

CHAPITRE 19

L'économie des infrastructures des marchés financiers

Mis à jour le 17 décembre 2018

Au sens économique, les infrastructures désignent, de manière générale, des institutions servant de support à l'échange de biens, d'informations ou de droits entre agents. Les infrastructures des marchés financiers (ci-après FMI pour *Financial Markets Infrastructures*) servent plus spécifiquement le fonctionnement des marchés financiers ; elles font l'objet d'une attention particulière des législateurs et régulateurs à double titre. D'abord parce que qu'elles sont critiques pour le bon fonctionnement des marchés financiers et donc le financement de l'économie. Ensuite parce qu'elles présentent des caractéristiques qui empêchent les seuls mécanismes de marché de capturer l'ensemble des coûts et bénéfiques associés à leur fonctionnement.

L'analyse économique du fonctionnement des FMI peut largement s'appuyer sur les outils dégagés dans le champ de l'économie des réseaux (particulièrement de communication). Selon Economides¹, les réseaux peuvent se définir formellement comme un ensemble de liens connectant des nœuds et dont les différentes composantes fonctionnent de manière complémentaire². La définition s'applique aux FMI, qui reposent toutes sur l'établissement de liens entre leurs participants (les « nœuds ») *via* leur connexion à un même système technique, qui permet de fournir le service souhaité. Par exemple, un système de paiement met en relation des participants et la complémentarité entre les composantes du réseau vient de la nécessaire connexion des participants au système pour que le service de paiement puisse être fourni. De manière générale, les réseaux peuvent s'appréhender à travers deux prismes différents³ : un prisme technique, où ils renvoient à une interconnexion d'équipements « coopérants » afin de transporter des flux (d'actifs, d'informations) et un prisme économique, où ils sont avant tout le support physique d'une intermédiation économique (vision transactionnelle).

Ce chapitre tire les conséquences de la nature économique des FMI (I) sur

l'organisation et les dynamiques de marché (II) et expose les enjeux de leur tarification (III).

1. La nature économique des FMI

Les FMI ont une nature économique particulière, en ce qu'elles présentent d'importantes externalités, en particulier de réseau, ainsi qu'une structure de coûts permettant des économies d'échelle et d'envergure.

1.1. La présence d'importantes externalités, en particulier de réseau

La notion d'externalité désigne des situations où la production ou la consommation d'un agent affectent l'utilité d'un autre agent, sans que cet effet ne transite par un marché. Lorsque l'effet est positif, l'externalité est dite positive⁴ ; dans le cas contraire, elle est dite négative⁵. Ce phénomène implique que les mécanismes de prix ne permettent pas de capturer de manière satisfaisante l'ensemble des coûts et/ou bénéfiques associés à la production ou à la consommation en question. Cette discordance entre coûts sociaux et coûts privés entraîne *a priori* pour les agents une divergence entre le niveau de consommation ou de production optimal socialement et individuellement et complexifie la mesure des bénéfices sociaux engendrés par les activités concernées.

Les infrastructures des marchés financiers présentent deux grandes catégories d'externalités : des externalités, à la fois négatives et positives, liées à l'impact qu'elles peuvent avoir sur la stabilité financière, et des externalités dites « de réseau » (ou « de club »).

Les externalités positives liées à la stabilité financière renvoient à la contribution que les FMI peuvent avoir en soutien de la stabilité financière du fait de leur systémicité⁶, et justifie l'intervention des pouvoirs publics pour imposer leur usage pour certains

1 Economides N., « The economics of networks », *International journal of industrial organization*, 1996.

2 C'est-à-dire que toutes ces composantes sont nécessaires à la fourniture d'un service.

3 Curien N., *Économie des réseaux*, La Découverte, 2005.

4 Un exemple classique est celui des relations de voisinage entre un apiculteur et un agriculteur, les abeilles du premier pollinisant gratuitement les champs du second du point de vue de ce dernier.

5 Par exemple, des activités industrielles polluantes générant des maladies et des coûts pour les personnes vivant aux alentours.

6 Du fait de cette systémicité, évoquée plus en détails au chapitre 18, les efforts entrepris individuellement par chaque IMF profitent à l'ensemble du système financier.

marchés. On pourra prendre l'exemple du règlement EMIR⁷, qui impose entre autres une obligation de déclaration des transactions sur produits dérivés à des référentiels centraux de données. En effet, la déclaration des transactions sur un marché donné auprès d'un référentiel central de données entraîne une transparence accrue du marché en question, qui permet d'évaluer de manière plus appropriée les risques que son activité peut faire peser sur la stabilité financière et, pour les pouvoirs publics, d'y remédier en temps utile, ce qui bénéficie à l'ensemble des participants. Aussi, afin d'assurer la stabilité financière, ce même règlement EMIR impose aux contreparties une obligation de compensation multilatérale *via* des contreparties centrales de certaines catégories de dérivés de gré-à-gré⁸ (le rôle principal des CCP étant de faire bénéficier leurs participants d'un mécanisme de gestion des risques).

Les externalités négatives liées à la stabilité financière renvoient, inversement, aux potentiels risques systémiques que les infrastructures des marchés peuvent faire peser sur la stabilité financière du fait de leur systémicité⁹. Cela explique qu'elles fassent l'objet d'une régulation exigeante, destinée à faire en sorte que leurs opérateurs tiennent compte des risques qu'elles peuvent faire peser sur l'ensemble de leur écosystème. Cette volonté se traduit par exemple dans les dispositions des PFMI relatives à la gouvernance, qui disposent qu'« *une infrastructure de marché financier devrait être dotée de dispositions relatives à la gouvernance qui soient claires et transparentes, qui favorisent sa sécurité et son efficacité et qui soutiennent la stabilité du système financier dans son ensemble, d'autres considérations d'intérêt public et les objectifs des parties prenantes.* ».

En matière de FMI, d'une part, tous les acteurs de l'écosystème bénéficient de manière égale d'un système financier stable et résilient, sans que le fait pour l'un d'entre eux d'en bénéficier n'empêche les autres de faire de même, d'autre part, il est impossible d'en exclure quelqu'un que ce soit. Dès

lors, les externalités relatives à la stabilité financière produites par les FMI relèvent de la catégorie économique spécifique des biens publics, ce qui soulève le risque de possibles comportements de passager clandestin¹⁰ de la part des acteurs du marché pouvant, au niveau de l'industrie dans son ensemble, conduire à un niveau sous-optimal de gestion des risques financiers. Ce problème est accentué par les risques d'aléa moral¹¹ portés par la certitude des acteurs que les autorités publiques sauveront quoi qu'il arrive les entités de taille systémique (problématique du *too big to fail*, ou *too interconnected to fail*).

Les externalités de réseau signifient que l'utilité d'un participant pris individuellement dépend positivement du nombre (présent et futur) de participants au réseau¹² : la décision d'un utilisateur de rejoindre un réseau correspond ainsi à un autre type d'externalité positive. Cet effet peut être à la fois direct et indirect¹³. Il peut être direct, en ce que le raccordement de nouveaux utilisateurs affecte directement l'utilité des autres, en leur permettant d'être connectés à un nombre plus important de contreparties. Il peut être indirect, en ce que l'accroissement du nombre d'utilisateurs du réseau entraîne une amélioration des caractéristiques de son offre (amélioration de la qualité des services), ou de l'offre de biens et services complémentaires (amélioration de la variété) en rendant le réseau plus attractif pour les fournisseurs de services¹⁴.

En conséquence de ces effets de réseau, les décisions d'adoption par des utilisateurs sont liées au nombre d'utilisateurs, présents et anticipés, de la FMI. Or, par défaut, les utilisateurs ne prennent pas leurs décisions de rejoindre ou non un réseau en tenant compte de l'impact que cette décision peut avoir sur l'utilité des autres utilisateurs : c'est là que réside l'externalité. Il y a donc, pour les FMI, un risque que la taille du réseau à l'équilibre soit inférieure à celle qui serait socialement optimale.

Par exemple, dans le cas des bourses de valeur, les externalités de réseau proviennent,

7 Pour plus de détails, se référer au chapitre 16 (référentiels centraux de données) et au chapitre 11 (CCP).

8 Cf. chapitre 11 (CCP) pour plus de détails.

9 Pour plus de détails, se référer au chapitre 17.

10 Chacun peut avoir intérêt à ce que le système financier soit plus stable, mais faire peser sur les autres les coûts de sécurisation.

11 Possibilité qu'un agent assuré contre un risque se comporte différemment que s'il y était totalement exposé.

12 En termes économiques, on dira que les utilités des différentes parties sont donc interdépendantes, et leur consommation du bien ou service en réseau est complémentaire.

13 Katz M., Shapiro C., « Network Externalities, Competition and Compatibility », *American Economic Review*, 1985; Katz M., Shapiro C., « Systems Competition and Network Effects », *The Journal of Economic Perspectives*, 1994. Se référer également à Curien N., *Économie des réseaux*, La Découverte, 2005.

