 36,7	- 26,9
Verre	11,3	7,2	5,9	- 36,2	- 18,4
Produits minéraux non métalliques	40,7	34,2	28,8	- 16,1	- 15,9
Textile	17,7	10,2	8,8	- 42,6	- 13,6
Machines hors électriques)	20,4	7,9	6,9	- 61,5	- 12,1
Valeur médiane	19,5	18,8	13,1	- 13,6	- 12,1
Tabac	35,6	41,9	36,9	17,7	- 12
Métaux non ferreux	6,6	33,1	29,2	398,4	- 11,9
Produits métalliques	38,6	50,2	45	30,1	- 10,4
Faïence	6,6	14,2	12,9	115,9	- 8,7
Plastiques	50,2	50,4	52,9	0,4	4,9
Matériel de transport	6,3	7,6	8,3	19,6	9,5
Imprimerie Édition	89,7	114	126,4	27,1	10,8
Produits alimentaires	34,7	33,3	39,1	- 4,1	17,4
Papier	13,5	13,7	16,5	1,1	20,7
Chimie	8,3	3,7	4,6	- 54,9	24,4
Bois (hors meubles)	90,3	54,5	68,8	- 39,7	26,3
Raffinage de pétrole	7,8	27,1	80,1	249,1	195,2
Instruments de mesure	2,3	0,4	2,4	- 81,6	473,4

Note : L'effet frontière est estimé comme l'exponentielle de moins la constante obtenue après estimation de l'équation de gravité utilisant comme variable expliquée le log de X_{ij}/X_{jj} .

ECO434, Ecole polytechnique, 2e année
PC 4
Approche Intertemporelle du Compte Courant

Exercice 1 : Choix intertemporel et Taux de Change Réel (TCR)

On considère une *petite économie ouverte* dans laquelle deux biens, le bien échangeable T et le bien non-échangeable N , sont produits (Y_T et Y_N) et consommés (C_T et C_N). On note P_N et P_T les prix des deux biens et $Q \equiv \frac{P_N}{P_T}$ le prix relatif du bien non-échangeable.

Les exportations *nettes* sont notées X . Dans la première partie de l'exercice, on suppose que ces exportations nettes sont *exogènes*. Cette hypothèse simplificatrice permettra une première exploration du lien entre prix relatif des échangeables et balance commerciale. Dans la deuxième partie de l'exercice, cette hypothèse sera relâchée.

Modèle statique

1. Ecrire les conditions d'équilibre des marchés des deux biens. Exprimer R , la valeur de la production nationale, en unités du bien T (c'est-à-dire que le bien T est considéré comme le *numéraire*).
2. On suppose que la fonction d'utilité du consommateur s'écrit

$$u(C_T, C_N) = \frac{1}{4} \ln(C_T) + \frac{3}{4} \ln(C_N)$$

Montrer que le budget du consommateur s'écrit $R - X$ et interpréter. Résoudre le programme du consommateur. Calculer la part de la demande nationale dépensée pour chaque bien.

3. On considère la production de chaque bien comme *fixe*, et on suppose que $Y_N = 3Y_T$. Montrer que

$$\frac{X}{Y_T} = 1 - Q$$

Expliquer quel est l'impact d'une hausse du prix relatif des échangeables sur les exportations nettes.

4. Calculer la part de la production du bien T dans la production totale (exprimée en unités de numéraire) lorsque $Q = 1$. Même question lorsque $Q = 1,33$. Si une économie a un déficit courant égal à 6% du PIB (comme les Etats-Unis en 2007), quel ajustement suggère ce petit modèle?

Modèle de choix intertemporel On considère maintenant un modèle à deux périodes ($t = 1, 2$). Soit la fonction d'utilité intertemporelle

$$U(C_T^1, C_N^1, C_T^2, C_N^2) = u(C_T^1, C_N^1) + \beta u(C_T^2, C_N^2)$$

avec $0 < \beta < 1$. Le consommateur de la petite économie ouverte peut prêter ou emprunter à l'étranger au taux exogène r .

1. Quelle est l'interprétation de β ? Montrer que la contrainte de budget intertemporelle prend la forme

$$Q^1 C_N^1 + C_T^1 + \frac{Q^2 C_N^2 + C_T^2}{1+r} \leq \Omega$$

où on définira Ω .

