

TROIS ESSAIS EN MICROÉCONOMIE DE LA FINANCE ET DE LA BANQUE

Jean-Edouard Colliard

Résumé

Cette thèse inclut trois articles théoriques dont le but est de prendre en compte des frictions institutionnelles ou liées à la régulation pour expliquer certains dysfonctionnements du marché, d'épisodes temporaire d'illiquidité à des crises financières. Ce programme de recherches, associé à ce qu'on appelle parfois la "finance institutionnelle" (*institutional finance*), est développé ici dans trois applications: la régulation des banques, l'agrégation de l'information sur les marchés financiers, et la dynamique des marchés à ordres à cours limité. Le thème commun de ces trois articles est de montrer comment les marchés financiers sont affectés par l'interaction de frictions et de comportements stratégiques et, éventuellement, d'étudier de possibles remèdes à apporter.

Rational blinders: is it possible to regulate banks using their internal risk models? ("Des oeillères rationnelles : est-il possible d'utiliser leurs propres modèles de risque internes pour réguler les banques ?") étudie le secteur bancaire et l'impact du cadre réglementaire actuel, qui utilise l'information interne des banques contenue dans leurs modèles de risque pour calculer leurs ratios minimum de capital. Ceci donne aux banques des incitations à utiliser des modèles optimistes afin de contourner la régulation. Je montre que simplement rendre la régulation plus contraignante peut être contre-productif, menant à une utilisation plus large de modèles optimistes et à un risque plus élevé, j'étudie donc des mécanismes pour révéler l'information privée des banques sur le modèle correct.

Catching falling knives: speculating on market overreaction “Rattraper un couteau qui tombe, ou spéculer sur la surréaction des marchés”) présente un modèle dynamique d’un marché financier pendant un épisode de crise, le but étant d’étudier les effets informationnels de ventes forcées (*fire sales*) et de la présence sur le marché d’acteurs possédant de l’information privée sur l’importance quantitative de ces ventes. Je montre que ces acteurs agissent comme une assurance contre des krachs non liés aux fondamentaux de l’économie : ils rendent moins probable que des ventes forcées déclenchent un krach, mais au prix de davantage de sélection adverse et d’une convergence plus lente des prix dans le long terme.

Trading fees and efficiency in limit order markets (“Frais de transaction et efficacité des marchés à ordres à cours limité”), co-écrit avec Thierry Foucault (HEC Paris) et à paraître dans la *Review of Financial Studies*, étudie l’impact de la concurrence croissante entre plateformes d’échange sur la fourniture de liquidité par les investisseurs sur les marchés à ordres à cours limité. Nous montrons en particulier qu’une concurrence accrue mène à des frais de transaction plus faibles qui, en raison de la réaction stratégique des investisseurs, peuvent de manière contre-intuitive avoir un impact négatif sur la liquidité du marché et sur les gains des investisseurs.

Examinés au travers du prisme de la microéconomie de la finance, les marchés financiers sont sujets à des inefficacités dues à des contraintes institutionnelles ou technologiques, au comportement stratégique des acteurs et à des problèmes informationnels. La théorie des marchés financiers a accordé de plus en plus d’attention à ces problèmes au cours des dernières années, le comportement d’un marché théorique ne présentant aucune friction étant désormais bien compris et suffisamment étudié pour fournir un modèle de référence utile. Markus Brunnermeier va jusqu’à faire une distinction entre la “finance de l’équation d’Euler” (*Euler equation finance*), basée sur le comportement d’un agent représentatif sur des marchés

parfaits, et la “finance institutionnelle” (*institutional finance*), où les frictions institutionnelles sont au centre du processus de formation des prix¹. La microéconomie est un outil tout indiqué pour identifier à la fois la fonction de différentes frictions et les inefficacités dont elles sont la source, ainsi que pour donner des pistes de solutions. Inversement, la richesse des marchés financiers en problèmes stimulants et en anomalies empiriques offre un domaine d’applications unique pour de nombreuses branches de la microéconomie, qui pose de nouvelles questions et requiert parfois de nouveaux outils et de nouveaux développements théoriques. Une des ambitions de ce travail est de procurer un modeste exemple de comment la microéconomie et la finance peuvent réciproquement s’enrichir.

La suite de cette introduction a pour but de donner un résumé non technique des différents chapitres de la thèse et de les placer dans un contexte plus large, les chapitres eux-mêmes ne rappelant que les travaux les plus proches des trois articles présentés.

1 Préambule sur la microéconomie et la finance

Avant de se tourner vers les différents sujets étudiés dans cette thèse, il est utile, notamment pour des non spécialistes amenés à lire cette introduction, de consacrer quelques pages aux relations entre microéconomie et finance. Pour de nombreux économistes travaillant en dehors de l’économie de la finance, la finance est d’une manière générale vue comme une branche des mathématiques appliquées, et se caractérise par l’utilisation intensive du calcul stochastique et de lettres grecques, rendant peu compréhensible l’intérêt que peut présenter ce champ pour un microéconomiste. Ce jugement, que j’ai pu observer comme étant relativement

¹Lors d’une présentation de Brunnermeier et Pedersen (2009) et Tobias et Brunnermeier (2008) à la Banque de la Réserve fédérale d’Atlanta, voir http://www.frbatlanta.org/news/CONFERENCE/09fmc/brunnermeier_slides.pdf.

répandu, est d'autant plus paradoxal que les modèles classiques de la finance sont des applications de la théorie microéconomique, qu'il s'agisse par exemple de la frontière d'efficience, du *Capital Asset Pricing Model*, ou de l'*Arbitrage Pricing Theory*. Des auteurs comme Harry Markowitz, Merton Miller, Stephen Ross ou William Sharpe ont obtenu leur thèse en économie². Même les branches de la finance perçues comme plus mathématiques et moins économiques doivent leurs fondations à la théorie économique : l'utilisation de martingales dans l'évaluation des produits dérivés repose sur l'hypothèse que les marchés sont efficients et que seules des innovations aléatoires dans les valeurs des actifs sont à prendre en compte, tout le champ intitulé *consumption-based asset pricing* est basé sur l'étude "d'économies d'arbres de Lucas" (*Lucas tree economies*, Lucas (1978)).

La "finance néoclassique" (Ross (2004)) trouve ses racines dans la théorie microéconomique, et a en même temps ouvert la voie à des modèles d'évaluation des actifs d'une sophistication mathématique croissante, ce sont ces modèles que nombre d'économistes ont en tête lorsque ils entendent parler de travaux en "finance". La citation suivante, extraite de Ross (2002), donne une définition très complète du *benchmark* théorique qu'offre la finance néoclassique :

Les deux piliers de la finance néoclassique sont l'hypothèse des marchés efficients et la théorie de l'évaluation des actifs qui lui est très liée, en particulier l'absence d'arbitrage et l'évaluation risque-neutre. Les marchés efficients sont l'ensemble des hypothèses exprimant l'intuition selon laquelle l'information des acteurs du marché est intégrée dans les prix des actifs. [...] Dans la mesure où les prix sont la conséquence des actions des agents, ils reflètent l'information détenue par ces agents et, pour cette raison, il n'y a rien de plus à gagner en analysant les prix, ou même l'information publique également accessible aux autres investisseurs. [...] Des stratégies d'investissement reposant sur l'information déjà contenue dans les

²Notons néanmoins que, selon la légende, Milton Friedman se serait opposé à l'obtention de sa thèse par Harry Markowitz, au motif que ses travaux ne relèveraient pas de l'économie.

*prix, qui inclut toute l'information publiquement disponible, sont vouées à l'échec.*³

La microéconomie a été un outil important dans la construction de ces “deux piliers”, mais les microéconomistes n’ont jamais déserté le champ de la finance après avoir établi le cadre théorique de base. Au contraire, les microéconomistes ont très tôt essayé de comprendre des caractéristiques des marchés financiers qui ne trouvent pas d’explication dans un monde walrasien sans frictions. Un certain délai fut néanmoins nécessaire. Exception faite de travaux antérieurs comme [Markowitz \(1952\)](#), les grandes étapes de la “finance sans frictions” datent des années 1960 et du début des années 1970: [Treyner \(1962\)](#), [Sharpe \(1964\)](#) et [Lintner \(1965\)](#) pour le CAPM, [Fama \(1970\)](#) qui fait la synthèse de travaux des années 1960 sur l’hypothèse d’efficacité des marchés, [Black et Scholes \(1973\)](#) qui lancent le champ de l’évaluation des produits dérivés et des mathématiques financières, [Ross \(1976\)](#) qui introduit l’*arbitrage pricing theory*. En revanche l’étude des frictions était difficile avant le développement de l’économie de l’information et l’adoption des outils de la théorie des jeux.