14 Phénomène connu dans le domaine de la programmation comme le paradigme *hardware-software* : une hausse de la demande pour un *hardware* peut entraîner une hausse du développement de *softwares* compatibles (et vice-versa), ou de leur qualité. Dans le domaine des cartes de crédit, il peut se traduire par le fait que plus le nombre d'utilisateurs de cartes de crédit est important, plus le nombre de commerçants disposant d'un terminal permettant cette utilisation sera important (complémentarité), ce qui en retour accroît l'utilité de leur carte de crédit pour leurs possesseurs (cf. chapitre 4).

du point de vue des investisseurs, de la recherche de la plateforme de négociation offrant la liquidité la plus importante possible pour un type de titres donné. Cela entraîne directement une liquidité encore plus importante pour ce même type de titre et donc un renforcement de l'attrait de la plateforme choisie pour les autres investisseurs¹⁵. Du point de vue des émetteurs, elles proviennent de ce que les besoins en capitaux seront d'autant plus facilement absorbés par un réseau d'investisseurs que la taille de celui-ci est importante¹⁶.

Au stade de la compensation et du règlement, les externalités de réseau proviennent notamment des gains de temps dans le traitement des transactions, tandis que la mise en place de mécanismes de compensation (*netting*) permet de réduire des coûts d'opportunités associés aux exigences de capital réglementaire.

1.2. Une structure de coûts favorisant des économies d'échelle et d'envergure

1.2.1. Les FMI génèrent des économies d'échelle

L'une des raisons premières des FMI consiste à permettre aux acteurs financiers d'externaliser en un point unique des traitements qui étaient auparavant faits de façon décentralisée. Ce faisant, elles génèrent des économies d'échelle.

En effet, la mise en place d'une FMI implique d'importants coûts fixes – au sens où ils sont largement indépendants des volumes de transactions. Ces coûts sont ceux associés aux infrastructures physiques (serveurs, bâtiments les hébergeant), aux développements informatiques (écriture des programmes, tests, correction des erreurs, etc.), et à la mise en place du système de gouvernance (documentation juridique, etc.). Dans une certaine mesure, ils peuvent également comprendre les coûts de maintenance et de surveillance du système ainsi que les coûts de formation du personnel – même si ces postes ne sont pas totalement

indépendants du nombre de transactions. Enfin, les utilisateurs subissent aussi des coûts fixes initiaux de connexion à l'infrastructure. En revanche, les coûts variables liés à l'exploitation des réseaux sont généralement assez faibles.

Le caractère à la fois élevé des coûts fixes et faible des coûts variables implique des coûts moyens par utilisateur généralement décroissants avec leur nombre – donc, des économies d'échelle, ou rendements croissants, également liées à des effets d'apprentissage¹⁷ (maîtrise accrue des processus).

Empiriquement, les études sur le sujet ont ainsi jugé importantes les économies d'échelle dans les systèmes de paiement (de détail ou de gros), avec une part des coûts fixes représentant de 50 à 80 % des coûts totaux¹⁸. La possibilité d'économies d'échelles a aussi été un important argument mis en avant lors de la mise en place de TARGET2¹⁹.

Au stade de la négociation, les bourses de valeur présentent aussi des économies d'échelles liés à la conjonction de coûts fixes élevés de mise en place de la plateforme d'échange et de faibles coûts incrémentaux par transaction une fois celle-ci mise en place. Ainsi, à coûts fixes comparables, la plateforme enregistrant le volume le plus important de transactions pourra bénéficier des coûts de transaction moyens les plus faibles.

Enfin, ces économies d'échelle se vérifient aussi au stade de la compensation et du règlement-livraison. Ainsi, le coût marginal de la compensation est proche de zéro, tandis que l'intervention d'une chambre de compensation permet de réaliser des économies dans la collecte et l'analyse des informations sur ses membres. En effet, la surveillance des risques de crédit et de liquidité implique la mise en place de techniques sophistiquées de suivi de ces risques, qui nécessite des investissements coûteux (informatique, de modélisation, organisationnels...). Il est plus rationnel de les mutualiser au sein d'un seul

15 Pour un exposé de la littérature relative à la liquidité comme externalité, se référer à : Serval T., « Lorsque les réseaux d'information deviendront des bourses », Presses de Sciences Po, 2001.

16 Cf. Curien N., *Économie des réseaux*, La Découverte, 2005.

17 Ces effets désignent des rendements croissants du facteur travail avec la répétition de certaines tâches dans le temps.

18 Khiaonarong T., « Payment systems efficiency, policy approaches, and the role of the central bank », Bank of Finland *Discussion Papers*, 2003.

19 Bolt (W.) et Beijnen (C.), « Size matters : economies of scale in European payment processing », *DNB Working Paper*, 2007 ; Bolt W. et Humphrey, D., « Payment Network Scale Economies, SEPA, and Cash Replacement », *Review of Network Economics*, 2007 ; Bolt W. et Humphrey, D., « Public good issues in TARGET », *ECB Working Paper Series*, 2005.

acteur, la CCP. En outre, la position centrale de celle-ci lui donne une capacité de suivi et de gestion de ces risques supérieure à celle que pourraient déployer chaque acteur pris individuellement. Dans la mesure où les appels de marge sont réalisés sur la base des positions nettes des participants, cette situation permet aux participants de réaliser des économies en liquidités. Par ailleurs, elle permet de réduire les volumes au niveau du règlement-livraison en proportion du nombre de transactions compensées : plus les volumes de transactions sont importants sur un instrument donné, plus la probabilité que certaines de ces transactions soient en sens contraire est importante.

1.2.2. Les FMI génèrent des économies d'envergure

Les FMI peuvent par ailleurs être le support d'économies d'envergure (ou de gamme), qui désignent des situations où il est plus avantageux, en termes de coûts, de produire plusieurs types de biens ou services conjointement, *via* une firme unique, que par des opérateurs séparés. Toutefois, l'exploitation de telles économies peut avoir un impact sur les autres formes d'efficacité, et notamment systémique – se référer au chapitre 17 portant sur les risques dans les FMI.

Dans le cadre des fonctions d'opérateur de systèmes de règlement de montant élevé des banques centrales, il a pu être avancé qu'existaient pour ces dernières des économies d'envergure découlant du double fait i) de gérer pour les établissements de crédit un système de comptes destinés à des fins de réserves et de règlements interbancaires et ii) d'opérer le système permettant des règlements entre ces comptes²⁰. Par ailleurs, ces systèmes permettent aux banques centrales de fournir des services à la fois aux gouvernements²¹ et aux établissements bancaires privés, les faisant ainsi bénéficier de synergies liées à la fourniture jointe de ces services²².

Du côté des plates-formes de négociation de titres, les économies d'envergure peuvent

découler de la possibilité d'effectuer, sur la même plateforme, des transactions sur plusieurs types de titres. Une fois l'infrastructure mise en place, l'extension des services d'échange à un type de titre supplémentaire peut être offerte à un coût incrémental modeste, surtout si un réseau d'acheteurs et de vendeurs actifs est déjà mobilisé. Il est aussi économiquement optimal, pour les utilisateurs, de regrouper leurs activités sur différents types de titres sur le moins de plateformes de transaction possible, notamment afin de réaliser des économies de back-office.

Au stade du post-marché, effectuer la compensation de plusieurs classes d'instruments financiers permet, sous réserve de cadre de gestion des risques qui restent en tout état de cause spécifiques à chaque classe, de mutualiser certaines ressources (les équipes risques ou juridiques par exemple, ou les infrastructures technologiques, qui sont un centre de coûts important). Par ailleurs, la diversification des instruments financiers compensés par la CCP peut permettre, du fait d'une moindre corrélation entre les facteurs de risque des différents instruments, de réduire la variance de l'exposition et en conséquence le montant du collatéral immobilisé en garantie contre ces risques. L'argument vaut aussi pour les participants choisissant de faire compenser leurs transactions auprès d'une seule CCP plutôt que plusieurs – le risque associé au portefeuille de chaque participant étant moins important que la somme des risques par instrument, une CCP unique pourra exiger une marge calculée sur la base d'une exposition nette, moins importante (modèles de *portfolio managing*). Ainsi, en Europe, LCH Ltd a récemment lancé l'offre *Spider*, qui permet de compenser conjointement des futures de taux listés et des swaps de taux OTC en calculant une marge nette sur l'ensemble des positions. Cela permet à LCH Ltd de se développer sur le marché listé en tirant parti de sa position forte sur les swaps de taux ; inversement, Eurex a lancé une initiative visant à pénétrer le marché des swaps à partir de sa position sur les futures.

20 Millard S., Haldane A., Saporta V., *The Future of Payment Systems*, Routledge, 2008 ; Green, E.J. et Todd R.M., « Thoughts on the Fed's Role in the Payments System », *Federal Reserve Bank of Minneapolis Quarterly Review*, 2001 ; Green, E.J., « The Role of the Central Bank in Payment Systems », 2005.

