Essays in financial economics

Edouard Chretien

NNT:2017SACLX025

THÈSE DE DOCTORAT
DE L'UNIVERSITÉ PARIS-SACLAY
PREPARÉE A L'ÉCOLE POLYTECHNIQUE

Ecole doctorale n°: 578
Sciences de l'homme et de la société - Pôle Éco-Gestion
Spécialité de doctorat: Sciences Économiques

Présentée et soutenue publiquement au Centre de Recherche en Economie et Statistiques,
le 23 Mai 2017 par

Édouard Chrétien

Essais en Économie Financière
Essays in Financial Economics

Directeur de thèse : M. Édouard Challe

Composition du jury :

Président du Jury :	M. François Geerolf	-	University of California Los Angeles
Rapporteurs :	M. Guillaume Plantin	-	Sciences Po Paris
	M. Andrei Shleifer	-	Harvard University
Suffrageant :	M. Olivier Loisel	-	ENSAE Paristech
Directeur de thèse:	M. Édouard Challe	-	École Polytechnique

Спросите, спросите их только, как они все, сплошь до единого, понимают в чем счастье? О, будьте уверены, что Колумб был счастлив не тогда, когда открыл Америку, а когда открывал ее; будьте уверены, что самый высокий момент его счастья был, может быть, ровно за три дня до открытия Нового Света, когда бунтующий экипаж в отчаянии чуть не поворотил корабля в Европу, назад! Не в Новом Свете тут дело, хотя бы он провалился. Колумб помер почти не видав его и, в сущности, не зная, что он открыл? Дело в жизни, в одной жизни, — в открывании ее, беспрерывном и вечном, а совсем не в открытии!

Идиот, Фёдор Михайлович Достоевский

Demandez, demandez-leur seulement comment tous, sans exception, ils comprennent le bonheur ? Ah ! soyez certains que ce n'est pas, quand il a découvert l'Amérique mais quand il a été sur le point de la découvrir que Colomb a été heureux. Soyez persuadés que le monument culminant de son bonheur s'est peut-être placé trois jours avant la découverte du Nouveau-Monde, lorsque l'équipage au désespoir s'est rebellé et a été sur le point de faire demi-tour pour revenir en Europe.
Il ne s'agissait pas ici du Nouveau-Monde, qui aurait pu s'effondrer. Colomb est mort l'ayant à peine vu et sans savoir, au fond, ce qu'il avait découvert. Ce qui compte, c'est la vie, la vie seule ; c'est la recherche ininterrompue, éternelle de la vie, et non sa découverte!

L'idiot, Fiodor Dostoïevski

Ask them, ask any one of them, or all of them, what they mean by happiness! Oh, you may be perfectly sure that if Columbus was happy, it was not after he had discovered America, but when he was discovering it! You may be quite sure that he reached the culminating point of his happiness three days before he saw the New World with his actual eyes, when his mutinous sailors wanted to tack about, and return to Europe! What did the New World matter after all? Columbus had hardly seen it when he died, and in reality he was entirely ignorant of what he had discovered. The important thing is life— life and nothing else! What is any 'discovery' whatever compared with the incessant, eternal discovery of life?

The Idiot, Fyodor Dostoyevsky

Remerciements et pensées

Ces mots sont les derniers ajustements apportés à cette thèse. Ils seront très probablement ceux qui retiendront le plus votre attention. Les mettre en première page, permet de vous épargner la lecture parfois aride des chapitres suivants, que je vous invite cependant à lire, commenter et critiquer. Malgré toute l'attention que je pourrais apporter aux paragraphes qui suivent, j'oublierai très certainement certains de mes lecteurs dans la longue liste de noms de ceux qui m'ont permis de réaliser ce travail, dans la longue liste de noms de ceux qui m'ont accompagné, entouré, aidé ou soutenu pendant les quelques années qui se sont écoulées depuis le début de cette thèse. Qu'ils acceptent mes excuses.

Je remercie tout d'abord Édouard Challe, d'avoir accepté de m'encadrer, me guider, et diriger cette thèse, avec une grande bienveillance. J'ai beaucoup appris de nos interactions, au cours notamment des deux articles que nous avons réalisés ensemble, et qui constituent une part importante de ce travail. Je le remercie pour sa porte, toujours ouverte, et pour la nécessaire liberté qu'il m'a laissée. Celle-ci m'a permis d'affiner mes goûts et mes choix de recherche, et m'a emmené vers le dernier chapitre, puis les travaux que je mène aujourd'hui. À lui la reconnaissance et l'affection d'un padawan à son maître Jedi.

Je remercie ensuite tout particulièrement les membres de ce jury d'avoir accepté de participer à cet exercice. Je remercie ainsi François Geerolf, et Olivier Loisel, pour les commentaires qu'ils ont pu m'apporter en vue de l'amélioration de ces articles, pour la grande justesse et la rigueur de leurs remarques, tant dans le cadre de cette soutenance, lors de la pré-soutenance, que lors des discussions que nous avons pu avoir auparavant. Je remercie Guillaume Plantin pour sa relecture attentive, ses commentaires précieux qui sont autant de pistes d'améliorations pour les articles présentés, et les différents discussions que j'ai pu avoir avec lui à diverses étapes durant le cours de cette thèse. *Last but not least, I would like to thank Andrei Shleifer, first for his invaluable guidance on the last work of this dissertation, for his continuous support and help over the years, for his impressive insights. His research tastes have greatly influenced this work and my own tastes, including the directions of my future research. I also thank him for having hosted me in Harvard,*

where I learned a lot, and for all the time spent in La Chambarelle, the runs, the talks and the wonderful food and wine tastings. I cannot be grateful or thankful enough for all the things he has done for me.

Le Corps de l'INSEE m'a offert de très belles conditions pour réaliser cette thèse, et mérite ma reconnaissance. Les trois premières années se sont déroulées essentiellement au CREST. Je remercie Francis Kramarz, pour m'avoir accueilli au CREST, et Pierre Cahuc pour m'avoir fait une place au LMA. Je remercie également les membres du CREST, pour les discussions éclairées, et les repas et cafés animés. Je pense en particulier aux membres passés ou présents du LMA: Gregory Corcos, François Fontaine, Étienne Lehmann, Franck Malherbet, Jean-Baptiste Michau, Isabelle Méjean, Eva Moreno-Galbis, Julien Prat, Alessandro Riboni. Pour leur bonne humeur, et leur présence chaleureuse, qui donnent au CREST cette ambiance si particulière, je pense également à Elise Coudin, Bruno Crépon, Laurent Davezies, Robert Gary-Bobo, Alessandro Iaria, Anett John, Thierry Kamionka, Thomas Le Barbanchon, Manasa Patnam, Vincent Pons, Mathilde Poulhes, Roland Rathelot, Arne Uhlendorff. Mes pensées vont également aux INSEE du fond du couloir, Jean-Baptiste Bernard, Benoît Campagne et Aurélien Poissonnier, que je remercie par ailleurs pour le squelette qui m'a servi de base à la structure de ce document. Enfin, je remercie les doctorants et doctorantes du CREST, compagnons de fortune et d'infortune, pour la bonne ambiance, la bonne humeur, et le soutien qu'ils ont su apporter tout au long de ces années. Je pense à Alicia, Andreea, Antoine, Arnaud, Audrey, Basile, Charles-Marie, Clémence, Daphné, Hélène, Ivan, Jamal, Jeanne, Jérémy, Jérôme, Jordan, Julie, Malka, Manon, Marianne, Mathilde, Mélina, Meriem, Pauline, Sandra, Sebastián, Xiaofei.