Une première génération de modèles à la fin des années 1970 et dans les années 1980 tente d’abord d’expliquer l’existence de certaines caractéristiques institutionnelles des marchés financiers. La première est la présence d’intermédiaires financiers, redondants dans un monde sans frictions. [Leland et Pyle \(1977\)](#) montrent que, en raison de l’information privée détenue par un entrepreneur sur la valeur de son investissement, financer un projet a un coût informationnel, ouvrant la voie à une fonction informationnelle de l’existence d’intermédiaires financiers si des économies d’échelle permettent de faire baisser de tels coûts ([Ramakrishnan et Thakor \(1984\)](#), [Diamond \(1984\)](#)). Les banques en particulier ont attiré beaucoup d’attention en raison de leurs nombreuses spécificités. [Diamond et Dybvig \(1983\)](#) montrent qu’elles ont un rôle dans la provision de liquidité, tandis

³Ma traduction.

que Diamond (1984) puis Calomiris et Kahn (1991) et Holmstrom et Tirole (1997) voient les banques comme des “surveillants délégués” (*delegated monitors*) dont la fonction est de discipliner les emprunteurs pour le compte des apporteurs de capitaux, trop dispersés pour jouer ce rôle. Le lecteur est renvoyé au chapitre 2 de Freixas et Rochet (1997), auquel cette courte présentation emprunte, pour plus de détails sur ces travaux. A peu près au même moment, le concept d’efficience des marchés financiers est attaqué par Grossman et Stiglitz (1980), qui montrent que si les prix des actifs incorporent toute l’information pertinente, alors acquérir cette information est inutile puisque il suffit de regarder les prix. Mais dans ce cas personne n’acquiert d’information, et les prix ne peuvent donc rien incorporer. Ce “paradoxe” a joué un rôle important dans l’introduction de frictions dans l’étude des marchés financiers, notamment le coût d’acquérir de l’information et la présence d’investisseurs avec des problèmes de liquidité venant bruyé le marché. Stoll (1978) et Glosten et Milgrom (1985) donnent deux explications différentes à l’existence de fourchettes de prix (*spreads*) sur les marchés financiers, une autre déviation importante par rapport au modèle de référence : pour le premier le *spread* compense les intermédiaires pour les positions déséquilibrées qu’ils accumulent, et les risques additionnels qu’elles leur font courir. Pour les seconds, la friction principale est la sélection adverse : les faiseurs de marché courent le risque de faire des pertes en faisant une transaction avec un vendeur ou un acheteur plus informés, et cotent un *spread* positif pour compenser ce risque de perte. Le champ de la microstructure des marchés financiers qui a suivi, défini par O’Hara (1995) comme “l’étude du processus et des résultats de l’échange d’actifs selon un ensemble bien spécifié de règles”⁴, est une façon d’appréhender les marchés financiers particulièrement attirante pour un microéconomiste. Pour citer le même auteur, “alors que la plus grande partie de la science économique fait abstraction du détail

⁴Ma traduction.

du processus d'échanges, la théorie de la microstructure étudie spécifiquement comment des mécanismes d'échanges spécifiques affectent le processus de formation des prix". Enfin, un dernier champ de la finance dans lequel la microéconomie a joué un grand rôle au cours de la même période est bien sûr la finance d'entreprise, où des articles comme Myers et Majluf (1984) ou Jensen (1986) ont grandement enrichi l'analyse du choix entre différents modes de financement par les entreprises.

Une fois développés les outils pour étudier les problèmes d'information et les interactions stratégiques sur les marchés financiers, les travaux sur les imperfections et les échecs de ces marchés ont fleuri : les contraintes sur les ventes à découvert (Diamond et Verrecchia (1987), Hong et Stein (1987)), la spéculation déstabilisatrice (Stein (1987), DeLong, Shleifer, Summers et Waldmann (1990)), les contraintes de levier (Chowdhry et Nanda (1998)), les attaques spéculatives (Corsetti, Dasgupta, Morris et Shin (2004)), le *predatory trading* (Brunnermeier et Pedersen (2005)), les logiques moutonnières (Avery et Zemsky (1998)), le risque systémique (Freixas, Parigi et Rochet (2000)) ne forment qu'une liste très courte et arbitraire des problèmes étudiés. La microéconomie des marchés financiers a suivi une évolution comparable à la microéconomie en général (qui étudie le plus souvent des marchés pour des biens et services réels) : élaboration d'une théorie générale et cohérente d'un monde sans frictions (théorie de l'équilibre général / CAPM, APT et hypothèse d'efficacité des marchés), contestation de la pertinence de cette théorie (économie de l'information et économie industrielle / un grand nombre des articles ci-dessus, ainsi que la finance comportementale, suivant un angle d'attaque différent), étude des échecs de marché et des frictions ainsi que de possibles solutions, le modèle sans frictions initial devenant un point de référence théorique utile. Il est facile de voir ce dernier point en finance à chaque fois qu'un article essaie d'expliquer une "anomalie", c'est-à-dire une déviation empirique par rapport au modèle de référence des marchés efficients.

Il est intéressant de constater que, alors que beaucoup de travaux ont développé une vision des marchés financiers bien différente de la première génération de modèles, les travaux plus appliqués en finance quantitative ont été lents à prendre en compte ce changement de paradigme, ce qui probablement explique en partie pourquoi de nombreux non spécialistes réduisent encore la recherche en “finance” à l’hypothèse des marchés efficients. Là encore cette hystérèse a un parallèle en économie : malgré l’évolution de la microéconomie, les modèles quantitatifs en macroéconomie ont longtemps reposé sur un cadre walrasien à agent représentatif avec peu ou pas de frictions, ce qui mène des commentateurs peu informés à croire que les économistes n’ont guère fait de progrès conceptuels depuis les années 1960. Dans les deux cas il est facile de comprendre la cause d’un tel délai : il est bien plus facile de construire un modèle quantitatif précis apportant des prédictions chiffrées sur la base d’un modèle sans frictions que sur la base d’une vision plus sophistiquée des marchés, délivrant de nombreux résultats qualitatifs intéressants, mais aussi de nombreux obstacles (relations non linéaires, discontinuités, multiplicité d’équilibres, complexité...). Dans les deux cas également, le processus d’intégration dans des modèles plus quantitatifs des frictions étudiées dans les articles théoriques a déjà commencé, mais est toujours en cours. En finance beaucoup de progrès a par exemple déjà été fait dans l’intégration des problèmes de liquidité ([Acharya et Pedersen \(2005\)](#), [Brunnermeier et Pedersen \(2009\)](#)), bien que beaucoup reste sans doute à faire.

Le but de ce préambule était de mieux identifier le champ auquel ce travail appartient : un large corpus de travaux théoriques en finance est à la finance sans frictions ce que l’économie industrielle et l’économie de l’information sont à la théorie standard de l’équilibre général. Il est relativement aisé d’identifier quels articles appartiennent à cet ensemble, qui souffre d’une trop grande abondance de noms, sans qu’aucun soit ni très précis ni très bon. Microéconomie de la finance comprend

également l'étude de marchés sans frictions et est à ce titre trop large. "Microstructure des marchés" est parfois employé dans ce sens, à tort car trop étroit. "Finance dysfonctionnelle"⁵ est un terme coloré mais imprécis dans la mesure où une friction ou une anomalie ne sont pas nécessairement dysfonctionnelles (au contraire, les frictions institutionnelles ont en général une fonction, au sens où elles résolvent un problème plus profond). "Finance institutionnelle", comme elle est opposée par Markus Brunnermeier à la "finance de l'équation d'Euler", est probablement le meilleur terme, avec le plus simple "finance avec frictions". Les deux s'appliquent parfaitement aux trois articles présentés ci-dessous.

⁵Du nom d'une école d'été à la *Duisenberg School of Finance* d'Amsterdam.

2 Rational blinders: is it possible to regulate banks using their internal risk models?

L'ambition de cet article est de bâtir un cadre théorique pour comprendre comment des institutions financières peuvent avoir des incitations rationnelles à utiliser des modèles prédictifs biaisés. Un exemple fascinant de ce phénomène est donné par Goldman Sachs, qui a publiquement expliqué que de lourdes pertes survenues en 2007 étaient dues à des fluctuations du marché qui représentaient, d'après leur propre modèle, des "événements de 25 sigmas" (*25 sigmas events*), soit des événements supposés survenir une fois tous les 10^{135} ans. Or de telles fluctuations survinrent trois jours de suite ! La sous-estimation du risque que cela représente semble trop importante pour que la simple malchance soit une explication admissible. Le risque de modèle, c'est-à-dire des erreurs de bonne foi faites dans la gestion et la prévision du risque, n'est pas une explication très satisfaisante non plus : il faut alors expliquer pourquoi tant d'institutions et de régulateurs semblent s'être reposés sur des modèles relativement optimistes quant aux risques extrêmes, alors que des modèles alternatifs utilisant des distributions à queues épaisses ou la théorie des valeurs extrêmes étaient disponibles depuis un certain temps. Cet article développe l'hypothèse que le problème repose sur de mauvaises incitations, hypothèse qui à ma connaissance a été peu explorée jusqu'à présent, si ce n'est de manière informelle, par exemple dans [Danielsson \(2008\)](#). J'étudie l'exemple particulièrement parlant de la régulation du risque de crédit dans le cadre des accords de Bâle : les banques peuvent développer des modèles statistiques de risque et produire des estimations de risque que le régulateur utilise pour calculer les fonds propres réglementaires. Le but d'un tel mécanisme est de rendre la régulation plus sensible au vrai risque de la banque, mais un effet pervers est qu'une banque adoptant un modèle pessimiste quant aux risques extrêmes voit ses fonds propres réglementaires

augmenter et son profit diminuer.