21 Représentant le rôle historique des banques centrales d'intermédiaire entre les gouvernements et leurs prêteurs.

22 Bolt, W. et Humphrey, D., « Public good issues in TARGET », *ECB Working Paper Series*, 2005.

Enfin, des économies d'envergure peuvent aussi provenir de la fourniture par un seul groupe de services de négociation de titres, de compensation et de règlement, qui sont fortement complémentaires²³. Cela permet un traitement des transactions de bout en bout (*straight-through-processing*) au sein d'un même groupe, entraînant

des réductions de coûts de communication entre les différentes activités (pouvant favoriser la mise en œuvre de standards communs pour la transmission des données entre les différentes étapes), et améliorant le processus de diffusion des innovations tout au long de la chaîne en réduisant les besoins de coordination. La mise en place

23 Au sens où la demande de chaque service n'est pas indépendante du prix des autres services.

Encadré n° 1 : Économies d'échelle et d'envergure attendues de la mise en place de T2S

La réalisation du projet T2S fournit un bon exemple des économies de coûts que peut permettre de réaliser une consolidation d'activités dans le domaine du règlement-livraison (pour plus de détails sur le fonctionnement de T2S, se référer au chapitre 14).

Les gains attendus de T2S en termes d'exploitation des économies d'échelle et d'envergure et d'externalités de réseau ont été détaillés dans une étude réalisée en 2007 par la BCE¹. La fourniture de cette plateforme commune devait permettre des économies d'échelles *via* une consolidation des coûts d'investissement ainsi que des coûts d'exploitation liés à la maintenance des plateformes (propres, avant la migration, à chacun des 24 CSD que comptent les pays de la zone euro).

Du côté des utilisateurs des CSD, la mise en place de T2S devait permettre de faire le choix de centraliser leurs comptes titres chez un seul dépositaire central de leur choix (à condition que ceux-ci aient mis en place les liens nécessaires) ou leur compte espèces au niveau d'une seule banque centrale, entraînant une baisse des coûts liés à la dispersion de leurs comptes. Ces acteurs devaient donc notamment bénéficier de possibilités de rationalisation de leurs activités transfrontières et d'économies de *back-office*, *via* en particulier la possibilité de s'appuyer sur un seul point d'entrée à plusieurs marchés, d'harmoniser leurs procédures en interne, d'utiliser leurs garanties de manière commune pour leurs activités cash et titres (création d'un pool unique de collatéral réduisant le coût de la mobilisation transfrontalière des garanties), et de s'appuyer sur un nombre moins élevé d'intermédiaires (CSD, mais aussi pour la gestion d'actifs). L'usage d'une plateforme unique et de protocoles de communication standardisés devait aussi offrir aux émetteurs la possibilité d'atteindre plus d'investisseurs, et ainsi d'augmenter la demande s'adressant à eux, tandis qu'il permettrait aux investisseurs de réduire le coût de leur gestion d'un portefeuille de titres répartis à l'international – et donc d'augmenter le rendement de ce portefeuille.

En ce qu'il ouvrait les marchés domestiques, le projet devait permettre d'accroître la concurrence entre CSD sur le marché unique, et entraîner une consolidation (des systèmes) en Europe – ce qui pourrait dégager des économies d'échelle et d'envergure supplémentaires. Toujours à moyen terme, les bénéfices liés à l'utilisation d'une plateforme technique unique se concrétiseraient aussi au moment de la faire évoluer – voire, le cas échéant, de la refondre –, quand il n'y aurait alors qu'un seul projet à conduire au lieu de 23.

Enfin, les travaux actuels en vue de la consolidation entre TARGET2 et T2S visent à exploiter les synergies entre les deux plateformes² (cas d'économies d'envergure), dans les domaines de l'utilisation de ressources et architecture IT, de possibilités de réutilisation de technologies de communications existantes, d'organisation des fonctions support et opérationnelles. L'exploitation de ces synergies pourrait aller jusqu'à une fusion future éventuelle des deux plateformes (cf. chapitre 7, section 6.3).

1 BCE, 2007, *TARGET2-Securities – Economic Feasibility*.

2 Détaillées par exemple dans ce document accessible sur <https://www.ecb.europa.eu/pub/pdf/other/t2seconomicfeasibility0703en.pdf?8e36385d37d399eaf9a3615292b80c08>

d'un « silo » vertical de FMI tout au long de la chaîne de traitement des titres, allant de la négociation au règlement-livraison en passant par la compensation, illustre l'existence de ce type d'économies. Elle pose toutefois des défis en termes de gestion des risques, notamment systémiques, ainsi que présenté dans le chapitre 17 (section 2.2.1), et de concurrence.

2. Les conséquences de la nature économique des FMI sur l'organisation et les dynamiques de marché

L'existence d'économies d'échelles du côté de l'offre et d'externalités de réseau du côté de la demande peut limiter les possibilités de concurrence dans les industries de réseaux, et favoriser la concentration des opérateurs, d'une part du fait d'une tendance à la concentration horizontale et verticale, et d'autre part du fait que ces deux caractéristiques s'alimentent l'une l'autre et créent des effets de rétroaction. Une telle tendance à la concentration soulève des enjeux essentiels de stabilité financière. En effet, la concentration de l'activité de compensation ou de règlement-livraison, par exemple, sur un très petit

nombre d'acteurs, voire sur un acteur unique par catégorie d'activité, aboutit à constituer des acteurs systémiques dont la défaillance aurait des conséquences déstabilisatrices extrêmes et ne peut donc pas être envisagée (« *too big to fail* »). La constitution de tels acteurs monopolistiques crée un aléa moral important, dans la mesure où elle impose aux autorités d'intervenir pour éviter ces défaillances (cf. chapitre 17 dédié aux risques dans les FMI). Ainsi, un enjeu de politique industrielle et de concurrence pour les pouvoirs publics dans le domaine des FMI est d'encourager l'organisation de marché la plus efficiente possible, en évitant la création d'acteurs monopolistiques à caractère trop systémique.

2.1. Une tendance à la concentration horizontale et verticale

En conséquence des externalités de réseau et des économies d'échelle et d'envergure engendrées par leurs activités, les FMI présentent des caractéristiques de monopole naturel, au sens où il peut être optimal qu'une seule entité prenne en charge la satisfaction de l'ensemble de la demande.

Les caractéristiques économiques fondamentales des FMI constituent dès lors

Encadré n° 2 : T2S et la délimitation des « pures » activités de fourniture d'infrastructures

Ainsi que décrit au chapitre 14, T2S n'est pas considéré comme un système de règlement titres ou un CSD, mais comme une plateforme technique de règlement-livraison permettant aux CSD de développer leurs services sur une base identique. La mise en place de T2S peut être rapprochée de choix effectués dans d'autres industries de réseau (télécommunications, transports ferroviaires, fourniture d'électricité etc.) historiquement organisées en monopoles, visant à cantonner les monopoles historiques à des activités de « pure » fourniture des infrastructures de réseau, où la fourniture par un seul acteur est la forme la plus efficace d'organisation du marché. Au contraire, la partie du réseau correspondant aux activités de fourniture de services commerciaux sur la base de cette infrastructure, aurait vocation à être ouverte à la concurrence, du fait notamment de coûts fixes moins importants.

Dans le cas de T2S, le raisonnement économique consistant à séparer la couche infrastructure de la couche service est semblable, mais le mouvement historique est inverse, avec une prise en charge par le secteur public d'une activité de pure fourniture de l'infrastructure (la plateforme technique) pour des motifs d'efficacité, tandis que les services commerciaux fournis à la clientèle des CSD (la couche « service ») restent à la main des CSD.

une barrière à l'entrée pour les nouveaux arrivants, dans la mesure où ceux-ci doivent être en mesure de répliquer les coûts fixes importants des opérateurs en place, où l'essentiel de ces coûts fixes est généralement très spécifique et ainsi irrécupérable (*sunk costs*) en cas de sortie du marché, et où un acteur déjà en place bénéficiera toujours de coûts moyens par utilisateur plus faibles. Elles exacerbent l'importance, pour un nouvel opérateur, de pouvoir atteindre rapidement une masse critique de clients qui lui permettent d'exploiter les économies d'échelle et les effets de réseau, l'incertitude sur les capacités d'un entrant potentiel à atteindre une telle masse critique pouvant elle aussi être un frein à l'entrée sur le marché. Ces caractéristiques suggèrent l'efficacité productive de la tendance à la concentration dans le domaine des FMI, à la fois horizontalement et verticalement.

En pratique, le domaine des paiements de montant élevé fonctionne souvent sous la forme d'un duopole au niveau domestique, comme on l'a vu aux chapitres 6 à 8 : TARGET2 et EURO1 en Europe, FEDwire et CHIPS aux États-Unis... De même, le marché des paiements de détail (cf. chapitre 10) est souvent dominé, au niveau domestique, par un seul acteur (exemple de CORE(FR), opéré par STET, en France). Pareillement, les marchés financiers domestiques européens sont généralement organisés autour d'un seul CSD national et une seule CCP – la mise en production récente de T2S pourrait toutefois entraîner des modifications de ce paysage.