Le personnel administratif du CREST mérite également mes remerciements pour l'aide discrète et indispensable qu'ils ont su m'apporter, et pour les moments de détente partagés. Je pense à Sébastien David, Martine Germond, Murielle Jules, Arnaud Richet, Fanda Traoré, et Edith Verger. J'ai une pensée particulière pour Sophie Magon, qui nous a quitté, et sa famille. Enfin, je pense à Saïda et son rire communicatif.

In 2015, I spent the Spring Semester visiting the Harvard Econ Department. I am grateful to Philippe Aghion, Emmanuel Farhi, Sam Hanson, Victoria Ivashina, David Scharfstein, *Jeremy Stein and Adi Sunderam for accepting me in their doctoral courses, as well as for the discussions and meetings we had over the course of the stay. I learned a lot with them. I would also like to thank Matteo Maggiori for his sponsorship, and Alexandra, Ansgar, Max, and Xavier for the time spent together and fruitful discussions.*

Depuis Septembre 2015, je travaille à la Direction des Études de l'ACPR, d'abord au Service d'Études Actuarielles et Simulations, puis à la Cellule Recherche. Je tiens d'abord à remercier Olivier de Bandt, Anne-Sophie Borie-Tessier, Dominique Durant, Henri Fraisse, Anne-Lise Bontemps-Chanel, et Anne-Laure Kaminski pour m'avoir accueilli dans leur direction et service, puis cellule. Je suis particulièrement reconnaissant à Olivier de Bandt et Henri Fraisse pour m'avoir fait confiance lors de la mise en place de la Cellule Recherche, qui a été déterminante pour me permettre d'achever ce travail, et d'initier de nouveaux projets de recherche. Je remercie également mes collègues, pour les échanges nombreux que nous avons pu avoir ensemble. Ils m'ont beaucoup appris notamment sur la régulation bancaire et assurantielle ainsi que sur le monde de la recherche et des études empiriques. Ils ont par ailleurs su créer à tout moment une atmosphère de travail chaleureuse, et conviviale. Pour cela, un grand merci à Eugenio Avisoa, Farida Azzi, Évariste Beigneux, Fabrice Borel-Mathurin, Boubacar Camara, François-Daniel Castellani, Julie Coudrel, Mamadou Coulibaly, Clément Denis, Geoffrey Devost, Sébastien Diot, Stéphane Esquerré, Sandrine Lecarpentier, Frédéric Malgras, Samuel Slama, Jean-Luc Thévenon, Oana Toader, Sophie Tryhoen, Eric Vansteenberghe, et Frédéric Vinas. Une mention spéciale à mes cobureaux Claire Labonne, Mathias Lé et Pierre Pessarossi, pour leur aide, leur soutien et leur disponibilité quotidiennes.

Au cours de cette thèse, j'ai effectué d'autres rencontres qui m'ont aidé, écouté, soutenu et fait progresser. Elles participent à transmettre cet amour de cette quête intellectuelle qu'est la recherche et ce goût permanent de l'effort. Je pense en particulier à Johan Hombert, et Pierre-Olivier Weill. Qu'ils en soient remerciés. *For the support, and the different discussions we have been through, for the doubts and the confidence, for all the time spent trying to find the right questions to answer, which are now the foundations for new exciting projects, I'd like to thank Raj Iyer, Daisuke Ikeda and Michael Kumhof.* Enfin, je remercie Victor Lyonnet, pour sa collaboration sur le dernier chapitre de cette thèse, et les projets à venir que nous développons ensemble. Je le remercie par ailleurs pour les discussions moins fructueuses que nous avons nourris tout au long de cette thèse et qui ont contribué à faire oublier les difficultés qui ont jalonné cet exercice.

Sur un plan plus personnel je tiens à remercier ceux qui, sans prendre une part directe à l'élaboration de cette thèse, en ont suivi les évolutions, et ont été à mes côtés tout au long de ces années. Leur présence, leur affection, et leur soutien constant m'ont été indispensables, ainsi que la fraîcheur que m'apportait nos interactions. Merci en particulier aux anciens, Adrien "Titi", Martin "Tinmar", Philippe "Pilip" et Virgile "Cifrice", au Burâl, David "La Pougne" et Hugo "Chary" [1], au capitaine, Paul "La Chev", a l'officier sécu, Pierre-Adrien "Pan", à l'escrimeur, Renaud "Renal", au gamer Attico, aux marins Antoine, Sébastien "Seb", Thomas "Dut", Timothée "Tim" et Tristan, aux midships Matthieu "Midship" et Matthieu "Teach", aux collégiens Alexandre "Ilio" et Ivan, aux prépa boys Fajwel et Guillaume "Barrackovitch", au blond Charles-Henri "Carlito" et à leurs conjointes. Je pense également aux couples d'AMUR Cécile G. et Hugo, Cécile S. et JB, Lulu et Rémi, et au Sciences Po Crew Alix, Caroline, Olivier, Juliette, Léa, Nina, Pierre, Scott, Timothy, Vanessa, Violaine.

Enfin, mes pensées et mes remerciements vont à ma famille, qui m'a toujours et partout accompagné et soutenu, et qui a su prendre soin de moi à chaque instant. Maman, Papa, Lorraine et Garance, merci pour votre présence. Je pense également à Baboum, qui a fêté hier ses 90 ans, mes oncles et tantes, cousins et cousines, petits-cousins, et petites-cousines beaucoup trop nombreux pour être cités ici. J'ai en particulier une pensée émue pour Thierry, parti trop tôt, et sa famille, Clairette, Laurette et Floraine. Je remercie également la famille de Claire, qui n'a cessé de se passionner pour cette thèse, en particulier Francis, Corinne et Thierry, Anne-Sophie et Arnaud. Pour finir je pense à Claire, qui partage ma vie depuis plus de sept ans, et a vécu avec moi chaque étape de cette thèse, en m'offrant un support de tous les instants. Je mesure en finissant ces lignes combien tout cela a été précieux et nécessaire et la remercie du fond du coeur.

[1] Dont je tairais par pudeur le surnom plus commun

Contents

List of Tables

List of Figures

Chapitre

1 Résumé

The Paris-Saclay University allows their PhD students to write their dissertation in English, but requires a summary in French to be included as the onset of the dissertation. The following presentation aims at fulfilling this requirement, and relies heavily on the introduction and conclusion of the current work.

L'Université Paris-Saclay autorise ses étudiants à rédiger leur thèse en anglais. Elle exige toutefois un résumé en français, placé en première partie de la thèse. Cette présentation vise à remplir cette contrainte. Elle a pour objet de proposer une vue d'ensemble des travaux réalisés au cours de cette thèse, en reprenant, en version française les introduction et conclusion de la présente thèse.

Ce présent volume comporte trois chapitres principaux reprenant trois articles de recherche développés au cours de cette thèse, complétés de chapitres introductif et conclusif. La thématique générale abordée avec ces projets s'articule autour de problématiques théoriques d'économie financière. Toutefois, trois champs distincts sont successivement étudiés, en cohérence avec l'évolution de mes recherches et de mes centres d'intérêt tout au long de ce cheminement. Le premier article est une contribution à la théorie de la microstructure de marché, le second aborde les questions relatives aux jeux de coordination à information imparfaite, tandis que le dernier propose une vue théorique appliquée aux questions d'intermédiation financière. Les trois chapitres vont être présentés successivement dans les sections suivantes, ainsi que l'articulation entre eux.

1.1 Une perspective de marché

La première approche, qui constitue le premier chapitre de cette thèse, est une contribution à la littérature de microstructure de marché. Travail joint avec Édouard Challe, il a été publié sous cette forme dans le *Journal of Economic Theory*. La question principale au coeur de cette article est celle de la relation entre les choix d'ordre (ordres limites ou ordres de marché) par des *traders* imparfaitement informés, échangeant un actif sur un marché, et le contenu informationnel du prix de cet actif.