Je construis dans cet article un modèle d'équilibre d'intermédiation financière, où des modèles trop optimistes peuvent être choisis de manière stratégique pour contourner la régulation. Un premier résultat surprenant est que durcir les exigences de capital réglementaire (par exemple en passant de Bâle II à Bâle III) pour compenser le "risque de modèle" peut mener à une plus large adoption de modèles optimistes, et accroître le nombre de défauts dans le secteur bancaire. En d'autres termes, la prise en compte de la possibilité que les modèles de risque soit choisis stratégiquement conduit à s'interroger sur l'évolution récente de la régulation. Dans une deuxième partie, cet article étudie le problème de sélection adverse d'un régulateur qui cherche à apprendre d'une banque le vrai modèle de risque et à l'utiliser pour définir les fonds propres réglementaires. Il s'agit en particulier de comprendre l'arbitrage entre utiliser plus intensément l'information apportée par les modèles internes pour la régulation, et le coût de s'assurer que les modèles utilisés ne soient pas biaisés.

Le cadre est le suivant : un régulateur choisit une formule reliant les estimations de risque fournies par les banques à un niveau de fonds propres réglementaires. Les banques apprennent quel est le vrai modèle de risque de crédit, mais peuvent choisir d'en utiliser un autre pour produire l'estimation envoyée au régulateur. Chaque banque reçoit ensuite un niveau de fonds propres à maintenir en fonction de l'estimation envoyée, et choisit combien elle emprunte à des investisseurs et combien elle prête à des emprunteurs finaux. Enfin, une proportion aléatoire de ces emprunteurs vont être incapables de rembourser leur prêt, suivant une distribution de probabilités qui correspond au vrai modèle de risque. Les banques dont les pertes dépassent le capital font alors défaut à leur tour. La proposition 1 résoud l'équilibre de ce jeu, soit des prix et une distribution de modèles tels qu'aucune banque ne modifie le volume des prêts qu'elle accorde ou des montants qu'elle

emprunte, ni ne change de modèle de risque. Le corollaire 1 montre qu'à l'équilibre le nombre de banques adoptant un modèle trop optimiste et la proportion espérée de banques faisant défaut s'accroissent tous deux lorsque la demande de prêts par les emprunteurs finaux augmente. On s'attend donc à ce que l'optimisme soit pro-cyclique.

Dans le cadre réglementaire actuel, la seule chose que peut faire le régulateur est d'augmenter les fonds propres réglementaires demandés pour une estimation de risque donnée, ou bien ne plus autoriser l'usage de modèles jugés trop optimistes. La situation est proche d'un jeu de délégation où le régulateur choisit un ensemble de niveaux de levier qu'une banque peut atteindre. La réponse des régulateurs à la crise financière et aux soupçons de biais optimiste portés sur les modèles de risque a été d'augmenter les fonds propres réglementaires, par exemple dans l'idée qu'il fallait couvrir le risque de modèle. Dans le cadre étudié ici, ce changement est équivalent à restreindre l'ensemble des niveaux de levier atteignables par les banques, et on pourrait s'attendre à ce que le résultat soit un niveau de risque dans le secteur bancaire plus faible. La proposition 2 montre que ce n'est pas toujours le cas : lorsque la régulation se durcit, l'offre de crédit décroît et le taux d'intérêt sur les prêts augmente. Il devient alors plus profitable de prêter davantage, si bien que davantage de banques choisissent des modèles optimistes. La combinaison de ces deux effets peut accroître le risque total, de manière contre-intuitive. Ainsi, si l'utilisation de modèles trop optimistes est due à un problème de sélection adverse et non de risque de modèle, les réformes réglementaires actuelles peuvent être largement contre-productives, les solutions à apporter étant bien différentes.

Il faut alors étudier comment inciter les banques à utiliser des modèles non biaisés, un problème de sélection adverse qui présente ici des spécificités intéressantes. Une possibilité naturelle est d'utiliser un menu de ratios de capital et de pénalités contingentes. Si une banque utilise un modèle très optimiste puis subit de lourdes

perdes, il s'agit d'un signal que le modèle était trop optimiste et la banque pourrait être pénalisée. La proposition 3 donne des conditions suffisantes pour que le régulateur soit capable de révéler le vrai modèle avec un tel mécanisme sans laisser de rente informationnelle à la banque régulée. Dans la régulation de Bâle, une banque peut toujours refuser d'utiliser l'approche utilisant les modèles internes et appliquer à la place une mesure standard de ses risques. En termes techniques, cela donne à la banque une option de sortie dépendant de son type : une banque refusant l'approche interne aura un profit plus élevé si elle sait par ailleurs que le risque de crédit est faible que si elle sait qu'il est élevé. La condition suffisante de la proposition 3 requiert que cette option de sortie ne réagisse pas trop à l'information privée de la banque, tandis que les prédictions issues des différents modèles possibles doivent au contraire être suffisamment différentes, ce qui rend l'observation des pertes finalement subies par la banque plus informative pour le régulateur.

Il faut ensuite prendre en compte le fait que la révélation de l'information repose ici sur la possibilité pour le régulateur de punir la banque après des pertes élevées. Si la régulation réagit beaucoup à l'information reçue, une banque avec un modèle optimiste se verra autoriser un niveau de levier élevé, de telle sorte que des pertes importantes ont de fortes chances de déclencher le défaut de la banque. Une fois en situation de défaut, la banque ne peut évidemment plus être punie. Ce problème peut empêcher le régulateur d'implémenter la régulation optimale de premier rang. La proposition 4 donne des conditions sous lesquelles ce problème se produit, en quel cas le régulateur fait face à un arbitrage entre laisser moins de rente à la banque ou implémenter une régulation plus sensible au modèle choisi.

S'il est difficile de s'appuyer sur des pénalités ex-post pour révéler l'information détenue par les banques sur le vrai modèle, une autre possibilité est d'essayer d'extraire cette information ex ante. Dans ce but la dernière section de cet article étudie un mécanisme d'audit, dans lequel le modèle utilisé par une banque est

d’abord audité par le régulateur, et la banque doit payer une pénalité si le modèle est identifié comme biaisé. L’hypothèse implicite ici est que pour utiliser un modèle biaisé une banque doit introduire quelques distorsions dans un modèle qui sans cela serait correct : utilisation de paramètres optimistes, choix de périodes particulières pour la calibration, d’une méthode d’estimation exotique... Si le modèle est audité pendant suffisamment longtemps, ces distorsions pourraient être identifiées. Mais l’audit est coûteux, si bien que le régulateur doit arbitrer entre le coût de la régulation et l’efficacité des ratios de capital. La proposition 5 montre que cet arbitrage conduit le régulateur à implémenter des ratios de capital moins sensibles au modèle interne choisi par la banque que dans l’optimum de premier rang, afin de diminuer les coûts d’audit. Dans le cas extrême où auditer est infiniment coûteux, si deux modèles impliquent des ratios différents il est impossible d’obtenir qu’une banque révèle le modèle le plus pessimiste, la seule possibilité est donc d’avoir une régulation “plate” dans laquelle les ratios ne dépendent pas du modèle choisi. Inversement, si l’audit n’a pas de coût alors l’utilisation d’un modèle biaisé est identifiée avec probabilité 1, si bien que le régulateur peut implémenter les ratios optimaux de premier rang.

Ce chapitre conclut à deux possibilités pour le régulateur. La première est de revenir à des ratios de capital plus simples, sans utilisation des modèles internes, ce qui implique de ne pas faire usage d’une information pertinente. La seconde est de rendre le cadre actuel plus complexe encore en donnant aux banques des incitations à améliorer leurs modèles, ce qui est à la fois difficile et coûteux. Une solution médiane est impossible, et traiter un problème de sélection adverse comme un problème de risque de modèle semble particulièrement dangereux. Une seconde conclusion est que le développement de modèles quantitatifs peut être affecté par les incitations, avec des effets potentiellement dévastateurs pour la stabilité du système financier. Enfin, une contribution d’intérêt plus général pour des

économistes est de comprendre un cas où là “demande” de modèles quantitatifs sur lesquels les agents basent leurs décisions et leurs anticipations est affectée par des considérations stratégiques.

Liens avec d’autres travaux : les travaux sur la régulation bancaire, aussi bien théoriques qu’empiriques, sont extrêmement nombreux. Sur le plan théorique, la régulation actuelle trouve l’une de ses sources dans les premiers travaux à avoir montré qu’imposer des contraintes sur le ratio de capital des banques allait donner des incitations à choisir des portefeuilles plus risqués, à moins que les actifs détenus par les banques ne soient “pondérés par les risques”, et que les poids utilisés soient corrects (Kim et Santomero (1988), Rochet (1992)). Certains de ces poids peuvent théoriquement être calculés à partir de l’information publique contenue dans les prix de marché, mais pour certains actifs typiquement détenus par des banques, notamment des prêts à des individus ou à des entreprises non cotées, une telle information n’est pas accessible au régulateur. Dans tous les cas, il est probable que la banque régulée ait davantage d’information que le régulateur sur la qualité des actifs qu’elle détient, si bien qu’on est en face d’un problème de sélection adverse (la banque a davantage d’information) et peut-être également d’aléa moral (plus ou moins d’effort peut être consacré à la surveillance des prêts). Chan, Greenbaum et Thakor (1992), Freixas et Rochet (1998) et Giammarino, Lewis et Sappington (1993) ont montré jusqu’à quel point une l’utilisation combinée de ratios de capital et de contributions au fond d’assurance des dépôts dépendant du risque déclaré par la banque pouvaient résoudre les deux problèmes.