Le paysage du post-marché pour les titres est logiquement plus intégré aux États-Unis qu'en Europe, les États-Unis étant à cet égard un seul marché domestique. Le règlement-livraison et la compensation s'y organisent principalement autour de DTCC²⁴ (*Depository Trust and Clearing Corporation*), qui assure aussi les fonctions de dépositaire central. En Europe, le cloisonnement des systèmes domestiques (pour des raisons historiques, techniques, institutionnelles, juridiques ou fiscales) entraînait des coûts relativement élevés pour les transactions transfrontalières²⁵ (du

fait de la complexité et du nombre d'intermédiaires potentiellement impliqués dans ces opérations), et c'est notamment ce qui a conduit à réaliser T2S.

Cette fragmentation empêchait que soient pleinement exploitées les externalités positives de réseaux et les économies d'échelles possibles. Un mouvement de consolidation est cependant observé en Europe depuis plusieurs années, tant horizontalement (concentration des entités fournissant des services similaires au même point de la chaîne de valeur du post-marché) que verticalement (entités fournissant des services différents mais intégrés le long de la chaîne de valeur du post-marché).

Des mouvements de concentration horizontale ont été notamment observés *via* les rapprochements entre plateformes de négociation : création du groupe Euronext en 2000 par rapprochement des bourses d'Amsterdam, Bruxelles et Paris ; puis rapprochement avec la Bourse de Lisbonne et la Bourse de Porto en 2002, et en 2017 avec la Bourse de Dublin. Ils ont été aussi observés au stade de la compensation (regroupement au sein d'un même groupe de LCH et de Clearnet en 2003), ou au stade de la conservation et du règlement-livraison avec en particulier la stratégie de rapprochement menée par Euroclear avec plusieurs CSD nationaux et la mise en place de la plateforme ESES (*Euroclear Settlement of Euronext-zone Securities*) à partir de 2009 pour toutes les transactions effectuées sur les marchés Euronext²⁶. Le trio formé par la plateforme Euronext au niveau de la négociation, la contrepartie centrale LCH SA (contrôlée depuis 2013 par le London Stock Exchange) au niveau de la compensation, et le dépositaire central Euroclear illustre, pour chaque couche de la chaîne de valeur du traitement des titres, le modèle d'intégration horizontale. Dans ces exemples, des FMI centrées sur leur métier (négociation, compensation, règlement/livraison) cherchent à servir plusieurs marchés géographiques. Le modèle horizontal peut ainsi bénéficier aux participants et autres utilisateurs en ce qu'il permet de dégager de fortes synergies commerciales : ouverture et

24 Créé en 1999 pour regrouper et intégrer les opérations du *Depository Trust Company* (DTC) et de la *National Securities Clearing Corporation* (NSCC).

25 Le différentiel de coût entre une transaction domestique et une transaction transfrontière était estimé en 2011 comme variant de 1 à 10 – « *Settling Without Borders*, » European Central Bank, 2011.

26 À l'exception du Portugal

positionnement sur différents segments du métier cœur de l'infrastructure ; possibilités de diversification ; innovation ; plus grande indépendance dès lors que la totalité de l'activité de l'infrastructure ne dépend pas d'un seul et même marché.

L'intégration verticale « en silos » s'observe également sur d'autres marchés. Il s'agit d'une organisation où toute la chaîne des opérations sur titres (plateforme de négociation, chambre de compensation et dépositaire central de titres gérant un système de règlement-livraison) est traitée de bout en bout par des infrastructures et des systèmes propriété d'un même groupe capitalistique. Cela correspond, par exemple, aux modèles allemand (avec le Deutsche Börse Group qui contrôle la plateforme Eurex, la CCP Eurex Clearing et le CSD Clearstream Banking Frankfurt) ou italien (avec le Borsa Italiana Group, où la société Borsa Italiana – elle-même contrôlée par le LSE – contrôle la plateforme de négociation MTS, la chambre de compensation Cassa di Compensazione e Garanzia et le CSD national Monte Titoli). Les gains peuvent provenir d'un alignement des intérêts stratégiques pouvant se traduire par une innovation accrue (baisse des coûts de coordination, internalisation des externalités liées à la R&D), ou une intégration des outils et processus permettant une meilleure coordination et diffusion d'effets d'apprentissage. Ce modèle pourrait tendre toutefois à encourager des relations quasi exclusives entre ces acteurs, ce qui soulèverait des risques concurrentiels, ce contre quoi la réglementation s'efforce de lutter (cf. infra). En outre, une telle intégration verticale de différentes infrastructures qui ne présentent pas le même profil de risque et doivent répondre à des exigences réglementaires différentes peut représenter une difficulté lorsqu'il s'agit de préserver la nécessaire indépendance de chacune des infrastructures.

2.2. Des conséquences qui pèsent sur les dynamiques d'innovation

Du fait des externalités de réseau et des économies d'échelle et d'envergure

décrits plus haut, l'offre et la demande interagissent généralement dans les industries de réseau selon un principe dit de rétroaction (*feedback*)²⁷ : plus un service en réseau est diffusé, plus il suscitera une demande importante, en raison de rendements croissants d'adoption, et renforcera ainsi sa diffusion au détriment de services concurrents (une hausse de l'offre entraîne une hausse de la demande, qui entraîne une hausse de l'offre etc.), et inversement.

En premier lieu, ces effets viennent renforcer la tendance au monopole naturel pour conduire à une polarisation du marché sur un nombre limité d'opérateurs de réseau et de fournisseurs de services : les firmes disposant d'une assise importante renforcent leur position. Ainsi, ils sont un amplificateur d'évolution et créent une logique de « *winner takes all* », où même un faible avantage peut faire basculer le marché en faveur de celui qui le possède, tandis qu'il sera très difficile pour les firmes dominées d'exister (on parle aussi de « *market tipping* »).

Une conséquence classique de tels effets est que la technologie dominante ne sera pas nécessairement la plus efficace : une technologie efficace peut se retrouver éliminée pour cause de « *time to market* » inadéquat, et un standard peut se retrouver dominant du fait d'une sélection précoce (« *first mover advantage* ») et non parce qu'il est le meilleur du fait de sa qualité.

En deuxième lieu, les effets de rétroaction ont pour conséquence que les dynamiques d'innovation peuvent souffrir d'un excès d'inertie lié à des défauts de coordination²⁸. Celui-ci découle à la fois de potentiels défauts d'incitations à adopter des nouvelles technologies du côté des utilisateurs, et de défauts d'incitations à investir dans les nouvelles technologies du côté de l'offre.

Les inerties d'adoption s'expliquent d'abord par le fait que les technologies de réseau se caractérisent généralement par des coûts de migration importants. Cela peut avoir

27 Cf. note n°5 ; voir aussi Foray D. « Innovation et concurrence dans les industries de réseau », *Revue française de gestion* ; Katz M., Shapiro C., « Technology adoption in the presence of network externalities », *Journal of Political Economy*, 1986 ; Shapiro C. et Varian R. *Information Rules*, Harvard Business School Press, 1999.

28 Katz M. et Shapiro, « Systems Competition and Network Effects », *Journal of Economic Perspectives*, 1994 ; Farrell J. et Saloner G., « Standardization, compatibility and innovation », *Rand Journal of Economics*, 1985 ; Milne A., « What is in it for us? Network effects and bank payment innovation », *Journal of Banking and Finance*, 2005.

pour effet de rendre captifs les consommateurs, et constituer une barrière à l'entrée supplémentaire pour des fournisseurs de technologies alternatives. Les coûts de migration peuvent se définir²⁹ comme la somme du coût supporté par le consommateur (apprentissage, investissements spécifiques) pour changer de technologie, et du coût supporté par le nouveau fournisseur afin de placer le consommateur dans la même situation qu'avec le fournisseur précédent. Ainsi, dans le cas des technologies informatiques sous-jacentes aux FMI, le participant a généralement dû réaliser des investissements durables dans des biens complémentaires³⁰ spécifiques à la technologie considérée. Le changement de technologie principale oblige ainsi à changer ou du moins adapter les biens complémentaires, qui peuvent en outre avoir une durabilité différente (c'est-à-dire il n'y a pas de bon moment pour repartir à zéro), ce qui renforce l'effet d'inertie. Un exemple de coûts de migration dans le monde des FMI peut s'illustrer par les investissements qu'ont dû réaliser les CSD pour pouvoir utiliser T2S, incluant notamment des modifications d'architecture dans les systèmes qu'ils opèrent (par exemple, modification de l'architecture informatique de la plateforme ESES pour Euroclear). Ces coûts de migration viennent en appui du fait que les inerties technologiques peuvent parfois être économiquement efficaces, même si la technologie envisagée est supérieure. C'est notamment le cas si le coût de la migration est plus élevé que les avantages procurés par la nouvelle technologie.