1.1.1 Approche et littérature

La microstructure du marché est un champ de la littérature financière traitant de l'impact de la décision des agents financiers, et de la conception des marchés sur les métriques financièrs. Le Groupe de travail sur la Microstructure de marché du Bureau national de recherche économique (NBER) définit son rôle comme suit :

> Le groupe de recherche sur la microstructure du marché est consacré à la recherche théorique, empirique et expérimentale sur l'économie des marchés de titres, y compris le rôle de l'information dans le processus de découverte de prix, la définition, la mesure, le contrôle et les déterminants des coûts de

> liquidité et de transaction et leurs Implications pour l'efficacité, le bien-être et la réglementation des mécanismes commerciaux alternatifs et des structures de marché. [1]

L'approche développée dans ce chapitre est théorique et étudie un marché financier stylisé et statique avec trois types de participants : des *traders* informés qui participent au marché selon les informations privées qu'ils possèdent sur les retombées futures de l'actif, des opérateurs qui participent au marché uniquement pour des raisons de lissage de leur consommation, sans rapport avec l'information qu'ils possèdent sur l'actif, appelés *noise traders*, et des *market makers* qui ferment les transactions, en fixant les prix. Bien que la littérature sur la microstructure du marché ait été initialement principalement concernée par l'impact des *market makers* dans le processus d'agrégation de l'information (voir par exemple Stoll (19), Glosten (8), Glosten (9)), une branche ultérieure a souligné le rôle majeur des *traders*. Cette branche suit le travail pionnier de Kyle (15), Glosten and Milgrom (10), Easley and Maureen (7), parmi d'autres. Dans ce sens, nous étudions un modèle d'attentes rationnelles concurrentielles avec des informations asymétriques dans une configuration statique étroitement liée aux développements de Hellwig (14), Grossman and Stiglitz (11), Diamond and Verrecchia (6), Admati (1), qui considère un marché multiple, et Vives (21). Notre attention est centrée sur les types d'ordres que les *traders* informés peuvent poster sur le marché. La littérature sur la microstructure de marché étudie essentiellement deux types d'ordres, pour des raisons de tractabilité : les ordres de marché et les ordres limites, représentés comme un continuum d'ordres indexés par le prix d'exécution (dans la tradition des attentes rationnelles). Les *traders* informés qui choisissent de définir des ordres de marché ne placent qu'un seul ordre sur le marché, tandis que le continuum d'ordre consiste en un nombre infini d'ordres, chacun étant associé à un prix auquel l'échange a lieu. La question abordée par ce chapitre est le type d'allocation des ordres et la façon dont l'information est agrégée en prix lorsque les opérateurs sont laissés libre de choisir leurs types d'ordre, et doivent payer un surcoût pour placer le second type.

[1] Voir la page Web NBER : $http : //www.nber.org/workinggroups/groups_desc.html$

1.1.2 Résultats principaux

L'article étudie la détermination conjointe du contenu informationnel des prix d'actifs et de la composition du marché par type d'ordres sur un marché financier profond et concurrentiel au sein duquel l'information sur les fondamentaux économiques est dispersée parmi les investisseurs. Plus précisément, nous considérons une structure de marché dans laquelle les investisseurs reçoivent un signal bruité sur la valeur du dividende, à la suite de quoi ils peuvent investir dans l'actif, et, pour se faire, utiliser deux types d'ordres. Le premier type est celui qui est classiquement étudié dans la littérature sur l'agrégation de l'information : chaque investisseur est supposé soumettre une «courbe de demande» complète, c'est-à-dire un ordre qui spécifie la quantité acquise d'actifs pour toute valeur possible du prix d'équilibre. Un tel ordre revient à placer un continuum d'ordres limites (*limit orders*), et nous supposons qu'une telle complexité requiert le paiement d'un coût fixe (mais potentiellement faible). L'intérêt de ce type d'ordre est qu'il couvre l'investisseur contre le risque de prix d'exécution, c'est -à-dire l'incertitude sur le prix effectif d'échange résultant du fait que les ordres sont placés avant d'être agrégés par le marché pour donner le prix d'équilibre. Alternativement, les investisseurs peuvent placer des « ordres de marché » (*market orders*), qui sont des ordres d'achat conditionnels à l'information privée mais non au prix d'équilibre. Ces ordres sont beaucoup plus simples que des courbes de demande complètes, et nous normalisons leur coût à zéro. Les investisseurs utilisant ce type d'ordre sont exposés au risque d'exécution de prix (en plus du risque de dividende), et ce risque additionnel les conduit en général à limiter la taille de leurs ordres, à signal privé donné. Dans ce cadre, nous posons la question élémentaire suivante : quelle est la composition d'équilibre du marché par type d'ordre, c'est-à-dire la proportion d'investisseurs choisissant chacun de ces ordres ? Notre premier résultat est qu'il y existe une substituabilité stratégique dans le choix des types d'ordre lorsque l'information privée des investisseurs est suffisamment précise. Autrement dit, le fait que certains investisseurs placent des courbes de demande est de nature à dissuader les autres investisseurs de le faire.

L'explication de ce résultat est simple : les investisseurs plaçant des courbes de demande étant protégés contre le risque de prix d'exécution, ils répondent de manière plus agressive à leur information privée que les investisseurs plaçant des ordres de marché. Lorsque l'information est très précise, cette agressivité conduit à aligner efficacement le prix d'actif au fondamental (autrement dit, le prix devient plus informatif), ce qui réduit le risque de prix d'exécution et donc l'incitation à se couvrir contre ce risque. Nous montrons qu'à mesure que la précision des signaux s'accroit, et donc que le prix d'actif devient de plus en plus informatif, la fraction des investisseurs plaçant des courbes de demandes se réduit, pour devenir asymptotiquement résiduelle.

Ainsi, il apparaît que la microstructure de marché la plus couramment supposée dans la littérature n'est pas robuste à l'introduction d'un type d'ordre alternatif beaucoup plus simple et marginalement moins coûteux

1.2 Marchés et paniques

Le second chapitre adopte une vue plus large, en incluant le mécanisme développé dans le premier chapitre en première étape d'un jeu global à deux étapes. Ce chapitre, qui reprend également un article joint avec Édouard Challe, cherche à étudier avec davantage de précision l'impact de la structure de marché sur la sélection d'équilibre dans un jeu de coordination, qui considère le prix comme un signal informationnel public pour le jeu de coordination.