On pourrait penser que le même type de mécanisme pourrait être utilisé pour inciter les banques à révéler le vrai modèle, et donc leur vrai risque. La manière dont fonctionne la régulation en pratique est cependant assez différente : dans de nombreux pays les contributions payées par les banques ne dépendent pas du risque

que celles-ci font courir au fond d'assurance, et quand elles en dépendent c'est via des ratios simples ou des mesures plus sophistiquées résultant de l'inspection du régulateur (par exemple les *CAMEL ratings* aux États-Unis), mais rien n'est prévu pour que ces contributions dépendent du modèle choisi par la banque pour estimer son risque. Cela peut néanmoins arriver indirectement, par exemple lorsque la contribution au fonds dépend du ratio de capital sur actifs pondérés par le risque, et que les poids sont calculés à l'aide de modèles internes. Mais dans ce cas l'assurance des dépôts ne fait que renforcer le problème : utiliser un modèle plus optimiste permet de prêter plus, et également de payer moins (dans la logique de [Chan, Greenbaum et Thakor \(1992\)](#) il faudrait qu'un modèle plus optimiste permette de prêter moins et payer moins).

Plutôt que d'étudier un mécanisme où la combinaison d'un paiement lié au risque et d'un ratio de capital permet de révéler le type de la banque, les deux types de mécanisme que j'étudie, *backtesting* et audit, sont beaucoup plus proches de ce qui se fait déjà dans la réalité. Le mécanisme de *backtesting* en particulier a été étudié pour le risque de marché dans un cadre plus appliqué par [Lucas \(2001\)](#) et [Cuoco et Liu \(2006\)](#), à comparer aux études empiriques de [Berkowitz et O'Brien \(2002\)](#), [Pérignon, Deng et Wang \(2008\)](#) et [Pérignon et Smith \(2010\)](#). Par rapport aux deux premiers, en prenant en compte que la banque doit baser ses estimations de risque sur un modèle je montre que la façon dont les prédictions varient lorsque des modèles différents sont adoptés joue un rôle dans la capacité du régulateur à obtenir la révélation du vrai modèle. Le mécanisme d'audit est proche de ce qu'étudie [Prescott \(2004\)](#), mais dans un cadre légèrement plus général j'obtiens une incitation à donner des ratios de capital plus faibles qu'à l'optimum non contraint pour certains types, et pas seulement des ratios plus élevés.

Enfin ce chapitre est lié à des articles récents sur ce qu'on pourrait nommer le "marché" des modèles, voire de la production académique. La plupart des articles

existants essaient d'expliquer comment la rationalité limitée des agents permet d'expliquer qu'ils adoptent des modèles biaisés, par exemple [Hong, Stein et Yu \(2007\)](#). Le même phénomène peut se produire dans des cadres plus complexes comme [Cogley, Colacito et Sargent \(2007\)](#). Ce chapitre enrichit ce champ en donnant l'exemple d'incitations rationnelles à choisir des modèles biaisés en présence de certaines frictions, venant ici d'une régulation inadéquate, point qui à ma connaissance n'a guère été étudié, avec l'exception possible de [Ghosh et Masson \(1994\)](#). Notons enfin que ces articles étudient la "demande" de modèles, tandis que d'autres comme [Saint-Paul \(2011\)](#) étudient le côté "offre", celui des producteurs de modèles. Il est intéressant de noter que les banques sont souvent des deux côtés, et que les incitations étudiées dans ce chapitre peuvent selon les cas être vues comme des incitations à adopter certains modèles plutôt que d'autres parmi ceux qui sont déjà disponibles, ou comme des incitations à développer certains modèles, ce qui peut être plus grave (en biaisant l'ensemble des modèles qui seront disponibles dans le futur).

3 Catching falling knives: speculating on market overreaction

Ce chapitre présente un modèle dynamique d'un marché financier où des investisseurs avec de l'information privée non liée à la valeur fondamentale des actifs spéculent sur la possibilité que le marché sursuragisse ou sous-réagisse aux échanges observés. Dans une période de crise où la composition du marché n'est pas connue de tous les participants, cette *supply information*, dans la terminologie de [Genotte et Leland \(1990\)](#), peut être aussi importante que l'information fondamentale sur la valeur des actifs. Les exemples incluent le krach de 1987, lorsque l'incertitude sur l'impact des stratégies d'assurance de portefeuille a rendu difficile l'identification des ventes qu'elles entraînaient comme peu informatives ([Jacklin, Kleidon et Pfeiderer \(1992\)](#)) ; les suites de la faillite de Lehman Brothers en septembre 2008, lorsque des baisses brutales de prix pouvaient être dues soit à des anticipations pessimistes sur les fondamentaux de l'économie soit à un nombre inconnu d'institutions financières tâchant de vendre leurs actifs pour baisser leur levier ; ou plus récemment le *flash crash* du 6 mai 2010, pendant lequel la remontée rapide des cours dépendait des croyances des investisseurs sur la cause de ce krach soudain, soit de l'information négative soit des perturbations dans la fourniture de liquidité (voir [Easley, Lopez et O'Hara \(2011\)](#)).

L'article utilise un modèle à la [Glosten et Milgrom \(1985\)](#) dans lequel j'introduis des traders à retour positif (*positive feedback traders*), dont la proportion n'est connue que de quelques *supply-informed traders*. Comme le retour positif peut être sous-estimé par les faiseurs de marché, des crises ou des bulles peuvent se produire à court terme dans un tel cadre, lorsque les faiseurs de marché sursuragissent aux achats ou aux ventes qu'ils observent, les croyant plus informatifs qu'ils ne sont. Je montre que les *supply-informed traders* tendent à corriger ces déviations de court

terme : ils se comportent en effet à contre-courant (*contrarians*) lorsque le marché surréagit, et comme des traders à retour positif eux-mêmes lorsque le marché sous-réagit. Ce comportement rend néanmoins plus difficile pour les faiseurs de marché d'apprendre si le marché comporte beaucoup de traders à retour positif ou non, si bien que l'impact de long terme de ces spéculateurs est ambigu. Ils peuvent de plus ne plus échanger sur le long terme si leur information n'est pas suffisamment précise, et leur profit a typiquement une kurtosis élevée. La plupart du temps ils n'échangent pas en raison de fourchettes de prix trop larges, mais avec une faible probabilité un krach se forme et ces traders essaient de "rattraper des couteaux qui tombent"⁶ : en moyenne ils font un profit parce qu'ils savent que le krach vient probablement de traders à retour positif, mais il peut aussi d'agir de traders véritablement informés, en quel cas ils font des pertes.

Les traders à retour positif sont modélisés comme des agents qui doivent vendre après chaque vente sur le marché et acheter après chaque achat, peut-être avec des probabilités différentes. La proportion de ces traders sur le marché est soit faible soit élevée, selon certaines probabilités connues ex ante des faiseurs de marché. On peut considérer le cas dans lequel il y a plus de traders à retour positif après des ventes qu'après des achats (pendant une période avec beaucoup de ventes forcées par exemple), et seuls les *supply-informed traders* savent que la proportion de traders à retour positif est élevée. En moyenne, les *supply-informed traders* vont s'approcher de la vérité sur la valeur de l'actif échangé après chaque échange, mais pas nécessairement les faiseurs de marché : beaucoup de ventes vont probablement se produire à cause du retour positif, mais comme les faiseurs de marché n'ont pas encore appris qu'il y avait beaucoup de retour positif sur le marché, ils vont mettre à jour leurs prix davantage vers le bas que les *supply-informed traders* ne mettent à jour leurs croyances, puisque ces derniers savent que beaucoup de ventes

⁶D'après un adage populaire sur les marchés financiers, *Never try and catch a falling knife*, "n'essayez jamais d'attraper un couteau qui tombe".

ne sont probablement pas informatives. La proposition 3 montre que les faiseurs de marché vont finir par apprendre la véritable proportion de retour positif et la valeur de l'actif, mais peuvent s'éloigner en moyenne un peu plus de la vérité après chaque observation pendant un nombre limité de périodes au début du jeu. La raison en est qu'ils vont typiquement prendre des échanges initiés par des traders à retour positif pour des échanges initiés par des traders informés sur la valeur de l'actif. La proposition 5 donne des conditions suffisantes pour que les prix divergent de la valeur fondamentale de l'actif en moyenne pour un certain nombre de périodes, conditionnellement à une réalisation donnée de la valeur de l'actif et de la proportion de retour positif. Pour certains paramètres, un agent omniscient connaissant ces deux valeurs anticiperait qu'en moyenne un krach ou une bulle doit arriver, avant un renversement et un retour à la valeur fondamentale.