2. Montrer que

$$\frac{C_N^2}{C_N^1} = \beta(1+r) \frac{Q^1}{Q^2}$$

$$\frac{C_T^2}{C_T^1} = \beta(1+r)$$

et interpréter.

3. Montrer que si la production de chacun des deux biens est constante au cours du temps et si $\beta(1+r) < 1$, alors le solde courant est négatif à $t = 1$.
4. Sous ces deux mêmes hypothèses, comment le prix relatif des non-échangeables varie-t-il entre $t = 1$ et $t = 2$? Interpréter.

Exercice 2 : Modèle de Fisher avec investissement

On considère une extension simple du modèle de Fisher à 2 périodes vu en cours (section 4.2.1. du polycopié).

La principale différence est que le revenu national à chaque période n'est plus exogène, mais suit la fonction de production strictement croissante et strictement concave $F(K)$, où K désigne le stock de capital national, et satisfaisant les "conditions d'Inada" : $F(0) = 0$, $\lim_{K \rightarrow 0} F'(K) = +\infty$, $\lim_{K \rightarrow +\infty} F'(K) = 0$. On suppose que le bien de consommation peut aussi être utilisé comme capital sans cout d'ajustement, de sorte que le prix d'une unité de capital est égal au prix d'une unité de bien de consommation (hypothèse peu réaliste mais évitant des complications inutiles ici). On définit l'investissement I_t par :

$$K_{t+1} = K_t + I_t, t = 1, 2$$

On fait l'hypothèse simplificatrice que le capital ne se déprécie jamais. On suppose également $K_1 > 0$ exogène, et, comme il n'y a pas de période 3, $K_3 = 0$.

L'économie nationale est une *petite économie ouverte* qui prête et emprunte à l'étranger au taux exogène r . On note B_t la position extérieure nette. On suppose que le capital productif K_t est entièrement détenu par des résidents.

1. Ecrire la contrainte de budget intertemporelle et déterminer le solde courant aux périodes 1 et 2.
2. Ecrire le programme de l'agent représentatif comme un programme de maximisation d'utilité en C_1, C_2, K_2 . Montrer qu'à l'optimum on investit jusqu'au point où le rendement du capital est égal aux taux d'intérêt mondial r . Commenter.
3. Montrer que si le pays avait été en autarcie financière (avec pour seul actif le capital physique), la relation

$$C_2 = F[K_1 + F(K_1) - C_1] + (K_1 + F(K_1) - C_1)$$

aurait été vérifiée. Tracer une "frontière des possibilités de production intertemporelle" dans l'espace $\{C_1, C_2\}$. Représenter graphiquement les consommations et le taux d'intérêt à l'équilibre d'autarcie financière. Commenter brièvement.

4. Quelle est la pente de la frontière des possibilités de production à l'équilibre d'ouverture financière étudié plus haut ? Représenter graphiquement la situation d'ouverture financière et le gain à l'échange intertemporel. Prédire l'évolution du solde courant selon que le taux d'intérêt autarcique est inférieur ou supérieur au taux mondial.

ECO434, Ecole polytechnique, 2e année
PC 5

Flux de Capitaux Internationaux et Déséquilibres Mondiaux

Exercice 1 : Flux de capitaux dans le modèle de croissance néoclassique

Le modèle est en temps continu. On considère un agent représentatif ayant une durée de vie infinie (dynastie) avec la fonction d'utilité :

$$U_t = \int_t^{+\infty} e^{-\rho(s-t)} N_s \frac{(c_s)^{1-\frac{1}{\sigma}}}{1-\frac{1}{\sigma}} ds \quad (8.1)$$

où c_s désigne la consommation entre s et $s + ds$, $\sigma > 0, \sigma \neq 1$ est l'élasticité de substitution intertemporelle, $\rho > 0$ désigne la préférence pour le présent. On suppose que le nombre de travailleurs N_s croît dans le temps à un taux constant exogène, i.e. $\frac{\dot{N}_t}{N_t} = n, \forall t$ et $N_0 = 1$.

La production requiert du capital et du travail suivant la fonction de production

$$Y_t = (K_t)^\alpha (\xi_t N_t)^{1-\alpha} \quad (8.2)$$

où $K_0 > 0$ et $\xi_t > 0$ croît au taux constant et exogène g .

Le capital se déprécie au cours du temps au taux δ , de sorte que

$$\dot{K}_t = I_t - \delta K_t \quad (8.3)$$

On suppose que le marché du capital est concurrentiel.