1.2.1 Approche et littérature

Les jeux de coordination avec complémentarité stratégique constituent une classe de jeux dans laquelle les décisions et les actions des différents agents se renforcent mutuellement. Comme l'ont souligné Cooper and John (5), une conséquence naturelle est que cette classe de jeu a tendance à générer des équilibres multiples. Compte tenu de la variété des phénomènes économiques associés à de tels jeux, en particulier dans les domaines de la macroéconomie et de la finance (attaques spéculatives, crise de la dette, ...) et l'ambiguïté positive des résultats d'équilibres multiples, un champ de la littérature en théorie des jeux a cherché à développer des techniques de rafinnement permettant la séclection d'équilibre. A la suite de Carlsson and van Damme (4), le champ des jeux globaux a été développé. Il présente un raffinement grâce à des techniques de perturbation, qui contribuent à réduire la multiplicité d'équilibre. L'idée fondamentale est que les perturbations, qui prennent la forme d'information imparfaite privée, peuvent générer suffisamment de différenciations entre les agents pour permettre des comportements différenciés entre eux. En supprimant la coordination parfaite des agents entre eux, associée à l'utilisation d'un signal unique de coordination, la théorie des jeux globaux permet d'éliminer la multiplicité d'équilibres qui en découle. La question de la robustesse de ce résultat d'unicité à l'introduction d'une théorie des prix en première étape d'un jeu global de coordination a été soulevée par Atkeson (2), dans son commentaire à Morris and Shin (16). L'idée sous-jacente étant qu'un tel marché, en agrégeant l'information dispersée, pouvait conduire à une surpondération d'un signal public endogène dans les informations qu'utilisent les agents pour leur décisions dans la seconde étape du jeu. Ce point a été abordé dans Werning and Angeletos (22) et Tsyvinski, Mukherji, and Hellwig (20). Ils montrent ainsi que la multiplicité d'équilibre peut effectivement être restaurée par l'existence de prix agissant comme un signal public endogène, à condition que l'information privée soit suffisamment précise. La question que nous posons dans cet article est la robustesse de ce résultat à la modélisation du marché financier lui-même.

1.2.2 Résultats principaux

Les spéculateurs qui envisagent une attaque (par exemple sur une monnaie) doivent deviner les croyances d'autres spéculateurs, l'attaque ne pouvant réussir que si un nombre suffisant d'agents y prennent part. Ils peuvent utiliser le marché boursier pour établir leur décision. Cet article examine si ce processus de collecte d'informations a un effet stabilisateur en améliorant les anticipations ou a un effet déstabilisateur en donnant lieu à des équilibres multiples auto-réalisateurs. Pour étudier cela, nous analysons les résultats d'un jeu global en deux étapes dans lequel un prix de l'actif déterminé au stade initial du jeu fournit un signal public endogène sur le fondamental qui complémente l'ensemble informationnel à disposition des traders qui choisissent d'attaquer dans la phase de coordination du jeu . Au cours de la phase de marché, le placement d'un continuum d'ordres limites est coûteux, mais les commerçants peuvent utiliser des ordres de marché plus risqués, car soumis au risque d'éxécution du prix. Le risque d'exécution des prix réduit l'agressivité des traders et ralentit donc l'agrégation de l'information, ce qui rend les équilibres multiples dans la phase de coordination moins probables. Dans ce sens, les frictions de microstructure qui conduisent à une plus grande exposition individuelle (au risque d'exécution des prix) peuvent réduire l'incertitude globale (en supprimant un résultat d'équilibre unique). Le résultat de ce chapitre est que les frictions de microstructure sur le marché financier peuvent avoir une incidence sur l'information intégrée dans un signal public utilisé à des fins de coordination. Même si dans notre configuration simple, cette friction n'élimine pas toutes les possibilités d'équilibres auto-réalisateurs, il est capable de réduire la zone d'équilibres multiples. Une raison probable de la persistence d'une zone d'équilibre multiple lorsque les signaux privés sont extremement précis est que la configuration que nous proposons offre une option binaire en ce qui concerne le risque d'exécution de prix : les agents peuvent choisir à un coût fixe d'éliminer complètement ce risque d'exécution de prix, ce qui favorise l'agrégation d'information dans le prix, et tempère notre effet de sélection d'équilibre.

1.3 Crises et intermédiation financière

Le point de vue du dernier chapitre est davantage orienté vers la littérature macro-financière. La question au coeur de ce travail est celle de la coexistence et l'interaction entre le secteur bancaire traditionnel, et le secteur dit *shadow banking*, qui s'est développé depuis les années 70 pour atteindre une taille actuelle d'un ordre de grandeur similaire à la taille du secteur bancaire classique. L'analyse effectuée ici considère l'interaction entre ces deux secteurs en temps de crise comme prisme d'analyse et en tire une série de conclusions sur les possibles canaux d'interactions entre eux. Travail joint avec Victor Lyonnet, il a reçu le Prix du Cercle K2 en Finance, et a bénéficié d'un financement de recherche de la part de l'Institut Louis Bachelier et de l'Institut Europlace de Finance.

1.3.1 Approche et littérature

Ce chapitre s'oriente davantage vers une litérature de type macro-finance, en particulier en théorie bancaire et théorie de l'intermédiation appliquée. Il se rapproche particulièrement de la littérature actuelle qui cherche à étudier le développement du *shadow banking*. Le *shadow banking* se définit comme un système bancaire parallèle, constitué de longues chaînes d'intermédiation, qui, prises dans leur ensemble, effectuent des activités similaires à celles qui caractérisent le secteur bancaire traditionnel (octroi de crédit à l'actif, création d'un actif sans risque et liquide au passif), sans avoir recours à la liquidité de la banque centrale ou aux garanties - explicites - du secteur public. Une partie de cette littérature vise à mesurer et à définir avec précision ce système financier, ainsi que la manière dont la crise de 2007-2009 l'a affecté et a alteré ses tendances en terme de développement. Les références comprennent Pozsar, Adrian, Ashcraft, and Boesky (18), He, Khang, and Krishnamurthy (13), Begenau, Bigio, and Majerovitz (3). Un autre aspect de la littérature a étudié la façon dont les banques traditionnelles interagissent avec les *shadow banks*. L'accent a été mis sur la vision de l'arbitrage réglementaire que procure ce système bancaire parallèle (voir par exemple Plantin (17)). Enfin, un aspect de la littérature s'est concentré sur la manière dont ces deux types d'institutions coexistent. Hanson, Shleifer, Stein, and Vishny (12) développent un modèle dans lequel ces deux types d'intermédiaires coexistent en investissant dans différents types d'actifs. Notre contribution s'inscrit à l'interaction de ces deux derniers volets.

1.3.2 Résultats principaux

Nous présentons un modèle des interactions entre banques traditionnelles et *shadow banks* qui explique leur coexistence. Lors de la crise financière de 2007, certains actifs et passifs des *shadow banks* se sont déplacés vers des banques traditionnelles et les actifs ont été vendus à des prix de *fire sale*. Notre modèle se fonde sur ces faits stylisés, et les réplique. La différence entre les banques traditionnelles et les *shadow banks* est double. Tout d'abord, les banques traditionnelles ont accès à un fonds de garantie qui leur permet d'émettre de la dette aux ménages en cas de crise. Deuxièmement, les banques traditionnelles doivent faire face à une réglementation coûteuse. Nous montrons qu'en temps de crise, les *shadow banks*liquident les actifs pour rembourser leurs créanciers, tandis que les banques traditionnelles achètent ces actifs aux prix de *fire sale*. Cet échange d'actifs en temps de crise génère une forme de complémentarité stratégique entre les banques traditionnelles et les *shadow banks*, où chaque type d'intermédiaire bénéficie de la présence de l'autre. Nous trouvons deux effets concurrents d'une petite diminution du soutien des banques traditionnelles en temps de crise, que nous appelons un effet de substitution et un effet de revenu. Ce dernier effet domine le premier, de sorte qu'un soutien anticipé plus faible aux banques traditionnelles en temps de crise induit plus de banquiers à gérer une banque traditionnelle *ex ante*. Ce chapitre propose également une approche normative. Il étudie l'allocation des banquiers dans les deux types de technologie d'intermédiation. Lorsque les deux types de banques coexistent dans l'allocation d'équilibre, il y a, en général, une inefficience : les banquiers ont tendance à trop s'allouer dans le secteur *shadow banking*. L'inefficience est liée au fait que les banquiers, en s'allouant dans le secteur *shadow banking*, ne parviennent pas à internaliser le fait que cette allocation réduit le support total dont peuvent bénéficier les *shadow banks* en temps de crise, ce qui réduit la capacité de levier de l'ensemble des *shadow banks*, qui ne sont donc pas aptes à investir suffisamment. Ce chapitre a fourni une première tentative théorique de prise en compte de la complémentarité de la banque traditionnelle et parallèle, pour comprendre la façon dont le système financier est construit et façonné. Cette complémentarité s'ajoute à la substituabilité habituelle proposée dans la littérature bancaire (à savoir la vue d'arbitrage réglementaire de la banque parallèle). Cette complémentarité comporte des implications postives et normatives que nous mettons en avant, et dont les implications sur la réglementation bancaire restent à étudier.