Dans le même exemple où la valeur de l'actif est élevée et les ventes forcées nombreuses, les *supply-informed traders* mettent à jour leurs anticipations moins vers le bas que les faiseurs de marché et vont donc acheter, c'est-à-dire aller à contre-courant du marché. Plus généralement, la proposition 4 donne des conditions suffisantes pour que les *supply-informed traders* se comportent à contre-courant ou au contraire avec un retour positif sur le long terme. Intuitivement, le premier cas va avoir tendance à se produire lorsque le marché surréagit, et le second lorsqu'il sous-réagit. Dans une section consacrée à l'étude de simulations du modèle, je vérifie que ce comportement rend la réalisation d'un krach non fondamental moins probable lorsque le nombre de *supply-informed traders* s'accroît. Néanmoins, il devient plus difficile pour les faiseurs de marché de mettre à jour leurs croyances sur la proportion de traders à retour positif : après une vente, une nouvelle vente est un signal que cette proportion est élevée, et un achat qu'elle est faible. Mais si elle est élevée alors les *supply-informed traders* achètent, si elle est faible au contraire ils vendent, ce qui rend les deux observations moins informatives pour les

faiseurs de marché. Le résultat est qu'avoir plus de *supply-informed traders* peut en fait ralentir la convergence des prix à long terme, au moins sous les conditions énoncées dans le corollaire 2.

Par ailleurs, les *supply-informed traders* sont une seconde source de sélection adverse : les faiseurs de marché savent qu'ils risquent d'échanger avec quelqu'un qui sait si le marché surréagit ou sous-réagit. Le lemme 2 et le corollaire 1 montrent que si l'avantage informationnel des *supply-informed traders* est trop faible, ils ne vont échanger que pour un nombre fini de périodes parce que le *spread* demandé par les faiseurs de marché pour compenser le risque d'échanger avec un trader informé sur la vraie valeur de l'actif est suffisant pour décourager les *supply-informed traders* d'échanger. Le gros de leur activité se concentrera donc sur des épisodes de court terme de grandes déviations des prix dues aux traders à retour positif. Ils tendront à acheter (respectivement, à vendre) après de longues suites de ventes (respectivement, d'achat), en moyenne les prix tendront après quelques périodes à se redresser (respectivement, à retomber), en quel cas ils feront des profits importants, mais parfois les ventes (respectivement, les achats) seront effectivement déclenchées par à la fois des traders à retour positif et des traders avec de l'information négative (respectivement, positive), en quel cas les *supply-informed traders* feront au contraire des pertes élevées. Les profits issus de cette activité sont donc extrêmement risqués, ce que l'on peut vérifier au moyen de simulations (résultat numérique 1).

Dans une période agitée, le marché peut donc bénéficier de la présence d'agents informés sur la composition même du marché, en raison par exemple de leurs contacts dans des institutions financières ou avec des intermédiaires importants. De tels agents apportent une sorte d'assurance contre bulles ou krachs non fondamentaux : ils réduisent la probabilité d'occurrence de tels événements, au prix d'une sélection adverse accrue et parfois d'une convergence plus lente des prix sur le long terme. Malgré de récents travaux comme [Glosten \(2010\)](#), il reste difficile de

mesurer le bien-être des différents agents dans les modèles à la Glosten et Milgrom. Mais en raison de la forme du profit de ces traders, en particulier sa forte kurtosis, on peut penser que trop peu d’agents investiront dans l’acquisition de ce type particulier d’information, on a donc là un exemple de “limite à l’arbitrage”, due à la forme du pari que des spéculateurs doivent prendre dans ce contexte pour équilibrer le marché.

Liens avec d’autres travaux : ce chapitre est au croisement de deux champs de recherche différents et en général relativement séparés. Le premier rassemble des travaux sur l’incertitude “non-fondamentale” (ici, sur la présence plus ou moins prononcée de traders à retour positif) et leur impact sur le processus des prix, notamment [Easley et O’Hara \(1992\)](#). D’autres articles ont étudié plus précisément comment ce type d’incertitude pouvait mener à des bulles ou à des krachs à court terme, notamment [Genotte et Leland \(1990\)](#), [Jacklin, Kleidon et Pfleiderer \(1992\)](#) et [Romer \(1993\)](#). D’autres modèles partent de l’hypothèse qu’un volume élevé d’échanges même sans motivation fondamentale aura nécessairement un impact sur les prix, positif ou négatif selon qu’il s’agit d’achats ou de ventes. Il s’agit d’un raccourci qui suppose implicitement l’une des deux hypothèses suivantes : soit cet impact est dû à une fourniture imparfaite de liquidité et repose donc sur une hypothèse d’arbitrage limité et requiert de comprendre pourquoi il n’y a pas plus de capital alloué à l’arbitrage (voir [Gromb et Vayanos \(2010\)](#)) ; soit il est dû au contenu informatif de l’achat ou de la vente opérés, en quel cas il faut comprendre pourquoi plus de capital n’est pas alloué à l’arbitrage de situations où la mauvaise interprétation d’échanges non informés mènent à des prix éloignés de leurs fondamentaux, activité exercée par les *supply-informed traders*.

Le deuxième champ lié à ce chapitre concerne la spéculation, entendue comme le comportement d’agents plus informés que le marché. Plusieurs articles se fo-

calisent sur la question de la possibilité de “comportements moutonniers” sur les marchés financiers et, bien que le sujet du chapitre soit différent, des intuitions contenues dans [Avery et Zemsky \(1998\)](#), notamment la partie sur “l’incertitude de composition”, ainsi que dans [Park et Sabourian \(2011\)](#), qui donnent des conditions sous lesquelles un trader se comportera de manière contraire au marché ou aura un retour positif, jouent un rôle important dans le modèle. Enfin [Brunnermeier et Pedersen \(2009\)](#) étudient un problème proche où des “spirales de liquidité” vont se produire lorsque les spéculateurs font face à des appels de marge, ce qui peut produire des krachs temporaires. Mais dans ces trois articles les traders sont toujours soit non informés, soit informés à la fois sur la valeur fondamentale de l’actif et sur le paramètre “non fondamental” du marché s’il y en a un (incertitude de composition chez [Avery et Zemsky \(1998\)](#), choc de liquidité chez [Brunnermeier et Pedersen \(2009\)](#)). Or on peut penser que ces deux types d’information peuvent être détenus par des agents différents, et mener à des comportements différents. De plus, des agents qui pourraient acquérir de l’information sur la composition du marché pourraient ne pas avoir la possibilité d’acquérir de l’information sur la valeur des actifs, ce qui rend nécessaire de comprendre les incitations à acquérir le premier type d’information.

Le résultat selon lequel le comportement des *supply-informed traders* peut compliquer l’apprentissage des faiseurs de marché évoque les travaux sur la “spéculation déstabilisatrice”, notamment [Stein \(1987\)](#) où l’introduction de spéculateurs imparfaitement informés sur un choc de liquidité peut déstabiliser le marché en compliquant la situation informationnelle pour les agents non informés. D’autres travaux comme [DeLong, Shleifer, Summers et Waldmann \(1990\)](#), ou un peu différemment [Brunnermeier et Pedersen \(2005\)](#), avancent un argument différent en montrant qu’il peut être rationnel pour un spéculateur qui sait qu’un actif est sur-évalué de l’acheter et renforcer la surévaluation avant de vendre. D’un certain point de

vue le résultat de ce chapitre est encore plus négatif dans la mesure où, même dans l'hypothèse optimiste où les *supply-informed traders* n'ont pas la possibilité de suivre ce type de stratégie dynamique, ils ont un impact stabilisateur à court terme mais pas nécessairement à long terme.

Enfin ce chapitre est complémentaire à de nombreux travaux qui microfondent la présence d'agents exerçant un retour positif sur le marché, que ce soit en raison de contraintes de levier et d'appels de marge comme dans [Chowdhry et Nanda \(1998\)](#) ou [Brunnermeier et Pedersen \(2009\)](#), ou bien parce que des gestionnaires de portefeuille ont des incitations à aller dans le sens du marché pour préserver leur réputation ([Dasgupta, Prat et Verardo \(2011\)](#)). Plutôt que d'introduire de telles microfondations dans le modèle en même temps que des *supply-informed traders*, ce qui s'avère peu praticable, le modèle présenté dans ce chapitre prend comme donnée l'hypothèse que certains agents exercent du retour positif, et étudie l'impact informationnel de ces agents et dans quelle mesure des traders informés sur cette dimension non fondamentale atténuent ou renforcent cet impact.

4 Trading fees and efficiency in limit order markets

Cet article, co-écrit avec Thierry Foucault (HEC Paris), étudie l'impact de la concurrence croissante entre plateformes d'échanges sur la fourniture de liquidité sur les marchés à ordres à cours limité. L'entrée de nouvelles plateformes électroniques *low-cost* dans les dernières années, aussi bien en Europe qu'aux États-Unis, a conduit à la fois à une réduction des frais de transaction totaux sur les marchés financiers, et à des changements dans la structure de ces frais : beaucoup de plateformes choisissent en effet de subventionner les offreurs de liquidité (appelés *makers*) et de faire payer les demandeurs de liquidité (*takers*). Au printemps 2007 par exemple le NASDAQ faisait payer 30 cents par *round lot* (100 actions) pour les *takers* et payait au contraire 20 cents aux *makers*, tandis qu'à la fin 2010 le coût était de 29 cents pour les premiers et la subvention de 28,50 cents pour les seconds. Les frais totaux gagnés par la plateforme ont donc diminué (de 10 cents à 0,50 cents pour 100 actions), et l'asymétrie entre *makers* et *takers* s'est accrue. Ceci conduit à s'interroger sur l'impact de telles évolutions sur les stratégies des acteurs du marché et sur sa liquidité.