Les inerties d'adoption s'expliquent ensuite par le fait que, même si tous les participants peuvent avoir collectivement intérêt à adopter une nouvelle technologie, le risque de se retrouver isolé dans un nouveau réseau peut être tel que chacun décide de reporter sa décision de changement, jusqu'à ce que la base d'utilisateurs du nouveau réseau soit suffisamment importante. Si tous les utilisateurs suivent le même raisonnement, le changement est bloqué. Cette impasse peut aussi se voir comme une forme de dilemme du prisonnier, dont le

résultat serait plus favorable si les acteurs coopéraient en s'entendant sur le choix du changement technologique, mais où en l'absence de coopération, et en situation d'incertitude sur les choix que feront les autres, il est optimal pour chaque participant de ne pas changer de technologie (sous-optimalité de l'équilibre de Nash³¹).

Ainsi, lorsque le marché demeure livré à lui-même, ses dynamiques n'engendrent pas toujours des résultats économiques efficaces, ce qui fonde l'intervention des pouvoirs publics sur les marchés, en particulier, dans le domaine des FMI, avec le rôle de catalyseur que peuvent jouer les banques centrales.

2.3. Les réponses des autorités publiques aux problématiques concurrentielles

2.3.1. FMI et politique de concurrence : l'enjeu de l'accès aux FMI

Outre le lien entre la taille des FMI et les risques systémiques qu'elles peuvent faire peser sur le système financier, traité spécifiquement au chapitre 17 sur les risques dans les FMI, la tendance naturelle à la concentration du secteur pose aussi problème du fait des pratiques de prix que les opérateurs peuvent mettre en place du fait de leur pouvoir de marché au détriment des consommateurs et des échanges économiques³². Ainsi, la politique européenne de la concurrence³³ s'attache à combattre la possibilité que le pouvoir de marché des acteurs puisse constituer l'instrument d'exercice de pratiques restrictives de concurrence, visant à créer ou renforcer les barrières à l'entrée.

Dans leur rôle de catalyseur, les instances européennes ont ainsi fait le choix de construire le marché unique sur un modèle fondé sur une mise en concurrence des différents fournisseurs de services financiers, que ce soit au niveau des établissements eux-mêmes ou des FMI, notamment sur le plan transfrontalier. Cette volonté s'est traduite notamment par des efforts entrepris pour supprimer les barrières identifiées dans

29 Shapiro et Varian (1999)

30 Au sens où ils sont nécessaires à l'accès au service fourni par l'infrastructure.

31 Désigne, en théorie des jeux, une situation dans laquelle chaque joueur adopte la meilleure stratégie possible compte-tenu de la stratégie adoptée par l'autre joueur.

32 Cf. le corpus économique théorique sur la charge morte du monopole.

33 Désigne l'ensemble des dispositifs, notamment juridiques, visant à organiser et contrôler les marchés de sorte à favoriser le maintien d'un fonctionnement suffisamment concurrentiel pour que l'efficacité économique soit maximale.

les rapports Giovannini, la directive MIF³⁴ qui a consacré cette approche, ou encore l'élaboration du code de conduite du post-trading menée sous l'égide de la Commission européenne. Il est à souligner que cette approche n'entre pas en contradiction avec la reconnaissance d'une efficacité productive de la concentration dans certains cas, dans la mesure où l'objectif poursuivi est de mettre en place un marché financier européen unique : la dé-fragmentation des différents marchés géographiques permet de mettre en concurrence différents acteurs (en déplaçant les frontières géographiques des marchés) et des mouvements de concentration peuvent, à plus long terme, être une conséquence de cette mise en concurrence (cf. *supra*, encadré 2 relatif à T2S).

Les conditions d'accès aux FMI font l'objet d'une attention toute particulière des législateurs et des régulateurs. En effet, l'accès aux FMI peut constituer une condition de participation à certains marchés. Elles constituent ainsi une infrastructure essentielle³⁵, dont les opérateurs peuvent se trouver en mesure de jouer comme d'un goulet d'étranglement³⁶ (*bottleneck facility*) pour les marchés qui ont un besoin essentiel d'accéder à l'infrastructure. Le caractère d'infrastructure essentielle de certaines FMI s'appuie aussi parfois sur des exigences réglementaires : ainsi, la loi Dodd-Frank aux États-Unis (titre VII) ou le règlement EMIR en Europe imposent aux contreparties de certaines opérations de produits dérivés une obligation de compensation. Il est donc important que les acteurs de marché assujettis à cette obligation puissent avoir accès aux infrastructures qui proposent ces services. La théorie des infrastructures essentielles fonde les obligations d'accès imposées dans le cadre de réglementations sectorielles, qui trouvent leur application dans le secteur des FMI au principe 18 des PFMI, qui dispose qu'« une infrastructure de marché financier devrait avoir des critères de participation objectifs, fondés sur une analyse des risques et rendus publics, et qui permettent un accès équitable et ouvert », afin de favoriser la concurrence entre les participants au marché.

Par ailleurs, le pouvoir de marché des opérateurs peut permettre la mise en œuvre de stratégie d'effet de levier³⁷, voire de stratégies d'éviction pure et simple³⁸ sur des marchés amont, aval ou connexes à celui qu'elle dessert – ces risques étant particulièrement importants pour le cas d'infrastructures verticalement intégrées. Par exemple, les CCP sont généralement possédés par des opérateurs aussi présents sur le marché aval de la négociation (exemple du groupe Deutsche Börse), et le choix de la CCP est souvent à la main de l'opérateur de la plateforme de négociation, qui peut donc décider de recourir à une unique CCP, évinçant en conséquence les autres.

Le principe d'un accès équitable et ouvert en ce qu'il concerne les relations entre infrastructures de marché s'incarne dans la notion d'*open access*, déclinée dans divers pans de la réglementation financière post-crise (EMIR³⁹, CSDR, MiFID II/MiFIR).

Le principe d'*open access* permet par exemple à une plateforme de négociation de faire compenser ses transactions par la chambre de compensation de son choix, et à une chambre de compensation d'avoir accès aux flux de transaction de toute plateforme de négociation. En effet, un opérateur sur le marché aval se voyant refuser l'accès à une chambre de compensation, ou imposer des conditions d'accès moins favorables que certains membres possédant l'infrastructure, pourra voir ses coûts se détériorer, voire se faire évincer du marché aval. Inversement, en l'absence de réglementation, le risque existe qu'une plateforme de négociation puisse refuser à une CCP la compensation des transactions négociées sur sa plateforme, tarissant ainsi les flux de transaction arrivant à la CCP. C'est pour pallier ce type de risque et favoriser l'ouverture des silos verticaux que les réglementations sectorielles prévoient ce principe d'*open access*. En sens inverse, ce principe permet à un CSD d'avoir accès aux transactions compensées par toute chambre de compensation et à une chambre de compensation d'avoir accès aux flux de transactions de toute plateforme de négociation.

34 Directive 2004/39/CE du Parlement européen et du Conseil du 21 avril 2004 concernant les marchés d'instruments financiers.

35 À l'origine une création des tribunaux américains appliquant les dispositions du Sherman Act prohibant la monopolisation, cette théorie a été reprise en droit communautaire et français dans le cadre de la lutte contre les abus de position dominante ainsi que de l'ouverture à la concurrence de marchés jusqu'alors gérés par des monopoles publics, puis a été étendue. Cf. Cour suprême des États-Unis, *United States v. Terminal Railroad Association*, 224 US 383, 1912; Rapport annuel 2005 de la Cour de Cassation citant un arrêt du 12 juillet 2005 de la chambre commerciale (n°04-12388); Chang F.B., « Financial Market Bottlenecks and the Openness Mandate », *University of Cincinnati*, 2015; affaire *Sealink*, Commission Européenne, 94/19/CE, 1993, où la Commission utilise l'expression d'« installation essentielle »; arrêts de la CJCE n°C-241/91, n°C-07/97, n°C-418-01.

36 Cf. Curien N., *Économie des réseaux*, La Découverte, 2005.

37 Capacité pour une entreprise d'augmenter ses ventes sur un marché – le « marché lié » – en exploitant sa position dominante sur un marché adjacent, rendant plus difficile l'accès à ces derniers marchés.

38 L'efficacité de telles stratégies a été débattue dans la littérature économique, et notamment contestée par l'École de Chicago – pour un exposé de la manière dont le contrôle d'une infrastructure essentielle peut permettre d'accroître les coûts des concurrents, se référer par exemple à Patrick Rey et Jean Tirole, *Handbook of Industrial Organization*.

39 Articles 7 et 8.

Encadré n° 3 : Quelques cas évoquant la problématique de l'accès aux FMI dans la jurisprudence

Les FMI ont pu être amenées à réviser leur politique d'accès pour se conformer aux exigences légales. Ce n'est par exemple qu'après l'édition de telles règles sur l'accès que la plateforme SwapClear de LCH Clearnet a levé la clause exigeant de ses membres de maintenir un encours de portefeuille de swap de taux d'intérêt d'un montant de 1 000 milliards de dollars, clause suspectée de réserver le marché aux importants vendeurs institutionnels ¹.