1.4 Suite de la dissertation

Pour la suite de la disseration j'effectue un retour à l'anglais et je reprends dans le détail les trois articles succintement présentés ici. Le chapitre 2 est une introduction, en anglais, au reste de la dissertation. Les chapitres 3, 4, 5 reprennent dans l'ordre les trois travaux présentés ici. Le chapitre 6 conclut.

Ce dernier chapitre présente également en exemple un projet de recherche courant, qui fait suite à cette thèse. Joint avec Rajkamal Iyer et Victor Lyonnet ce projet cherche à étudier avec davantage de précision ce que les ménages valorisent dans les dépôts bancaires qui constituent une part majoritaire de leurs actifs financiers. En effet, leur détention d'actifs exigibles à vue semble extrêmement importante au vue de leurs besoins de liquidités. Une étude expérimentale cherchera a étudier la valorisation de l'exigibilité par les agents dans leur détentions d'actifs. Une étude théorique propose la généralisation de ce types de contrats, sûrs et exigibles à vues, comme une réponse par les intermédiaires financiers à la demande par les agents d'actifs uniquement sûrs. Les conséquences normatives de ce type de réponse y seront étudiées. Ce projet bénéficie d'un financement de la Fondation Banque de France, à laquelle j'exprime ma gratitude.

Bibliographie

[1] Anat Admati. A noisy rational expectations equilibrium for multi-asset securities markets. *Econometrica*, 53(3) :629–57, 1985.

[2] Andrew Atkeson. Rethinking multiple equilibria in macroeconomic modeling : Comment. *NBER Macroeconomics Annual*, 15 :162–171, 2000.

[3] Juliane Begenau, Saki Bigio, and Jeremy Majerovitz. What can we learn from the financial flows of the 2008-2009 crisis ? *mimeo, Harvard University*, 2016.

[4] Hans Carlsson and Eric van Damme. Global games and equilibrium selection. *Econometrica*, 61(5) :989–1018, 1993.

[5] Russell Cooper and Andrew John. Coordinating coordination failures in keynesian models. *The Quarterly Journal of Economics*, 103(3) :441–463, 1988.

[6] Douglas Diamond and Robert E. Verrecchia. Information aggregation in a noisy rational expectations economy. *Journal of Financial Economics*, 9(3) :221–235, 1981.

[7] David Easley and O'Hara Maureen. Price, trade size, and information in securities markets. *Journal of Financial Economics*, 19(1) :69–90, 1987.

[8] Lawrence R Glosten. Insider trading, liquidity, and the role of the monopolist specialist. *The Journal of Business*, 62(2) :211–35, 1989.

[9] Lawrence R Glosten. Is the electronic open limit order book inevitable ? *Journal of Finance*, 49(4) :1127–61, 1994.

[10] Lawrence R. Glosten and Paul Milgrom. Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics*, 14 (1) :71–100, 1985.

[11] Sanford Grossman and Joseph Stiglitz. On the impossibility of informationally efficient markets. *American Economic Review*, 70(3) :393–408, 1980.

[12] Samuel G. Hanson, Andrei Shleifer, Jeremy C. Stein, and Robert W. Vishny. Banks as patient fixed-income investors. *Journal of Financial Economics*, 117(3) :449–469, 2015.

[13] Zhiguo He, In Gu Khang, and Arvind Krishnamurthy. Balance sheet adjustments during the 2008 crisis. *IMF Economic Review*, 58(1) :118–156, August 2010.

[14] Martin Hellwig. On the aggregation of information in competitive markets. *Journal of Economic Theory*, 22(3) :477–498, 1980.

[15] Albert Kyle. Continuous auctions and insider trading. *Econometrica*, 53(6) :1315–35, 1985.

[16] Stephen Morris and Hyun Song Shin. Rethinking multiple equilibria in macroeconomic modeling. In *NBER Macroeconomics Annual 2000, Volume 15*, pages 139–182. National Bureau of Economic Research, Inc, 2001.

[17] Guillaume Plantin. Shadow Banking and Bank Capital Regulation. *Review of Financial Studies*, 28(1) :146–175, 2015.

[18] Zoltan Pozsar, Tobias Adrian, Adam B. Ashcraft, and Hayley Boesky. Shadow banking. *Economic Policy Review*, (Dec) :1–16, 2013.

[19] Hans Stoll. The supply of dealer services in securities markets. *Journal of Finance*, 33(4) :1133–51, 1978.

[20] Aleh Tsyvinski, Arijit Mukherji, and Christian Hellwig. Self-fulfilling currency crises : The role of interest rates. *American Economic Review*, 96(5) :1769–1787, 2006.

[21] Xavier Vives. The speed of information revelation in a financial market mechanism. *Journal of Economic Theory*, 67(1) :178–204, 1995.

[22] Iván Werning and George-Marios Angeletos. Crises and prices : Information aggregation, multiplicity, and volatility. *American Economic Review*, 96(5) :1720–1736, 2006.

Chapter 2

General introduction

The three chapters of this dissertation are grounded in very diversified fields of the financial economics literature, ranging from market microstructure, to global games and banking theory. They follow the developments of my reflexion and research interests over the past few years and propose a theoretical approach on three distincts questions, at a micro and macro level, associated to the impact of financial frictions on prices, self-fulfilling crises and financial architecture.

The purpose of this introduction is to propose a brief summary of the different results obtained, and the literature they are embedded into, while emphasizing the links between them.

2.1 A market perspective

The first approach constitutes the first chapter of this dissertation. It is a contribution to the market microstructure literature, which is joint work with Edouard Challe, and has been published in the *Journal of Economic Theory*. The central question adressed in this paper is the one of the relationship of the choice of orders (limit orders, market orders) made by imperfectly informed traders trading an asset on a market, and the informationnal content of the price.

2.1.1 Modelling approach and related literature

Market microstructure is a financial subfield dealing with the impact of financial market decision and design on financial outcomes. The National Bureau of Economic Research Market Microstructure Working Group defines his role as follows:

> The market microstructure research group is devoted to theoretical, empirical, and experimental research on the economics of securities markets, including the role of information in the price discovery process, the definition, measurement, control, and determinants of liquidity and transactions costs, and their implications for the efficiency, welfare, and regulation of alternative trading mechanisms and market structures. [1]

The approach developped in this chapter is theoretical, and studies a stylised, static financial market with three types of participants: informed traders who participate in the market according to the private information they own on the security's future payoff, noise traders who participate in the market only for liquidity consumption smoothing reasons, unrelated to the information they own about the security, and market makers who close the deals, setting prices up. While the early market microstructure literature has been mainly concerned with the impact of market makers in the information aggregation process (see e.g. Stoll (19), Glosten (8), Glosten (9)), a later branch has emphasized the role of privately informed traders. This branch follows pioneering work of Kyle (15), Glosten

[1] See the NBER webpage: $http://www.nber.org/workinggroups/groups_desc.html$

and Milgrom (10), Easley and Maureen (7), among many others. In this vein, we study competitive rational expectations model with asymmetric information in a static setup closely related to developments by Hellwig (14), Grossman and Stiglitz (11), Diamond and Verrecchia (6), Admati (1), who considers a multiasset market, and Vives (21). Our focus is on the types of orders informed traders are allowed to post in the market. The market microstructure literature studies essentially two types of orders: market orders and full demand schedule (in the rational expectations tradition). Informed traders who choose to set market orders place only one order in the market, while full demand schedule consists of an infinite number of orders, each being associated with a price at which the trade takes place. The question this chapter adresses is the type of orders allocation, and the way information is aggregated into prices when you let traders choose their order types.