Nous étendons le modèle de marchés à ordres à cours limité de [Foucault \(1999\)](#) en introduisant de l'hétérogénéité entre investisseurs, des frais de transaction et de la concurrence entre plusieurs plateformes. Notre résultat le plus surprenant est que des frais de transaction plus faibles, et donc davantage de concurrence entre plateformes, peut en fait être néfaste aux investisseurs. A chaque fois qu'un échange a lieu entre deux investisseurs, il est évident que leurs gains totaux sont plus élevés lorsque les frais de transaction sont plus faibles. Néanmoins, pour des raisons détaillées ci-dessous, des frais de transaction plus faibles peuvent diminuer le nombre d'échanges à l'équilibre. Nous montrons également qu'en l'absence de frictions

supplémentaires, comme le fait de devoir placer son prix sur une grille discrète, la répartition des frais entre *makers* et *takers* est neutre : elle n'a d'impact ni sur les échanges à l'équilibre, ni sur le bien-être des investisseurs. Elle a un impact en revanche sur la fourchette de prix (*spread*), si bien que les études empiriques utilisant celle-ci pour mesurer l'impact de la fragmentation du marché sur la liquidité doivent être interprétées avec prudence.

Dans notre modèle, les investisseurs arrivent de manière séquentielle sur un marché à ordres à cours limité et peuvent être soit acheteurs soit vendeurs. Ils peuvent choisir de soumettre un ordre à cours limité et attendre une période pour l'exécution (non garantie) de leur ordre, ou bien soumettre un ordre au marché, ou encore se tourner directement vers un marché de *dealers* compétitifs. Les investisseurs diffèrent de plus dans leur capacité à attendre pour voir leur ordre exécuté : certains peuvent être par exemple des arbitragistes pour qui la rapidité de l'exécution est très importante. Nous montrons dans la proposition 1 que plusieurs régimes sont possibles à l'équilibre, selon les paramètres du modèle et en particulier le niveau total des frais de transaction et des *spreads* choisis par les *dealers*. Dans l'équilibre avec le volume d'échanges le plus élevé, les investisseurs patients comme les investisseurs impatients acceptent les offres dans le carnet d'ordres, et soumettent des ordres à cours limité eux-mêmes quand le carnet manque de liquidité. Il est aussi possible que les deux types d'investisseurs soumettent des ordres à cours limité, mais fassent des offres qui ne "visent" que les investisseurs les plus impatients. S'il devient plus profitable d'utiliser le marché de *dealers* que de soumettre un ordre à cours limité, on atteint un équilibre où seuls les investisseurs patients utilisent ces ordres et visent les investisseurs impatients, tandis que quand le carnet manque de liquidité les investisseurs impatients se tournent vers les *dealers*. Il est également possible que seuls les investisseurs patients soumettent des ordres à cours limité, mais que ces ordres soient acceptés par les deux types d'investisseurs.

Enfin il existe un équilibre dans lequel tous les investisseurs utilisent seulement le marché de *dealers*. Nous avons donc cinq types d'équilibre possibles, mais pour des valeurs données des paramètres l'équilibre obtenu est unique.

Pour chaque type d'équilibre nous pouvons calculer le nombre moyen d'échanges dans chaque période. Le corollaire 3 montre que ce nombre, le taux d'échanges, n'est pas une fonction monotone des frais de transaction choisis par la plateforme. En d'autres termes, augmenter ces frais peut accroître le taux d'échanges. Pour comprendre l'intuition de ce résultat, on peut considérer un acheteur patient voulant soumettre un ordre à cours limité : il peut soit faire une offre à un prix faible B , qui ne sera acceptée que si le prochain investisseur est impatient et veut échanger immédiatement, ou bien un prix plus élevé $B' > B$ qui sera accepté aussi bien par les patients que par les impatients. Un ordre au prix B' rapporte un profit moins important, mais est exécuté avec une probabilité plus élevée. Nous montrons dans l'article qu'une baisse des frais de transaction peut augmenter la différence entre B' et B et donner davantage d'incitations à choisir le prix élevé et la probabilité d'exécution faible, parce que les frais affectent davantage l'option des patients s'ils choisissent de refuser l'offre que l'option des impatients. Ce point est plus aisé à comprendre dans le cas où les investisseurs impatients le sont au point que, s'ils refusent un ordre, ils préfèrent échanger avec les *dealers* plutôt que de passer un nouvel ordre. Dans ce cas, leur option est par définition inchangée lorsque les frais de transactions augmentent, tandis que celle des investisseurs patients correspond typiquement à passer un nouvel ordre à cours limité, option qui a moins de valeur lorsque les frais de transaction augmentent.

Par ce mécanisme, la concurrence accrue entre les plateformes d'échanges peut mener les *makers* à essayer d'extraire davantage de rente des *takers*, entraînant un plus faible volume total et une liquidité moindre. Pour le montrer, nous endogénéisons le niveau des frais de transaction dans le cas d'une plateforme

d'échanges en monopole puis de deux plateformes se faisant concurrence sur les frais de transactions avant que les échanges ne commencent. Traiter le second cas nous conduit à étendre le jeu d'origine pour permettre aux investisseurs de choisir sur quelle plateforme ils veulent échanger, la proposition 4 montre alors qu'à l'équilibre les frais totaux sont de zéro lorsqu'il y a deux plateformes, tandis qu'ils sont strictement positifs dans le cas du monopole (étudié dans la proposition 2). De plus, le monopole fixe toujours ses frais de manière à éviter l'équilibre avec le volume d'échanges le plus bas, puisque en augmentant les frais il peut également atteindre un équilibre avec un volume plus élevé, et donc accroître à la fois prix et quantité. Il en résulte que sous certains paramètres l'entrée d'une nouvelle plateforme conduit à une diminution des frais de transaction et à une baisse du volume d'échanges. Nous proposons également une mesure du bien-être des investisseurs dans ce modèle, et montrons que celle-ci peut être maximisée avec des frais de transaction non nuls (corollaire 4). Cependant l'entrée d'une nouvelle plateforme augmente toujours le bien-être : comme le montre le corollaire 5, lorsqu'une plateforme en monopole conduit à plus d'échanges qu'un duopole il s'avère que le monopole extrait tout le surplus des investisseurs. Il est donc toujours préférable pour les investisseurs d'avoir deux plateformes plutôt qu'une, mais sous certains paramètres il peut être encore mieux d'avoir un *plancher* sur les frais de transaction imposé par le régulateur.

Nous regardons ensuite l'impact de la répartition des frais entre offreurs de liquidité (*makers*), et demandeurs (*takers*). Supposons une situation initiale avec un prix *ask* A et un prix *bid* B , et que la plateforme change la répartition des frais tout en gardant constant le total, par exemple augmente les frais pour les *takers* de ϵ et diminue les frais pour les *makers* du même montant. Si un *taker* acceptait dans la situation initiale l'ordre au prix B , il doit maintenant payer ϵ de plus qu'avant à la plateforme. Pour s'assurer qu'il accepte toujours son offre, le *maker*

doit proposer un nouveau prix $B' = B + \epsilon$, si bien que le *taker* obtient exactement le même surplus qu'auparavant. En faisant cela, le *maker* perd ϵ si l'échange est réalisé, perte exactement compensée par le fait qu'il doit alors payer ϵ de moins à la plateforme. Au final, l'équilibre reste inchangé, mais les prix sont différents : le *bid* a augmenté de ϵ et le *ask* a diminué de ϵ , si bien que le *spread* a diminué de 2ϵ . La répartition des frais de transactions n'affecte donc pas les gains des investisseurs, mais subventionner les *makers* et faire payer les *takers* rend les prix *apparemment* plus attractifs. Ces résultats sont résumés dans les corollaires 6, 7 et 8 de l'article, qui sont importants pour les études empiriques sur les effets de la fragmentation du marché sur la liquidité : si l'entrée de nouvelles plateformes implique une baisse des frais de transaction totaux et un changement dans leur répartition entre *makers* et *takers*, des *spreads* plus faibles ne reflètent pas nécessairement une liquidité plus élevée puisqu'ils peuvent être simplement dus à des coûts plus faibles pour les *makers*. Ces implications empiriques claires sur l'impact des frais de transaction sur les *spreads* sont aussi un moyen de tester la théorie, et certaines ont d'ailleurs été déjà confirmées par [Malinova et Park \(2011\)](#).

Une conclusion possible de cet article est que la répartition des frais de transaction entre *makers* et *takers* a peut-être reçu trop d'attention de la part des régulateurs et des praticiens, tandis que l'importance du niveau *total* de ces frais a été sous-estimée. Alors que la première a un impact clair sur le *spread* et la liquidité apparente du marché, elle est sans effet sur les gains des investisseurs et sur les échanges, tout changement dans la structure des frais étant neutralisé par les changements de prix. Des frictions comme l'obligation de faire une offre à un prix sur une grille discrète peuvent changer ce résultat, mais l'effet de la répartition des frais reste probablement plus faible qu'il ne semble à première vue. Inversement, l'effet du niveau total ne peut pas être vu comme un simple transfert de richesse entre les investisseurs et les plateformes : il affecte les stratégies des investisseurs et

donc les échanges à l'équilibre, parfois de façon contre-intuitive. En particulier, des frais de transaction plus faibles ne devraient pas nécessairement être vus comme un résultat bénéfique pour les investisseurs de la concurrence accrue entre plateformes.