On a l'identité comptable :

$$Y_t = C_t + I_t + B_t \quad (8.4)$$

avec $C_t = N_t c_t$ et B_t la position nette extérieure, égale à zéro en situation d'autarcie.

On suppose que le travail est immobile entre pays et pour simplifier que le bien de consommation a un prix constant normalisé à 1.

1. Résoudre le programme de maximisation d'utilité du planificateur central en autarcie et montrer que

$$\frac{\dot{c}_t}{c_t} = \sigma \left(\alpha \left(\frac{K_t}{\xi_t N_t} \right)^{\alpha-1} - \delta - \rho \right)$$

Interpréter cette expression en rapport avec le taux d'intérêt d'équilibre en autarcie et la patience de l'agent représentatif.

2. On appelle capital par unité de travail efficient la grandeur $\tilde{k}_t = \frac{K_t}{\xi_t N_t}$ et revenu par unité de travail efficient la grandeur $\tilde{y}_t = \frac{Y_t}{\xi_t N_t}$.

En utilisant la condition précédente, calculer les valeurs de \tilde{k}_t et \tilde{y}_t à l'état stationnaire en autarcie, qu'on notera respectivement \tilde{k} et \tilde{y} . Montrer que le taux d'intérêt d'équilibre à l'état stationnaire vérifie $r^a = \rho + \frac{g}{\sigma}$ et commenter brièvement.

3. On considère maintenant le cas d'une petite économie ouverte. On note r le taux d'intérêt mondial. On note r^a le taux d'intérêt en autarcie à l'état stationnaire. Quelles valeurs prennent \tilde{k} et \tilde{y} ? Montrer que

$$\frac{\dot{c}_t}{c_t} = \sigma(r - \rho) = \sigma(r - r^a) + g$$

Commenter brièvement.

4. On note W_t la richesse nationale du pays définie comme la somme du stock de capital et de la position nette extérieure notée B_t . Montrer que la richesse nationale par unité de travail efficient \tilde{w}_t suit

$$\dot{\tilde{w}}_t = (r - n - g)\tilde{w}_t + (1 - \alpha)\tilde{y} - \tilde{c}_t$$

où \tilde{c}_t désigne la consommation agrégée par unité de travail efficient.

5. On fait l'hypothèse que le reste du monde a atteint l'état stationnaire et croît au taux du progrès technique g^w , avec une croissance de la population égale à n . On admettra que lorsque $r > n + \max\{g, g^w\}$ on peut écrire la consommation par unité de travail efficient :

$$\tilde{c}_t = (r - n - g^w) \left(\tilde{w}_t + \frac{(1 - \alpha)\tilde{y}}{r - n - g} \right)$$

Montrer que la position extérieure nette (exprimée en unités de travail efficient) suit

$$\dot{\tilde{b}}_t = \left(\tilde{w}_0 + \frac{(1 - \alpha)\tilde{y}}{r - n - g} \right) e^{(r - r^a)\sigma t} - \frac{(1 - \alpha)\tilde{y}}{r - n - g} - \tilde{k}$$

et commenter brièvement. Sachant que le solde courant s'écrit \dot{B}_t , décrire comment le solde courant et la position extérieure nette varient selon que la différence $r - r^a$ est positive, négative ou nulle.

6. En conclusion, comment la croissance de la productivité affecte-t-elle le sens des flux financiers internationaux ?

Exercice 2 : Flux de capitaux et démographie

On s'intéresse aux flux de capitaux dans une petite économie ouverte à deux générations : travailleurs (T) et retraités (R). On se place en temps continu. On fait plusieurs hypothèses simplificatrices pour que l'évolution démographique soit capturée par les paramètres $0 < \lambda < 1$ et $0 < \theta < 1$:

- à chaque période, une fraction λ du nombre existant de travailleurs naît, et devient immédiatement travailleur
- à chaque période, une fraction λ des travailleurs prend sa retraite (indépendamment de l'âge)
- à chaque période, une fraction θ des retraités meurt ; aucun travailleur ne meurt
- le ratio de dépendance économique (nombre de retraités par travailleur) initial est égal à $\frac{\lambda}{\theta}$

On suppose aussi pour simplifier que les agents ne consomment qu'à leur dernière période (au moment de mourir) et que le revenu par période (exogène) vaut Y_t , dont une fraction δ représente la taille du secteur financier. On note r_t le rendement du seul actif financier disponible. Enfin, on suppose que la productivité croît au taux constant g . On note W_t^R et W_t^T la richesse financière d'un retraité et d'un travailleur en t , respectivement.