2.1.2 Main results

The article examines the joint determination of the informational content of asset prices and market composition by order type on a deep and competitive financial market in which information on economic fundamentals is dispersed among investors. Specifically, we consider a market structure in which investors receive a noisy signal on the value of the dividend, following which they can invest in the asset, and, in order to do so, use two types of orders. The first type is the one which is conventionally studied in the literature on the aggregation of information: each investor is supposed to post a complete demand curve, that is to say an order that specifies the quantity of acquired/sold for any possible value of the equilibrium price. Such an order is tantamount to placing a continuum of limit orders, and we assume that such complexity requires the payment of a fixed (but potentially low) cost. The interest of this type of order is that it covers the investor against the risk of the execution price, that is to say the uncertainty about the effective exchange price resulting from the fact that the orders are placed before being aggregated by the market to give the equilibrium price. Alternatively, investors can place market orders, which are conditional purchase orders to private information but not to the equilibrium price. These orders are much simpler than complete demand curves, and we normalize their cost to zero. Investors using this type of order are exposed to the risk of price execution (in addition to the dividend risk), and this additional risk generally leads them to limit the size of their orders, with a given private signal. In this framework, we ask the basic question: what is the equilibrium composition of the market by type of order, i.e. the proportion of investors choosing each of these orders? Our first result is that there is a strategic substitutability in the choice of order types when the private information of investors is sufficiently precise. In other words, the fact that some investors place demand curves is likely to dissuade other investors from doing so.

The explanation for this result is simple: investors placing demand curves are protected against the run-price risk, they respond more aggressively to their private information than investors placing market orders. When the information is very precise, this aggressiveness leads to efficient alignment of the asset price to the fundamental (in other words, the price becomes more informative), which reduces the risk of execution price and therefore the incentive to hedge against this risk. We show that as the accuracy of the signals increases, so that the asset price becomes more and more informative, the fraction of investors placing demand curves is reduced to become asymptotically residual. It thus appears that the market microstructure most commonly assumed in the literature is not robust to the introduction of a much simpler and marginally less costly alternative order type. Introducing endogenous order's choice in a standard model of market microstructure yields to distortions in the information aggregation by prices, compared to the one usually found in this literature: as people tend to choose less costly orders, the information aggregation is less efficient than in the standard approach with full demand schedule, implying that observed prices in markets can in fact be a much more blurry signal about asset's future dividends than in a situation where agents cannot choose their orders. When considering markets where severe pricing adjustments occur in times of crises (e.g. sovereign debt market), the individual reluctance to pay for costlier types of order might be a driver of bad information aggregation into prices.

2.2 Markets and runs

The second chapter adopts a broader view, by embedding the mechanism developped in the first chapter into a two-stage global game. This chapter, which is also joint work with Edouard Challe, has been submitted. The question we wish to adress is the impact of market microstructure in a financial market on the equilibrium selection in a coordination game which considers the price outcome observed in the financial market as a public information.

2.2.1 Modelling approach and related literature

Coordination games with strategic complementarities (i.e such that the decisions and actions of the different agents mutually reinforce one another) tend to generate multiple equilibria, as emphasized by Cooper and John (5). Given the variety of economic phenomena associated to such games, in particular in the fields of macroeconomics and finance (speculative attacks, debt crises,..), and the positive ambiguity of multiple equilibria outcomes, a strand of the game theoretic literature has seeking to develop equilibrium selection techniques. Following Carlsson and van Damme (4), global games technique have been developed. They introduce refinement through perturbation techniques, which help pinning down unique equilibrium. The fundamental idea is that perturbations generate enough differentiations between the agents to allow for differentiated behaviors between the agents. One concern with the robustness of this result is that the introduction of a theory of prices in global coordination games may reintroduce multiplicity of equilibria. The concern about the robustness of the global game result has been raised by Atkeson (2) in his comment toMorris and Shin (16) . This point has been formalized and addressed in Werning and Angeletos (22) and Tsyvinski, Mukherji, and Hellwig (20). They show that equilibrium multiplicity may be restored by the existence of prices acting as an endogenous public signal, provided that private information is sufficiently precise. The question we are after in this paper is the robustness of this result to the modelling of the financial market itself.

2.2.2 Main results

Speculators contemplating an attack (e.g., on a currency peg) must guess the beliefs of other speculators, which they can do by looking at the stock market. This paper examines whether this information-gathering process is stabilizing by better anchoring expectations ñor destabilizing by creating multiple self-fulfilling equilibria. To do so, we study the outcome of a two-stage global game wherein an asset price determined at the trading stage of the game provides an endogenous public signal about the fundamental that a§ects tradersí decision to attack in the coordination stage of the game. In the trading stage, placing a full demand schedule (i.e., a continuum of limit orders) is costly, but traders may use riskier (and cheaper) market orders, i.e., order to sell or buy a fixed quantity of assets unconditional on the execution price. Price execution risk reduces traders aggressiveness and hence slows down information aggregation, which ultimately makes multiple equilibria in the coordination stage less likely. In this sense, microstructure frictions that lead to greater individual exposure (to price execution risk) may reduce aggregate uncertainty (by pinning down a unique equilibrium outcome). The outcome of this chapter is that microstructure frictions in the financial market may impact the information embedded into a public signal used for coordination purpose, even if in our simple setup, this friction does not eliminate all possibilities for self-fulfilling equilibria, it is able to reduce the multiple-equilibria zone. A potential reason for this is that the setup we propose offers a binary option with respect to price execution risk: agents are allowed to choose at a fixed cost to eliminate completely their price-execution risk, which is likely to favorize information aggregation into prices, hence dampening our equilibrium selection effect.

2.3 Crises and financial intermediation

The point of view of the last chapter is more oriented towards the macro-financial literature. The issue at the heart of this work is the coexistence and interaction between the traditional banking sector and the *shadow banking* sector, which has developed since the 1970s to the current size whoseorder of magnitude is similar to the size of the conventional banking sector. The analysis carried out here considers the interaction between these two sectors in times of crisis as a prism of analysis and draws a series of conclusions about the possible channels of interactions between them. Joint work with Victor Lyonnet, it has been awarded the Prize of the K2 Circle in Finance, and beneftied from a research grant from the Institut Louis Bachelier and the Europlace Institute of Finance.

2.3.1 Modelling approach and related literature

This chapter is embedded into the literature on shadow banking, which broadly defines a parallel banking system, consisting in whole chains of intermediation performed outside the traditional banking system, which taken together provide similar services as the traditional banking system without access to central bank liquidity or public sector credit guarantees. Part of this literature aims at measuring, and defining closely the Shadow Banking system, as well as the way the 2007-2008 crisis affected it. References include Pozsar, Adrian, Ashcraft, and Boesky (18), He, Khang, and Krishnamurthy (13), Begenau, Bigio, and Majerovitz (3). Another strand of the literature has studied the way shadow banks interact with banks. An emphasis has been made on the regulatory arbitrage view of shadow banks (see for instance Plantin (17)). Finally, a strand of the literature has focused on the way these two types of institutions coexist. Hanson, Shleifer, Stein, and Vishny (12) develop a model where these two types of intermediaries coexist by investing in different types of assets. Our contribution falls at the intersection between these two last strands.