Liens avec d'autres travaux : de nombreux articles ont étudié la concurrence entre différents "marchés" pour attirer les échanges. [Chowdhry et Nanda \(1991\)](#) par exemple étudient la concurrence entre des faiseurs de marché présents sur des marchés différents pour attirer les échanges. [Gehrig \(1993\)](#) étudie un modèle à plusieurs égards proche du nôtre, où l'on peut échanger soit avec des intermédiaires concurrentiels soit en cherchant une contrepartie (ce qui dans notre cadre reviendrait à passer un ordre à cours limité). [Yavas \(1992\)](#), [Glosten \(1994\)](#), [Hendershott et Mendelson \(2000\)](#), [Rust et Hall \(2003\)](#), [Parlour et Seppi \(2003\)](#) et [Degryse, Van Achter et Wuyts \(2009\)](#) comparent des marchés organisés différemment, voire étudient s'il est possible pour des marchés de différents types de coexister (ces articles étudient respectivement le cas de *marketmakers* vs. *matchmakers*, carnet d'ordres vs. *dealers*, *dealers* vs. *crossing network*, pur carnet d'ordres vs. marché hybride avec spécialistes et carnet d'ordres, *dealers* vs. *passive crossing network*). [Foucault et Menkveld \(2008\)](#) étudient la concurrence entre deux marchés utilisant des ordres à cours limité, mais prennent les frais de transaction comme donnés. De manière générale, tous les articles précités (voir [Cantillon et Yin \(2011\)](#) pour un survol de ce champ) soit ignorent les frais de transaction, soit ne les endogénéisent pas, l'attention se concentrant sur le choix des investisseurs entre différents mécanismes d'échanges. Inversement, dans notre article les frais de transaction sont endogénéisés sous deux hypothèses différentes (une seule ou plusieurs plateformes), et la concurrence est à plusieurs niveaux : les plateformes d'échange se font concurrence entre elles, mais sont aussi en concurrence avec le marché des *dealers* pour attirer des investisseurs qui serviront de *makers*. Les *deal-*

ers eux-même sont en concurrence avec les *makers*, dans la mesure où les deux types d'agents essaient d'attirer des investisseurs qui choisiront d'être *takers*.

D'autres articles ont montré que la concurrence entre plateformes pouvait avoir un impact négatif, par exemple Mendelson (1987), Pagano (1989), Madhavan (1995), Parlour et Seppi (2003). En général l'argument consiste à montrer que pour un même niveau de frais de transaction avoir deux plateformes plutôt qu'une peut être néfaste, si bien que même lorsque davantage de concurrence entraîne des frais moindres l'effet d'une concurrence accrue est ambigu. Dans notre article, c'est la baisse des frais de transaction elle-même qui peut être néfaste, au sens où le niveau optimal des frais de transaction peut être strictement positif. Degryse, Van Achter et Wuyts (2012) et Pagnotta et Philippon (2011) étudient également des modèles où les frais de transaction sont rendus endogènes, mais les premiers regardent les frais de compensation et de règlement, tandis que les seconds se concentrent sur la vitesse des échanges.

Foucault, Kadan et Kandel (2009) analysent plus spécifiquement la répartition des frais de transaction entre *makers* et *takers*, tandis que le présent article s'intéresse davantage au niveau total. La comparaison est intéressante en ce que dans Foucault, Kadan et Kandel (2009) la répartition des frais de transaction n'est pas neutre, mais deux hypothèses sont différentes : d'une part les investisseurs ne peuvent pas choisir d'être *makers* plutôt que *takers*, d'autre part les prix doivent être positionnés sur une grille discrète. C'est probablement la deuxième hypothèse qui est à l'origine des résultats différents : de petites asymétries dans les frais de transaction ne peuvent pas être compensées par un ajustement des prix, parce que le *tick*, l'incrément minimal d'un prix, est d'un ordre de grandeur supérieur à celui des frais de transaction. La comparaison des deux articles nous apprend que l'asymétrie des frais de transaction observée sur les marchés doit probablement plus à cette friction additionnelle qu'à la forme particulière d'organisation des échanges

que représente un carnet d'ordres à cours limité. En particulier, l'argument selon lequel les *makers* doivent être subventionnés par la plateforme parce qu'ils sont nécessaires pour attirer des *takers* est faux, de même que d'autres arguments similaires reposant sur une application hâtive des travaux sur les marchés bifaces. La raison en est qu'il est aussi nécessaire que les *makers* s'attendent à voir des *takers* sur la plateforme, et peuvent ajuster leurs prix à une structure asymétrique des frais de transaction.

References

- ACHARYA, V. V., AND L. H. PEDERSEN (2005): “Asset pricing with liquidity risk,” *Journal of Financial Economics*, 77(2), 375–410. 8
- ADRIAN, T., AND M. K. BRUNNERMEIER (2008): “CoVaR,” Staff Reports 348, Federal Reserve Bank of New York. 3
- AVERY, C., AND P. ZEMSKY (1998): “Multidimensional Uncertainty and Herd Behavior in Financial Markets,” *American Economic Review*, 88(4), 724–48. 7, 23
- BERKOWITZ, J., AND J. O’BRIEN (2002): “How Accurate Are Value-at-Risk Models at Commercial Banks?,” *The Journal of Finance*, 57(3), 1093–1111. 16
- BLACK, F., AND M. SCHOLES (1973): “The Pricing of Options and Corporate Liabilities,” *Journal of Political Economy*, 81(3), pp. 637–654. 5
- BRUNNERMEIER, M. K., AND L. H. PEDERSEN (2005): “Predatory Trading,” *Journal of Finance*, 60(4), 1825–1863. 7, 23
- (2009): “Market Liquidity and Funding Liquidity,” *Review of Financial Studies*, 22(6), 2201–2238. 3, 8, 23, 24
- CALOMIRIS, C. W., AND C. M. KAHN (1991): “The Role of Demandable Debt in Structuring Optimal Banking Arrangements,” *The American Economic Review*, 81(3), pp. 497–513. 6
- CANTILLON, E., AND P.-L. YIN (2011): “Competition between exchanges: A research agenda,” *International Journal of Industrial Organization*, 29(3), 329–336. 30
- CHAN, Y.-S., S. I. GREENBAUM, AND A. V. THAKOR (1992): “Is Fairly Priced Deposit Insurance Possible?,” *The Journal of Finance*, 47(1), pp. 227–245. 15, 16
- CHOWDHRY, B., AND V. NANDA (1991): “Multimarket Trading and Market Liquidity,” *Review of Financial Studies*, 4(3), 483–511. 30

- (1998): “Leverage and Market Stability: The Role of Margin Rules and Price Limits,” *The Journal of Business*, 71(2), pp. 179–210. 7, 24
- COGLEY, T., R. COLACITO, AND T. J. SARGENT (2007): “Benefits from U.S. Monetary Policy Experimentation in the Days of Samuelson and Solow and Lucas,” *Journal of Money, Credit and Banking*, 39, 67–99. 17
- CORSETTI, G., A. DASGUPTA, S. MORRIS, AND H. S. SHIN (2004): “Does One Soros Make a Difference? A Theory of Currency Crises with Large and Small Traders,” *Review of Economic Studies*, 71(1), 87–113. 7
- CUOCO, D., AND H. LIU (2006): “An analysis of VaR-based capital requirements,” *Journal of Financial Intermediation*, 15(3), 362 – 394. 16
- DANIELSSON, J. (2008): “Blame the models,” *Journal of Financial Stability*, 4(4), 321 – 328. 10
- DASGUPTA, A., A. PRAT, AND M. VERARDO (2011): “The Price Impact of Institutional Herding,” *Review of Financial Studies*, 24(3), 892–925. 24
- DE LONG, J. B., A. SHLEIFER, L. H. SUMMERS, AND R. J. WALDMANN (1990): “Positive Feedback Investment Strategies and Destabilizing Rational Speculation,” *The Journal of Finance*, 45(2), pp. 379–395. 7, 23
- DEGRYSE, H., M. VAN ACHTER, AND G. WUYTS (2009): “Dynamic order submission strategies with competition between a dealer market and a crossing network,” *Journal of Financial Economics*, 91(3), 319–338. 30
- (2012): “Internalization, clearing and settlement, and liquidity,” Open Access publications from Katholieke Universiteit Leuven urn:hdl:123456789/335865, Katholieke Universiteit Leuven. 31
- DIAMOND, D. W. (1984): “Financial Intermediation and Delegated Monitoring,” *The Review of Economic Studies*, 51(3), pp. 393–414. 5, 6