1. Montrer que les hypothèses démographiques impliquent une population et un ratio de dépendance constants. Quel paramètre capture le vieillissement de la population ?
2. On se place initialement en autarcie. Montrer que

$$\begin{aligned} Y_t &= \theta W_t^R \\ \dot{W}_t^R &= (r_t - \theta)W_t^R + \lambda W_t^T \\ \dot{W}_t^T &= (r_t - \lambda)W_t^T + (1 - \delta)Y_t \end{aligned}$$

3. A l'état stationnaire, la richesse financière de chaque groupe doit croître au taux g . En déduire une équation du second degré déterminant la valeur de r_t en fonction des paramètres $\{\delta, \lambda, \theta, g\}$ et en déduire le taux d'intérêt autarcique.
4. Quel impact du vieillissement de la population sur le taux d'intérêt autarcique prédit le modèle ? Commenter brièvement. Dans quelle mesure ce résultat peut-il expliquer que certains pays à forte croissance exportent des capitaux à l'ouverture, contrairement à la prédiction du modèle de Ramsey ?

ECO434, Ecole polytechnique, 2e année
PC 7
Crises de change

Exercice : Une comparaison entre deux modèles de crises de change¹

On considère le modèle suivant d'une *petite économie ouverte* :

$$m_t - p_t = \phi y - \beta i_t \quad \phi > 0, \beta > 0, t = 1, 2 \quad (8.5)$$

$$p_t = p^* - s_t, t = 1, 2 \quad (8.6)$$

$$i_1 = i^* + s_1 - s_2^e \quad (8.7)$$

$$i_2 = i^* \quad (8.8)$$

où toutes variables sauf i_t et i^* sont en logarithmes. m_t désigne (le log de) l'offre de monnaie, p_t correspond au niveau des prix, y correspond au niveau de production de long terme. On note s_t le log des taux de change en monnaie étrangère (au certain), donc une hausse de s_t représente une appréciation nominale. s_2^e est la valeur anticipée de s_2 à la date $t = 1$.

On considère que la Banque Centrale (BC) maintient un taux de change nominal fixe $s_1 = s_2 = \bar{s}$. On cherche à étudier sa viabilité.

On suppose que la PPA est vérifiée initialement, ou

$$p_1 = p^* - \bar{s}$$

On suppose également que les prix sont rigides à $t = 1$, mais flexibles à $t = 2$, tandis que les taux d'intérêt sont toujours flexibles. Enfin, on suppose dans tout l'exercice que les anticipations sont rationnelles.

1. Un modèle d'attaque spéculative

1. Rappeler brièvement la signification des trois premières équations.
2. Soit $\bar{m} = \phi y + p^* - \beta i^* - \bar{s}$. Quelle est l'interprétation économique de cette variable ? Exprimer i_1 comme une fonction simple de m_1 , i^* , \bar{m} et des paramètres. A partir de cette relation, montrer que le taux de change flexible serait théoriquement égal à

$$s_1 = \bar{s} + \bar{m} - m_2^e - \frac{1}{\beta}(m_1 - \bar{m}) \quad (8.9)$$

$$s_2 = \bar{s} + \bar{m} - m_2 \quad (8.10)$$

1. Adapté de l'article d'Olivier Jeanne (1996), "Les modèles de crise de change : un essai de synthèse", *Economie et prévision*, no 123-124, pages 147-162.

Comment des chocs sur l'offre de monnaie en $t = 1$ et $t = 2$ affectent-ils ce taux de change flexible théorique ?

3. On considère que les actifs de la Banque Centrale se répartissent entre réserves de change r_t et prêts au secteur bancaire d_t (en logarithmes). On considère que le crédit aux banques croît au taux μ exogène, de sorte que

$$m_t = \theta r_t + (1 - \theta)d_t \quad 1 > \theta > 0 \quad (8.11)$$

$$d_2 = d_1 + \mu, \quad \mu > 0 \quad (8.12)$$

Montrer qu'il y aura une attaque spéculative sur le taux de change fixe si

$$\frac{1}{\beta}((1 - \theta)d_1 - \bar{m}) + ((1 - \theta)d_2^e - \bar{m}) > 0 \quad (8.13)$$

Pourquoi l'attaque spéculative se produit-elle en $t = 1$?