2.3.2 Main results

We present a model of the interactions between traditional and shadow banks that explains their coexistence. In the 2007 financial crisis, some of shadow banks' assets and liabilities have moved to traditional banks, and assets were sold at fire sale prices. Our model is able to accommodate these stylized facts. The difference between traditional and shadow banks is twofold. First, traditional banks have access to a guarantee fund that enables them to issue claims to households in a crisis. Second, traditional banks have to comply with costly regulation. We show that in a crisis, shadow banks liquidate assets to repay their creditors, while traditional banks purchase these assets at fire-sale prices. This exchange of assets in a crisis generates a complementarity between traditional and shadow banks, where each type of intermediary benefits from the presence of the other. We find two competing effects from a small decrease in traditional banks' support in a crisis, which we dub a substitution effect and an income effect. The latter effect dominates the former, so that lower anticipated support to traditional banks in a crisis induces more bankers to run a traditional bank ex-ante. This chapter also proposes a normative approach. It studies the allocation of bankers in both types of intermediation technology. When both types of banks coexist in the equilibrium allocation, there is generally an inefficiency: bankers tend to allocate too much into the shadow banking sector. The inefficiency is related to the fact that bankers, while allocating themselves towards the shadow banking sector, fail to internalize the fact that this allocation reduces the total support that shadow banks can obtain in bad times which reduces the leverage of all shadow banks, which are therefore too much constrained in investing. This chapter provided a first theoretical attempt to take into account traditional and shadow bank's complementarity, in understanding the way the financial system is built and shaped. This complementarity comes on top of the usual substituability put forward in the shadow banking literature (namely the regulatory arbitrage view of shadow banking). This complementarity has both postive and normative implications which we think as having potential regulatory implications. These need to be further studied.

2.4 Rest of the dissertation

In the following chapters of the dissertation, each of this topic is adressed more thoroughly. They are presented as their articles counterpart, which may give rise to overlaps or similarities across chapters.

Chapters 3, 4, 5 present the above articles in this order. Chapter 6 concludes and develops with an extension towards my future and current research projects.

Bibliography

[1] Anat Admati. A noisy rational expectations equilibrium for multi-asset securities markets. *Econometrica*, 53(3):629–57, 1985.

[2] Andrew Atkeson. Rethinking multiple equilibria in macroeconomic modeling: Comment. *NBER Macroeconomics Annual*, 15:162–171, 2000.

[3] Juliane Begenau, Saki Bigio, and Jeremy Majerovitz. What can we learn from the financial flows of the 2008-2009 crisis? *mimeo, Harvard University*, 2016.

[4] Hans Carlsson and Eric van Damme. Global games and equilibrium selection. *Econometrica*, 61(5):989–1018, 1993.

[5] Russell Cooper and Andrew John. Coordinating coordination failures in keynesian models. *The Quarterly Journal of Economics*, 103(3):441–463, 1988.

[6] Douglas Diamond and Robert E. Verrecchia. Information aggregation in a noisy rational expectations economy. *Journal of Financial Economics*, 9(3):221–235, 1981.

[7] David Easley and O'Hara Maureen. Price, trade size, and information in securities markets. *Journal of Financial Economics*, 19(1):69–90, 1987.

[8] Lawrence R Glosten. Insider trading, liquidity, and the role of the monopolist specialist. *The Journal of Business*, 62(2):211–35, 1989.

[9] Lawrence R Glosten. Is the electronic open limit order book inevitable? *Journal of Finance*, 49(4):1127–61, 1994.

[10] Lawrence R. Glosten and Paul Milgrom. Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics*, 14 (1):71–100, 1985.

[11] Sanford Grossman and Joseph Stiglitz. On the impossibility of informationally efficient markets. *American Economic Review*, 70(3):393–408, 1980.

[12] Samuel G. Hanson, Andrei Shleifer, Jeremy C. Stein, and Robert W. Vishny. Banks as patient fixed-income investors. *Journal of Financial Economics*, 117(3):449–469, 2015.

[13] Zhiguo He, In Gu Khang, and Arvind Krishnamurthy. Balance sheet adjustments during the 2008 crisis. *IMF Economic Review*, 58(1):118–156, August 2010.

[14] Martin Hellwig. On the aggregation of information in competitive markets. *Journal of Economic Theory*, 22(3):477–498, 1980.

[15] Albert Kyle. Continuous auctions and insider trading. *Econometrica*, 53(6):1315–35, 1985.

[16] Stephen Morris and Hyun Song Shin. Rethinking multiple equilibria in macroeconomic modeling. In *NBER Macroeconomics Annual 2000, Volume 15*, pages 139–182. National Bureau of Economic Research, Inc, 2001.

[17] Guillaume Plantin. Shadow Banking and Bank Capital Regulation. *Review of Financial Studies*, 28(1):146–175, 2015.

[18] Zoltan Pozsar, Tobias Adrian, Adam B. Ashcraft, and Hayley Boesky. Shadow banking. *Economic Policy Review*, (Dec):1–16, 2013.

[19] Hans Stoll. The supply of dealer services in securities markets. *Journal of Finance*, 33(4):1133–51, 1978.

[20] Aleh Tsyvinski, Arijit Mukherji, and Christian Hellwig. Self-fulfilling currency crises: The role of interest rates. *American Economic Review*, 96(5):1769–1787, 2006.

[21] Xavier Vives. The speed of information revelation in a financial market mechanism. *Journal of Economic Theory*, 67(1):178–204, 1995.

[22] Iván Werning and George-Marios Angeletos. Crises and prices: Information aggregation, multiplicity, and volatility. *American Economic Review*, 96(5):1720–1736, 2006.

Chapter

3

Market composition and price informativeness in a large market with endogenous order types

We analyse the joint determination of price informativeness and the composition of the market by order type in a large asset market with dispersed information. The market microstructure is one in which informed traders may place market orders or full demand schedules and where market makers set the price. Market-order traders trade less aggressively on their information and thus reduce the informativeness of the price; in a full market-order market, price informativeness is bounded, whatever the quality of traders' information about the asset's dividend. When traders can choose their order type and demand schedules are (even marginally) costlier than market orders, then market-order traders overwhelm the market when the precision of private signals goes to infinity. This is because demand schedules are substitutes: at high levels of precision, a residual fraction of demand-schedule traders is sufficient to take the trading price close to traders' signals, while the latter are themselves well aligned with the dividend. Hence, the gain from trading conditional on the price (as demand-schedule traders do) in addition to one's own signal (as all informed traders do) vanishes.

JEL-code : *A320, B747*

Keywords: *Market microstructure; Price informativeness; Market orders.*

3.1 Introduction

We analyse the joint determination of price informativeness and the composition of the market by order type in a large, competitive asset market with dispersed information. The market microstructure we consider is one in which informed traders may place either full demand schedules or more basic *market orders*, i.e., order to sell or buy a fixed quantity of assets unconditional on the execution price.[1] There are also "noise" traders who prevent the asset price from being fully revealing whenever the precision of private signals is bounded, as in, e.g., Grossman and Stiglitz [7,8], Diamond and Verrecchia [6] and others. After informed and noise traders have placed their orders, the trading price is set by a competitive, utility-maximising market making sector. We characterise the trading intensities associated with each order type, the *ex ante* utilities that they generate for the concerned traders (hence their preference for a particular type of order), and ultimately how the composition of the market interacts with the informativeness of the price.