- DIAMOND, D. W., AND P. H. DYBVIK (1983): “Bank Runs, Deposit Insurance, and Liquidity,” *Journal of Political Economy*, 91(3), pp. 401–419. 5
- DIAMOND, D. W., AND R. E. VERRECCHIA (1987): “Constraints on short-selling and asset price adjustment to private information,” *Journal of Financial Economics*, 18(2), 277–311. 7
- EASLEY, D., M. M. LOPEZ DE PRADO, AND M. O’HARA (2011): “The Microstructure of the ”Flash Crash”: Flow Toxicity, Liquidity Crashes, and the Probability of Informed Trading,” *The Journal of Portfolio Management*, 37(2), 118–128. 18
- EASLEY, D., AND M. O’HARA (1992): “Time and the Process of Security Price Adjustment,” *The Journal of Finance*, 47(2), pp. 577–605. 22
- FAMA, E. F. (1970): “Efficient Capital Markets: A Review of Theory and Empirical Work,” *The Journal of Finance*, 25(2), pp. 383–417. 5
- FOUCAULT, T. (1999): “Order flow composition and trading costs in a dynamic limit order market,” *Journal of Financial Markets*, 2(2), 99–134. 25
- FOUCAULT, T., O. KADAN, AND E. KANDEL (2012): “Liquidity cycles and make/take fees in electronic markets,” *The Journal of Finance*, (Forthcoming). 31
- FOUCAULT, T., AND A. J. MENKVELD (2008): “Competition for Order Flow and Smart Order Routing Systems,” *The Journal of Finance*, 63(1), 119–158. 30
- FREIXAS, X., B. M. PARIGI, AND J.-C. ROCHET (2000): “Systemic Risk, Interbank Relations, and Liquidity Provision by the Central Bank,” *Journal of Money, Credit and Banking*, 32(3), 611–38. 7
- FREIXAS, X., AND J.-C. ROCHET (1997): *Microeconomics of Banking*. MIT Press, Cambridge (Mass.). 6
- FREIXAS, X., AND J.-C. ROCHET (1998): “Fair Pricing of Deposit Insurance. Is it possible? Yes. Is it desirable? No,” *Research in Economics*, 52(3). 15

- GEHRIG, T. (1993): “Intermediation in Search Markets,” *Journal of Economics & Management Strategy*, 2(1), 97–120. 30
- GENNOTTE, G., AND H. LELAND (1990): “Market Liquidity, Hedging, and Crashes,” *The American Economic Review*, 80(5), pp. 999–1021. 18, 22
- GHOSH, A., AND P. MASSON (1994): *Economic Cooperation in an Uncertain World*. Blackwell Publishing, Malden. 17
- GIAMMARINO, R. M., T. R. LEWIS, AND D. E. M. SAPPINGTON (1993): “An Incentive Approach to Banking Regulation,” *The Journal of Finance*, 48(4), pp. 1523–1542. 15
- GLOSTEN, L. (2010): “Welfare cost of informed trade,” Working paper, NBER and Columbia Business School. 21
- GLOSTEN, L. R. (1994): “Is the Electronic Open Limit Order Book Inevitable?,” *The Journal of Finance*, 49(4), 1127–1161. 30
- GLOSTEN, L. R., AND P. R. MILGROM (1985): “Bid, ask and transaction prices in a specialist market with heterogeneously informed traders,” *Journal of Financial Economics*, 14(1), 71–100. 6, 18
- GROMB, D., AND D. VAYANOS (2010): “Limits of Arbitrage,” *Annual Review of Financial Economics*, 2(1), 251–275. 22
- GROSSMAN, S. J., AND J. E. STIGLITZ (1980): “On the Impossibility of Informationally Efficient Markets,” *The American Economic Review*, 70(3), pp. 393–408. 6
- HENDERSHOTT, T., AND H. MENDELSON (2000): “Crossing Networks and Dealer Markets: Competition and Performance,” *The Journal of Finance*, 55(5), 2071–2115. 30
- HOLMSTROM, B., AND J. TIROLE (1997): “Financial Intermediation, Loanable Funds, and the Real Sector,” *The Quarterly Journal of Economics*, 112(3), pp. 663–691. 6

- HONG, H., AND J. C. STEIN (2003): “Differences of Opinion, Short-Sales Constraints, and Market Crashes,” *The Review of Financial Studies*, 16(2), pp. 487–525. 7
- HONG, H., J. C. STEIN, AND J. YU (2007): “Simple Forecasts and Paradigm Shifts,” *The Journal of Finance*, 62(3), 1207–1242. 17
- JACKLIN, C. J., A. W. KLEIDON, AND P. PFLEIDERER (1992): “Underestimation of Portfolio Insurance and the Crash of October 1987,” *The Review of Financial Studies*, 5(1), pp. 35–63. 18, 22
- JENSEN, M. C. (1986): “Agency Costs of Free Cash Flow, Corporate Finance, and Takeovers,” *The American Economic Review*, 76(2), pp. 323–329. 7
- KIM, D., AND A. M. SANTOMERO (1988): “Risk in Banking and Capital Regulation,” *The Journal of Finance*, 43(5), pp. 1219–1233. 15
- LELAND, H. E., AND D. H. PYLE (1977): “Informational Asymmetries, Financial Structure, and Financial Intermediation,” *The Journal of Finance*, 32(2), pp. 371–387. 5
- LINTNER, J. (1965): “The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets,” *The Review of Economics and Statistics*, 47(1), pp. 13–37. 5
- LUCAS, A. (2001): “Evaluating the Basle Guidelines for Backtesting Banks’ Internal Risk Management Models,” *Journal of Money, Credit and Banking*, 33(3), pp. 826–846. 16
- LUCAS, ROBERT E, J. (1978): “Asset Prices in an Exchange Economy,” *Econometrica*, 46(6), 1429–45. 4
- MADHAVAN, A. (1995): “Consolidation, Fragmentation, and the Disclosure of Trading Information,” *The Review of Financial Studies*, 8(3), 579–603. 31
- MALINOVA, K., AND A. PARK (2011): “Subsidizing Liquidity: The Impact of Make/Take Fees on Market Quality,” Working paper, available at [ssrn: http://ssrn.com/abstract=1823600](http://ssrn.com/abstract=1823600), University of Toronto. 29

MARKOWITZ, H. (1952): “Portfolio Selection,” *The Journal of Finance*, 7(1), pp. 77–91.

5

MENDELSON, H. (1987): “Consolidation, Fragmentation, and Market Performance,” *The Journal of Financial and Quantitative Analysis*, 22(2), 189–207. 31

MYERS, S. C., AND N. S. MAJLUF (1984): “Corporate financing and investment decisions when firms have information that investors do not have,” *Journal of Financial Economics*, 13(2), 187 – 221. 7

O’HARA, M. (1995): *Market Microstructure Theory*. Blackwell, Cambridge (Mass.). 6

PAGANO, M. (1989): “Trading Volume and Asset Liquidity,” *The Quarterly Journal of Economics*, 104(2), 255–274. 31

PAGNOTTA, E., AND T. PHILIPPON (2011): “Competing on Speed,” NBER Working Papers 17652, National Bureau of Economic Research, Inc. 31

PARK, A., AND H. SABOURIAN (2011): “Herding and Contrarian Behavior in Financial Markets,” *Econometrica*, 79(4), 973–1026. 23

PARLOUR, C. A., AND D. J. SEPPI (2003): “Liquidity-Based Competition for Order Flow,” *The Review of Financial Studies*, 16(2), 301–343. 30, 31

PERIGNON, C., Z. Y. DENG, AND Z. J. WANG (2008): “Do banks overstate their Value-at-Risk?,” *Journal of Banking and Finance*, 32(5), 783 – 794. 16

PERIGNON, C., AND D. R. SMITH (2010): “Diversification and Value-at-Risk,” *Journal of Banking and Finance*, 34(1), 55 – 66. 16

PRESCOTT, E. S. (2004): “Auditing and Bank Capital Regulation,” *Economic Quarterly*, 90(4). 16

RAMAKRISHNAN, R. T. S., AND A. V. THAKOR (1984): “The Valuation of Assets under Moral Hazard,” *The Journal of Finance*, 39(1), pp. 229–238. 5

- ROCHET, J.-C. (1992): “Capital requirements and the behaviour of commercial banks,” *European Economic Review*, 36(5), 1137 – 1170. 15
- ROMER, D. (1993): “Rational Asset-Price Movements Without News,” *The American Economic Review*, 83(5), pp. 1112–1130. 22
- ROSS, S. A. (1976): “The arbitrage theory of capital asset pricing,” *Journal of Economic Theory*, 13(3), 341 – 360. 5
- ROSS, S. A. (2002): “Neoclassical Finance, Alternative Finance and the Closed End Fund Puzzle,” *European Financial Management*, 8(2), 129–137. 4
- ROSS, S. A. (2004): *Neoclassical Finance*. Princeton University Press, Princeton. 4
- RUST, J., AND G. HALL (2003): “Middlemen versus Market Makers: A Theory of Competitive Exchange,” *Journal of Political Economy*, 111(2), 353–403. 30
- SAINT-PAUL, G. (2011): “Toward a Political Economy of Macroeconomic Thinking,” in *NBER International Seminar on Macroeconomics 2011*, NBER Chapters. National Bureau of Economic Research, Inc. 17
- SHARPE, W. F. (1964): “Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk,” *The Journal of Finance*, 19(3), pp. 425–442. 5
- STEIN, J. C. (1987): “Informational Externalities and Welfare-Reducing Speculation,” *Journal of Political Economy*, 95(6), 1123–45. 7, 23
- STOLL, H. R. (1978): “The Supply of Dealer Services in Securities Markets,” *The Journal of Finance*, 33(4), pp. 1133–1151. 6
- TREYNOR, J. L. (1962): “Toward a Theory of Market Value of Risky Assets,” Unpublished manuscript. 5
- YAVAS, A. (1992): “Marketmakers versus matchmakers,” *Journal of Financial Intermediation*, 2(1), 33–58. 30