En Europe, la Commission a pu juger dans une décision du 2 juin 2004² (confirmée le 9 septembre 2009 par un arrêt du TPUE³) qu'en refusant de fournir des services transfrontaliers de compensation (au sens d'établissement des obligations contractuelles réciproques de l'acheteur et du vendeur) et de règlement pour des actions nominatives émises en vertu du droit allemand et en adoptant des pratiques de prix discriminatoires, au détriment d'Euroclear Bank, Clearstream Banking AG et sa société mère Clearstream International SA avaient abusé de leur position dominante sur les marchés concernés. La Commission a rappelé dans cette affaire que, si le droit de la concurrence, y compris pour le marché du règlement-livraison, reconnaissait la « *liberté des sociétés de choisir leurs partenaires commerciaux* », les sociétés en position dominante avaient une « responsabilité particulière ». Elle a considéré qu'en l'espèce, Clearstream était le seul dépositaire de titres allemands conservés en dépôt collectif et qu'une nouvelle entrée sur le marché n'était pas une hypothèse réaliste, ce qui en faisait un partenaire commercial incontournable, qu'Euroclear Bank n'était pas en mesure de produire les services qu'elle demandait et que le comportement de Clearstream avait eu pour effet d'entraver sa capacité à fournir à des clients du marché unique des services de compensation et de règlement sur une base transfrontalière de 1997 à 2002.

1 Felix B. Chang, « Financial Market Bottlenecks and the « Openess » Mandate », *University of Cincinnati*, 2015. Le *Core Principle C (v)* de la loi Dodd-Frank prohibe explicitement que les organismes compensant les dérivés imposent à leurs membres de maintenir un encours de position swap d'une valeur ou d'un volume particulier.

2 Communiqué de presse : http://europa.eu/rapid/press-release_IP-04-705_fr.htm.

3 <http://eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=CELEX:62004TJ0301&from=FR>

2.3.2. Stratégies d'interconnexion et de compatibilité pour les FMI

L'interconnexion, ou interopérabilité, désigne le raccordement des différents réseaux entre eux, et doit permettre aux utilisateurs d'un réseau de communiquer avec les utilisateurs d'un autre réseau, ou accéder à des services fournis par différents opérateurs⁴⁰. C'est cette compatibilité qui permet la réalisation d'externalités positives provenant des complémentarités entre les éléments du réseau⁴¹. Cette possibilité de « dialogue » entre systèmes peut être imposée par la réglementation, ou relever du choix des fournisseurs de service.

Dans le domaine des FMI, le manque d'interopérabilité correspond par exemple en Europe à la situation pré-migration à

T2S d'une fragmentation du marché du règlement-livraison entre les différents États-membres. Dans le cadre des travaux liés à la mise en place de l'espace SEPA dans le domaine des moyens de paiement, l'Euro-système a eu l'occasion de préciser que la notion d'interopérabilité, telle qu'entendue dans le règlement SEPA, impliquait qu'il devait suffire aux prestataires de services de paiement de participer à un seul système de paiement de détail pour pouvoir atteindre leurs contreparties dans l'ensemble de l'espace SEPA, même si celles-ci participent à un autre système, ce qui implique la mise en place de liens d'interopérabilité.

L'interopérabilité repose notamment sur la mise en œuvre d'une compatibilité de standards techniques (flux de messageries par exemple, ou compatibilité entre *hardware*,

40 On trouve par exemple une définition de la notion à l'article L.32 du Code des postes et des télécommunications.

41 Cf. Economides N., « The economics of networks », *International journal of industrial organization*, 1996 : « Links on a network are potentially complementary, but it is compatibility that makes complementarity actual. »

software et *operating systems*), opérationnels, et juridiques. L'utilisation de standards ouverts, transparents et non propriétaires facilitant l'interopérabilité stimule la concurrence sur le marché, en ouvrant la possibilité de transactions entre les utilisateurs des différentes infrastructures sans qu'ils aient besoin de participer à l'ensemble d'entre elles. Elle facilite l'activité simultanée de plusieurs infrastructures, permettant d'éviter le phénomène de concentration excessive décrit plus haut, ce qui peut entraîner une amélioration de la qualité des services proposés aux consommateurs et réduire les risques qu'un acteur en situation de position dominante abuse de cette dernière. Des standards communs peuvent permettre d'éviter des problèmes de coordination dans les choix technologies des firmes, et d'ancrer les anticipations des utilisateurs sur les chances d'adoption d'une technologie donnée. Cette réduction de l'incertitude des consommateurs permet elle-même de réduire leur risque d'être captifs de technologies finissant par ne pas être admises par la majorité du marché, et donc les comportements attentistes de non-adoption. Le fait que l'industrie se mette d'accord sur un standard entraîne généralement une augmentation du rythme d'adoption, et rend les conditions de concurrence plus symétriques.

La standardisation⁴² comporte toutefois des risques potentiels de verrouillage sur de « mauvaises » technologies, notamment du fait des mécanismes de rétroaction décrits plus haut, et peuvent accroître les coûts de transition vers des technologies nouvelles. Par ailleurs, ses coûts peuvent varier selon les acteurs de marché concernés – selon leur taille, ou selon que les acteurs sont déjà en place et ont déjà pu engager des coûts irrécupérables d'élaboration et de mise en conformité avec un jeu de standards qui peut être plus ou moins éloigné de celui visé dans le processus. Enfin, l'imposition d'une compatibilité totale peut présenter le risque d'encourager des comportements de passager clandestin dans la R&D et l'innovation, entraînant *in fine* une baisse des incitations à investir pour améliorer le fonctionnement du réseau.

Pour un opérateur pris individuellement⁴³, le choix d'utiliser des standards communs à ceux de ses concurrents peut être poussé par la perspective de pouvoir accéder à leurs clients, et donc augmenter grâce à l'*effet réseau* décrit plus haut l'utilité que ces clients pourraient retirer de l'accès à son réseau et/ou à ses services (donc notamment leur propension à payer), et d'atteindre plus rapidement une masse critique de consommateurs. *A priori*, l'intérêt d'une stratégie d'interconnexion sera donc d'autant moins élevé pour une firme que sa base installée de clients est importante, ce qui place les acteurs installés et les nouveaux entrants potentiels dans une situation asymétrique⁴⁴. À l'inverse, un refus d'interconnexion peut être motivé économiquement pour un opérateur par une volonté d'offrir des services différenciés par rapport à ses concurrents (*effet concurrence* dû au fait que la compatibilité des services offerts par deux fournisseurs les rend plus substituables), ou parfois même de les évincer du marché. Une telle stratégie pourrait par exemple être choisie par un nouvel entrant disposant d'une technologie supérieure aux acteurs en place, et souhaitant l'imposer en bénéficiant d'un effet d'entraînement (spirale de rétroactions positives). En ce sens, la mise en place de standards peut constituer une forme renouvelée de monopole permettant aux acteurs remportant les « guerres de standards » de bénéficier d'une rente économique. Pour cette raison, la législation relative aux FMI, notamment MIFID et MIFIR, ainsi qu'EMIR, imposent aux FMI de s'interconnecter en cas de demande d'une autre FMI en ce sens. Le seul motif de refus d'interconnexion accepté par ces règlements repose sur des considérations de risques. Une FMI qui refuse une interconnexion ne peut le faire que si elle établit qu'une telle interconnexion susciterait des risques pour sa sécurité et l'efficacité de son fonctionnement.

Se pose alors la question de savoir si la compatibilité des standards est un équilibre naturel de marché – c'est-à-dire si les participants sont incités à l'assurer par eux-mêmes afin de maximiser les externalités pouvant

42 Se référer à Milne, A., « Standards setting and competition in securities settlement », *Bank of Finland Research Discussion Papers*, 2005. L'auteur souligne que des standards propriétaires pourraient inciter plus fortement à l'innovation, être promus de manière plus agressive par les entreprises qui les ont développés, et présenter donc plus de chances d'être acceptés largement par l'industrie, et que, par ailleurs, des entreprises utilisant des standards incompatibles pourraient se livrer une concurrence plus intense sur les prix.

43 Cf. Curien N., *Économie des réseaux*, La Découverte, 2005, ainsi que Kemppainen K., Salo S., « Promoting Integration of European Retail Payment Systems : Role of competition, Cooperation and Regulation », *Bank of Finland*, 2006.

44 Cf. Curien N., *Économie des réseaux*, La Découverte, 2005.

être retirées d'une extension du réseau, ou non. Faute d'incitations suffisantes, ce n'est pas forcément le cas.