4. Discuter brièvement quelles variables économiques fondamentales devraient faire l'objet d'une surveillance pour évaluer le risque d'attaque spéculative contre une monnaie.

2. Un modèle de crise de change autoréalisatrice

On considère un second modèle dans lequel la Banque Centrale (BC) est moins passive face au risque de crise de change.

On suppose que la BC défend le taux de change fixe à $t = 1$ dans tous les cas, mais choisit de dévaluer ou pas en $t = 2$. Pour simplifier les calculs on fixe le montant de la dévaluation potentielle de sorte que

$$s_2 = s_1 - \Delta \quad \Delta > 0$$

si la BC choisit de dévaluer.

On considère que la BC prend à la date $t = 2$ la décision qui minimise la fonction de perte L_2 définie par

$$L_2 = [U_2(s_2)]^2 + C\mathbf{1}_{s_2 - s_1 = \Delta} \text{ avec } C > 0$$

où U_2 désigne le taux de chômage en $t = 2$, $\mathbf{1}$ est une variable indicatrice qui prend la valeur 1 si la BC dévalue ($s_2 - s_1 = \Delta$), et 0 sinon, et C représente un coût politique à la dévaluation.

On suppose que le chômage est persistant et dépend de l'inflation non anticipée $\pi - \pi^e$, par exemple parce que les salaires sont fixés à l'avance par négociation et l'inflation non anticipée réduit le salaire réel. Plus précisément, U_2 suit l'équation

$$U_2 = \rho U_1 - \lambda(\pi_2 - \pi_2^e) \text{ avec } 0 < \rho < 1, 0 < \lambda < 1$$

U_t, π_t, π_t^e, C ne sont pas en logarithmes.

1. Si les agents anticipent que la BC ne dévaluera pas (i.e. que le taux de change restera constant), sous quelle condition la BC défendra-t-elle effectivement le taux de change en $t = 2$? [il peut être utile de définir $\Psi(U_1) = \frac{C}{\lambda\Delta} - 2\rho U_1$]
2. Si les agents anticipent une dévaluation, quelle est la valeur de π_2^e ? Sous quelle condition la BC dévaluera-t-elle en $t = 2$?
3. Pour quelles valeurs des paramètres a-t-on des équilibres multiples (c'est-à-dire plusieurs actions différentes cohérentes avec des anticipations)? Commenter.
4. Par rapport à la question 1.3., quels "fondamentaux élargis" doivent être surveillés pour évaluer le risque d'une crise de change?
5. Fin 1992, le franc français subit une première attaque spéculative, à la suite d'autres monnaies européennes comme la livre sterling et la lire italienne. Malgré la courte victoire du "oui" au référendum sur le traité de Maastricht en septembre, et plusieurs mois de hausse des taux d'intérêt de la Banque de France en coopération avec la Bundesbank, l'entrée de la France en récession en 1993 accroît les tensions sur le marché de changes. En août 1993 les gouverneurs des Banques Centrales du SME décident d'élargir les bandes de fluctuation du mécanisme de change à 15%.

Commenter brièvement l'extrait suivant tiré de l'article de Jeanne (1996) sur la crise du franc français en 1992-1993 :

"La crise du Système Monétaire Européen (SME) a étonné par son ampleur, à un moment où beaucoup de commentateurs parlaient du "nouveau SME" comme d'un havre de stabilité. Toutes les monnaies du SME, certaines après une résistance coûteuse et prolongée, durent finalement décrocher de leurs marges étroites de fluctuation. Si certaines d'entre elles, comme la lire, étaient fragilisées par une inflation et des déficits publics excessifs, d'autres monnaies qui ne souffraient pas des mêmes problèmes ne furent pas épargnées par la spéculation.

Le franc français, par exemple, aurait pu être présenté comme la monnaie la plus vertueuse du SME à la veille de la crise. Il est frappant de constater aujourd'hui à quel point les fondamentaux pouvaient être à l'avantage du franc : l'inflation était nettement plus faible qu'en Allemagne; le franc était légèrement sous-évalué par rapport au mark

selon la plupart des mesures, et notre pays connaissait un excédent de la balance courante. Enfin, la France était le seul pays (avec le Luxembourg) à satisfaire les critères de Maastricht à la veille de la crise.”