Le remède à cette défaillance de marché est l'impulsion de standards communs pour une pleine interopérabilité de l'industrie qui peut provenir d'une action des pouvoirs publics ou d'organismes internationaux, de comités émanant de l'industrie elle-même, ou être sponsorisée par une ou plusieurs firmes. À titre d'exemple, dans le domaine du règlement-livraison de titres, le rapport Giovannini de 2003 a conféré à SWIFT ainsi qu'au *Securities Market Practice Group*⁴⁵ un rôle de facilitateur d'un projet d'harmonisation des standards de messagerie et d'interopérabilité à l'échelle de l'industrie, afin de supprimer l'une des barrières identifiées à la réalisation du marché unique. Dans le domaine des paiements, l'EACHA (*European Automated Clearing Houses Association*) a ainsi travaillé à l'élaboration d'un cadre d'interopérabilité entre les différents systèmes de paiement de détails traitant des transactions de paiements instantanés.

Du côté des pouvoirs publics, il ressort du mandat de l'Eurosystème d'encourager le dialogue avec et entre les différentes parties prenantes ainsi que la poursuite de solutions favorisant l'interopérabilité. Ce rôle de catalyseur s'est illustré par exemple avec le projet SEPA et, plus récemment, avec le projet TIPS (*Target instant payments settlement*) dans le cadre de la stratégie « Vision 2020 » de l'Eurosystème (cf. chapitre 7, section 6). La mise en place de T2S a aussi été l'occasion pour l'Eurosystème de jouer un rôle de catalyseur pour harmoniser le fonctionnement des différents marchés nationaux, en donnant l'impulsion à un alignement des cadres législatifs, opérationnels et techniques (ex : harmonisation horaire du cycle de règlement, formats de messages et de données, etc.), qui était nécessaire au bon fonctionnement de T2S. Ce travail a été effectué en s'appuyant sur une collaboration étroite et une implication forte de l'industrie, les enceintes de concertation ayant validé tous les choix de standards et exercé une pression par les pairs pour faire progresser ces efforts d'harmonisation.

3. L'enjeu de la tarification pour les opérateurs des FMI

Les conséquences du fonctionnement du marché sur les contraintes et stratégies de tarification possibles appellent une intervention publique, que ce soit en termes de régulation des acteurs eux-mêmes ou directement comme opérateur. Dans les faits, les principales obligations en matière de tarification imposés par les surveillants des FMI consistent en une exigence de transparence de leur tarification.

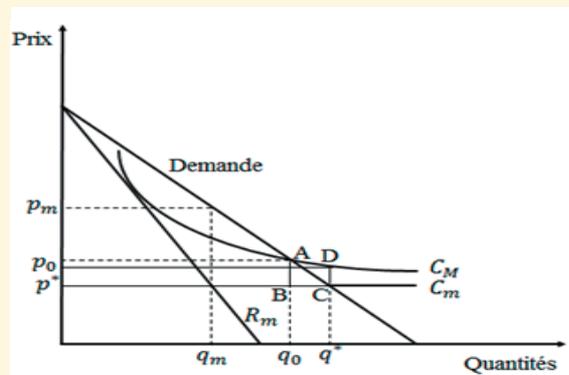
Le graphique 1⁴⁶ illustre les problématiques de tarification pour un monopole monoproduit. Plusieurs niveaux de prix appellent des commentaires. La tarification au coût marginal (prix fixé à p^*) représente en théorie une solution « de premier ordre », efficiente au sens de Pareto d'un point de vue allocatif, car maximisant le surplus collectif permettant d'épuiser tous les gains de l'échange⁴⁷. En pratique, une telle tarification a peu de chances de s'observer dans le domaine des FMI, pas uniquement du fait du pouvoir de marché des opérateurs décrit plus haut, mais aussi parce que dans une industrie qui se caractérise par d'importants coûts fixes et de faibles coûts variables (caractéristiques de monopole naturel), elle ne

45 Groupe de travail de praticiens visant à favoriser l'harmonisation des pratiques de marché pour l'industrie des titres. Il s'appuie sur les *National Market Practice Groups* présents dans plus de 30 pays.

46 Source : *Economie de la réglementation*, de F. Lévêque, La Découverte, 2004.

47 C'est la définition d'un optimum de Pareto.

G1 : Tarification d'un monopole monoproduit



Note : R_m est le revenu marginal du monopole, et la fonction de coût est supposée de la forme $C(q) = F + C_m q$, avec F les coûts fixes et C_m le coût marginal, supposé constant et toujours inférieur au coût moyen C_M , qui décroît au fur et à mesure que les coûts fixes sont amortis (cela découle de la structure de coûts décrite en partie 1). p^* , p_0 et p_m sont respectivement les prix fixés au coût marginal, au coût moyen et au coût de monopole (celui qui conduit à une maximisation de ses profits), avec les quantités correspondantes q^* , q_0 et q_m .

permet pas à l'opérateur qui la met en place de couvrir ses coûts (puisque ses coûts moyens sont par définition supérieurs à ses coûts marginaux). Graphiquement, les recettes, égales à p^*q^* , sont inférieures aux coûts C_{Mq^*} , et les pertes sont égales au rectangle p^*p_0DC , qui correspond aux coûts fixes.

Une tarification à un niveau inférieur au coût moyen peut constituer un moyen temporaire d'inciter les consommateurs à rejoindre le réseau à court terme afin de lui permettre de réaliser pleinement les externalités de réseau positives. Plus spécifiquement, une politique de prix possible peut consister à fixer les prix au coût unitaire attendu de traitement de transaction le plus bas une fois pleinement réalisées les économies d'échelles⁴⁸. Dans le cas où la banque centrale est opérateur de l'infrastructure, cette tarification correspond à une subvention publique, ce qui, lorsque le service fourni est de nature contestable, va décourager les opérateurs privés potentiels d'entrer sur le marché (car ils ne peuvent pas couvrir leurs coûts). Dans le cas où l'infrastructure est opérée par un acteur privé, une telle politique peut, comme dans le cas d'un opérateur public, correspondre aussi à un objectif d'atteinte de la masse critique à partir de laquelle les externalités de réseau sont positives. Elle peut toutefois être la conséquence d'une stratégie de prédation visant à empêcher un concurrent entrant ou potentiel d'être rentable, afin d'exploiter dans un second temps un pouvoir de marché et lever des profits sur-concurrentiels⁴⁹.

Un enjeu – complexe – pour les opérateurs de FMI, est donc, dans ce cadre, la détermination de la structure de prix optimale pour permettre la couverture des coûts.

En premier lieu, en présence d'effets de réseau, une tarification uniforme au coût moyen (qui correspondrait graphiquement au prix fixé à p_0), qui *a priori* permet à l'opérateur de couvrir les coûts sans faire de profits, peut décourager certains utilisateurs de rejoindre le réseau (ceux dont la disposition

à payer est située entre le coût moyen et le coût marginal), ce qui a pour conséquence que les externalités positives de réseau ne sont pas entièrement exploitées – ce qui constitue aussi un problème du point de vue de l'efficacité allocative. La perte sèche de bien-être est représentée graphiquement par le triangle ABC.

En deuxième lieu, en théorie, un monopole naturel choisira de lui-même de vendre la quantité q_m au prix de monopole p_m qui maximise son profit et correspond à l'égalité de son revenu marginal avec son coût marginal. Ce niveau de prix n'est pas Pareto-optimal, dans la mesure où le monopole pourrait servir, en réalisant des profits, tous les consommateurs dont la disposition à payer se situe entre p_m et p_0 .

Enfin, la tarification de l'usage de l'infrastructure peut aussi être fixée de manière non uniforme : prix différent pour un même service⁵⁰, ou en fonction des volumes consommés. Le choix de tarification dominant dans le domaine des FMI est celui d'une politique de prix non linéaire, faisant varier le prix par transaction supportée par l'infrastructure en fonction du nombre de transactions initiées par un participant.

Une première variante de cette tarification consiste en la formulation d'une tarification à deux paliers (*two-part pricing*) non-linéaire de forme $t(q)=F+cq$, comprenant une partie forfaitaire fixe F (dont l'objectif est de couvrir les coûts fixes et reflète les économies d'échelle qui en découlent), et une partie variable, proportionnelle au nombre de transactions (dont l'objectif est de couvrir les coûts variables moyens). C'est le type de tarification choisie par la Banque d'Angleterre dans sa fourniture de services au système de paiement de montant élevé privé, CHAPS⁵¹, avec pour objectif une couverture des coûts à long terme, sans perte ni profit, et sans subventions croisées entre les différentes lignes de services.

Une seconde variante consiste en une double tarification à deux paliers, selon le

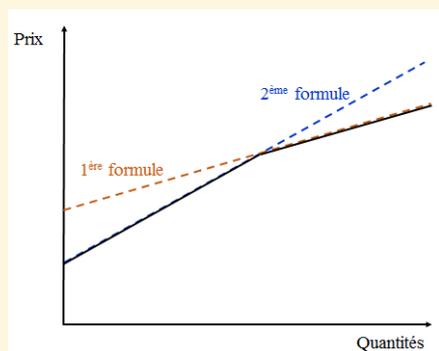
48 Bolt W. et Humphrey D., « Public Good Aspects of TARGET : Natural Monopoly, Scale Economies, and Cost Allocation », *DNB Working Paper*, 2005.

49 Lorsqu'elle est pratiquée par un opérateur en position dominante, une telle pratique est susceptible d'enfreindre les règles de concurrence, mais les autorités de concurrence ne considèrent établie la preuve d'une telle infraction qu'en appliquant des critères particulièrement difficiles à vérifier.

50 L'exemple le plus évident consisté à facturer une majoration par rapport au coût marginal de production d'autant plus importante que l'élasticité-prix de la demande est faible. Une telle discrimination par les prix nécessite toutefois en pratique d'évaluer l'élasticité-prix de la demande. Elle peut conduire, dans un cadre multiproduit, à mettre en œuvre une politique de subventions croisées où les services dont la demande est la plus élastique sont subventionnés par ceux où elle l'est moins. Une telle politique contrevient aux objectifs de certaines banques centrales de couverture des coûts par ligne de services (cas de la Fed par exemple), et (ce qui est lié) peut donner la possibilité à certains acteurs de pratiquer une politique d'écémage (*cream skimming*) en ne servant que les segments de marchés où l'élasticité-prix de la demande est faible et recouvrant sur ceux-ci leurs coûts tout en bénéficiant d'un avantage comparatif par rapport à un opérateur mettant en œuvre une politique de subvention croisée, ce qui finit par rendre celle-ci inefficace.

51 Prix annuel d'adhésion pour CHAPS de 15000 £ et pour le DvP de 15000 £, *per item fee* fixé à 15,5 p pour CHAPS et 1,90 £ pour le DvP.

G2 : Tarification d'un monopole monoproduit



Note : Les ordonnées à l'origine correspondent au coût fixe de chaque formule et les pentes des droites aux coûts marginaux constants.

même principe, où chaque participant choisirait le mode de tarification lui procurant le niveau d'utilité le plus élevé.

Ces deux formules peuvent être illustrées par le graphique 2. L'on voit sur la figure que pour des quantités faibles, il est plus intéressant d'adopter la formule 1 (première partie de la ligne pleine), et pour des quantités plus importantes, la formule 2 (seconde partie de la ligne pleine) : l'opérateur qui souhaite discriminer tient compte du fait

que chaque type de participant (dans un tel cas, les participants effectuant un nombre important de transactions *via* le système, et les autres) choisira la formule qui l'avantage le plus, et fixera les deux barèmes de tarifications en fonction.

C'est le mode de tarification des services de base choisi pour TARGET2, où les participants ont le choix entre deux options. La première, dans laquelle le tarif est unique, composé d'une redevance mensuelle de 150 euros et d'un prix à la transaction de 0,80 euro, est conçue à l'intention des établissements à faible volume d'échange. La seconde, dans laquelle la redevance mensuelle est de 1 875 euros et qui présente des tarifs dégressifs par seuil de volume de transactions, est conçue à l'intention des grands établissements.

Dans ces deux variantes, le prix moyen décroît avec les volumes consommés, ce qui a les mêmes effets qu'une politique de remise de volume, et limite les subventions croisées entre les différents types de participants. Une telle structure est généralement considérée comme efficiente⁵², et encourage l'usage de l'infrastructure chez les participants ayant des volumes importants.

52 Pour son application aux systèmes de paiement, se référer à Holthausen C. et Rochet J.-C., « Efficient Pricing of Large Value Interbank Payment Systems », *Ohio State University Press*, 2006.

Encadré n° 4 : La couverture des coûts par les banques centrales pour la fourniture de services de RTGS

La majorité des banques centrales opérant un RTGS appliquent une politique de couverture partielle ou totale des coûts. Certaines mettent en avant l'opportunité de subventionner les RTGS compte tenu des bénéfices qu'ils peuvent apporter à l'ensemble de la communauté en termes de stabilité économique (le Japon, par exemple). Les banques centrales sont d'autant plus sensibles à l'opportunité de la mise en place d'un objectif de couverture des coûts que les services qu'elles fournissent sont de nature contestable, dans le sens où ils pourraient dans une certaine mesure être fournis par le secteur privé. Un objectif de couverture seulement partielle des coûts peut s'expliquer au contraire par une absence de contestabilité du service fourni par la banque centrale (par exemple, parce que les paiements en monnaie de banque centrale ne peuvent être effectués que dans des RTGS opérés par la banque centrale), ou par une crainte qu'une politique de couverture totale ne conduise à une tarification décourageant l'usage du RTGS pour des systèmes de paiement plus risqués.

Cet objectif est dans certains cas une obligation légale. Aux États-Unis, le *Monetary Control Act* de 1980¹ impose à la Réserve fédérale un objectif général de long terme de couverture de ses coûts

.../...

1 https://www.federalreserve.gov/paymentsystems/pfs_pricingpol.htm.

directs et indirects. L'objectif poursuivi par le Congrès était à la fois de stimuler la concurrence (notion de *fair competition*), afin que les services soient fournis au plus bas coût possible pour la société, et de faire en sorte que ces services soient fournis à un niveau adéquat (rôle du prix comme signal dans un objectif d'efficience allocative). Ce principe est interprété de manière contraignante par la Fed, qui se donne un objectif de couverture totale des coûts (de production, d'investissement, d'opérations – maintenance et exploitation –, et un facteur d'ajustement par rapport aux coûts qui auraient été supportés par le secteur privé²) par ligne de service.

Dans l'Eurosystème, l'article 2 des statuts du SEBC et de la BCE dispose que le SEBC « agit conformément au principe d'une économie de marché ouverte où la concurrence est libre, en favorisant une allocation efficace des ressources », ce qui suggère que la tarification des services qu'il fournit doit tendre le plus souvent possible vers un objectif de couverture des coûts lorsqu'une subvention est susceptible d'entraver la concurrence du secteur privé. L'objectif affiché pour TARGET2 est celui d'une couverture des coûts importante³ afin d'éviter une concurrence déloyale avec les systèmes de paiement privés.

La couverture de ses coûts pour TARGET2 n'a pas pour autant vocation à être totale du fait de la prise en compte d'un *public good factor* correspondant aux externalités positives générées par son usage, notamment en termes de réduction de risque systémique. Ce *public good factor* s'interprète et se mesure comme la subvention qui serait nécessaire pour que le secteur privé « internalise » les coûts liés à la prise en compte des risques systémiques dans la manière dont il opère le système de paiement⁴. Au terme de plusieurs études économiques menées préalablement au lancement de TARGET2, ce public good factor a été estimé à 20% des coûts totaux. En effet, outre les services de règlement des paiements de montant élevé pour les transactions s'y effectuant directement, TARGET2 fournit aussi :

- des services de tenue de comptes de réserve pour le règlement final en monnaie de banque centrale des positions nettes des transactions effectuées dans d'autres FMI (au cœur des activités de banque centrale et ne pouvant par nature être fournis que par elle) ;
- des facilités de crédit intrajournalier, constituant une extension des pratiques de crédit *overnight*, activité de banque centrale par nature non contestable) ;
- des services de transfert de fond pour les comptes de réserve des banques commerciales.

Compte tenu de l'ensemble de ces fonctions additionnelles au seul règlement, imposer un objectif de couverture totale des coûts pourrait conduire à pratiquer une tarification dissuasive par rapport aux systèmes privés qui ne jouent pas les mêmes rôles et n'ont pas les mêmes contraintes, ce qui pourrait ne pas être socialement optimal. En pratique, dans l'évaluation du *public good factor*, le coût de production des prestations assimilables à des prestations commerciales s'établit toutefois plus facilement que pour les dimensions relatives aux externalités liées au renforcement de la sécurité globale de l'économie.

Pour T2S en revanche, l'objectif est celui d'une couverture totale des coûts, sans que l'opération de la plateforme ne poursuive un but lucratif. Cet objectif a notamment été choisi pour des raisons d'ordre concurrentiel, dans la mesure où il s'agissait d'« internaliser » des activités (exploitation de la plateforme technique fournissant le service de règlement livraison) auparavant du ressort des CSD, donc *prima facie* en partie contestable⁵.

2 Défini comme « an allowance for the taxes that would have been paid and the return on capital that would have been provided had the Federal Reserve's priced services been furnished by a private-sector firm ».

3 https://www.ecb.europa.eu/paym/target/target2/profuse/fees/shared/pdf/TARGET2_Pricing_Guide_v6.0.pdf, p4.

4 Pour plus de détails, voir : Holthausen C. et Rochet J.-C., « Incorporating a « public good factor » into pricing of large-value payment systems », European Central Bank Working Paper, July 2005.

5 En partie car quel que soit le modèle, la banque centrale demeure le seul acteur à pouvoir fournir un service de paiement de gros en monnaie banque centrale.