We first consider the case where *exogenous* measures of demand-schedule and market-order traders coexist in the market. In a pure market-order market (as in, e.g., Vives [16]), the informativeness of the price is bounded above, however precise private information about the dividend is. In contrast, whenever demand-schedule traders are in positive mass the informativeness of the price is unbounded as the quality of private information goes to infinity. The reason for this difference lies in the way private information is incorporated into the price in either case. Because market-order traders face price risk –since their orders are unconditional on the effective trading price–, they trade less aggressively on their private information than demand-schedule traders, which reduces the informativeness of the price. In contrast, demand-schedule traders are insulated from price risk, so their trading intensity grows without bound as their private information becomes infinitely precise; in the limit, they perfectly align the trading price of the asset with the dividend (formally, the trading price is at least as informative of the dividend as the signals received by informed traders, and sometimes more).

[1] See Brown and Zhang [4], Wald and Horrigan [17] and Vives [15] for further disscussion of the importance of market orders in actual asset markets.

Motivated by these observations, we examine informed traders' choices of order type and the impact of these choices on the composition of the market and the informativeness of the price. Since demand schedules are more complex than market orders (due to the full conditionality of the amount of trade on the price), we assume that they are more costly, at least marginally. Our main result is that, when the precision of private signals is large, then *the equilibrium is necessarily interior* (i.e., market-order and demand-schedule traders are both in strictly positive measures), but *market-order traders overwhelm the market* (i.e., their measure tends to one as precision goes to infinity). In other words, when the quality of information is high, the gain from conditioning one's trades on the price (as demand-schedule traders do) in addition to conditioning on one's own signal only (as market-order traders do) vanishes –and thus falls short of the cost, however small, for most traders.

There are two potential reasons for which this could be the case, and it is the purpose of the information structure that we assume –with potentially correlated noise in the signals received by informed traders– to disentangle them. First, knowledge of the price could become less and less valuable because one's own signal becomes more and more aligned with the information of others as the quality of private information improves (since *all* signals then get closer to the true value of the dividend). In other words, the advantage of acquiring information about the distribution of the signals received by others through the price (as in, e.g., Grossman and Stiglitz [8]; Diamond and Verrecchia [6]) is reduced when this distribution tightens, and vanishes in the limit. The other reason why knowing the price could become less valuable when private information become very precise is related to the price impact of noise trading. Demand-schedule traders trade against observed discrepancies between the trading price and their signals, and their trading intensity rises with the precision of the signal. As this precision goes to infinity, they trade so aggressively against noise traders that they effectively close the gap between the trading price and the dividend. By eliminating noise trader risk, demand-schedule traders reduce the value of knowing the price in addition to the signal itself. Considering the full spectrum of signal correlations allows us to identify which of the two informational roles of the price drives our main result. We show that the use of demand schedules vanishes at high levels of information precision even when signals are perfectly correlated, i.e. when information about the dividend is public. In this situation the price no longer plays any role as an aggregator of dispersed information. It follows that it is the reduction in the impact of noise trading on the equilibrium price that explains why knowledge of the price loses value as the information about the dividend becomes very accurate. Put differently, our analysis uncovers a form of substituability between demand schedules: when signals about the dividend are accurate, a small fraction of demand-schedule traders is enough to keep the price close to the signals they receive, which are themselves close to the true value of the dividend; hence, it is less useful to know the price in addition to one's own signal, so the incentive to purchase a demand-schedule is reduced.[2] In this sense, the pure demand schedule specification (the benchmark in the literature on price informativeness) is not innocuous and may not be stable to plausible changes in the microstructure of the market (here: the availability of a simpler, but cheaper, alternative order type).

Our analysis relates to at least two strands of the literature: one that explores the properties of asset prices under alternative order types and market microstructures, and one that studies the joint determination of information acquisition and equilibrium prices. The focus on market orders –as opposed to limit orders, stop orders or full demand schedules– in the market microstructure literature can be traced back to Vives [16], Medrano [10] and Brown and Zhang [4].[3] Vives [16] studies a pure market-order market while Vives [15] considers a market with exogenous sets of trader types. Medrano [10] analyses the order choice of a single monopolistic trader, in the tradition of the "insider trading" literature. In contrast, we consider the endogenous determination of the sets of market-order versus demand-schedule traders in a large, competitive asset market. Brown and Zhang [4] study traders' order choice in a large market, but in their model those who do not place mar-

[2] This relies on demand schedule traders not reaching measure zero, in which case the price would no longer be well aligned with their signals. We show, however, that this cannot be the case under endogenous order types. The reason is that in a full market-order market the aggressiveness of informed traders is bounded above, hence these traders no longer eliminate noise trader risk when signals are very precise; this makes demand schedules valuable again and ensures that demand-schedule traders have strictly positive measure.

[3] See Medrano [10] and the references therein for a detailed disucssion of the early literature on market orders, and Vives [15] for the more recent papers.

ket orders are “dealers” who observe the order flow but are uninformed about the asset dividend. The interest in the joint determination of equilibrium prices and information acquisition started with Grossman and Stiglitz [7], followed by Verrecchia [13] and more recently Barlevy and Veronesi [3], Peress [12] and Vives [14]. While this literature has traditionally focused on information acquisition about asset payoffs, our focus is on traders' willingness to purchase an information set that includes the trading price (as is the case with a demand schedule) –in addition to a free signal about the asset payoff.

Many of our results follow from the fact that market-order traders trade less aggressively than demand-schedule traders and thereby reduce price informativeness. Let us stress that this is by no means the only reason why price informativeness may be impeded relative to the competitive, full demand-schedule benchmark. First, there might be some unlearnable residual uncertainty about the dividend, a possibility explored by Angeletos and Werning [1] in a somewhat different context. They show that this causes the precision of the price signal to be bounded above, whatever the precision of the private signals on the learnable part of the dividend. Market frictions may also limit traders' reaction to their information and thus price informativeness. For example, short-sale constraints limit traders' responsiveness to bad news (Miller [11]; Diamond and Verrecchia [5]; Bai et al. [2]). Similarly, under imperfect competition traders reduce their trading intensity so as to avoid revealing their private information (Kyle [9]). Our approach is closer in spirit to the latter contribution in that limited trading aggressiveness follows from the microstructure of the market, rather than outside restrictions about the learnability of the information or the size of trades.

Section 2 presents the trading game. Section 3 analyses the case where order types are exogenous, and Section 4 that where they are endogenous. Section 5 concludes the paper.

3.2 The model

We consider the following competitive model of asset trading. There are two assets: (i) a riskless bond in perfectly elastic supply and paying out a constant interest rate; and (ii) a risky asset with trading price p and terminal dividend θ, where θ is drawn from the distribution $\mathcal{N}\left(\bar{\theta}, \alpha_\theta^{-1}\right)$, $\alpha_\theta > 0$, before any trading takes place. Traders know the distribution of θ but not its realisation.

There is a continuum of *informed traders* $i \in I = [0,1]$, each of whom gets a free private, noisy signal about the dividend $x_i = \theta + \alpha_x^{-1/2}\xi_i$, with $\alpha_x > 0$, $\xi_i \sim \mathcal{N}(0,1)$ and $\operatorname{cov}(\xi_i, \theta) = 0$. We allow the noise in the private signals to be cross-correlated and parameterise this property by the correlation coefficient $\rho \in [0,1]$. We adopt the convention that the average of i.i.d. random variables with mean zero is zero and we let $\tilde{\xi} \equiv \int_0^1 \xi_i \mathrm{d}i \sim \mathcal{N}(0,\rho)$ and $\eta_i \equiv \xi_i - \tilde{\xi} \sim \mathcal{N}(0, 1-\rho)$ denote, respectively, the *aggregate* and *idiosyncratic* components of the noise in the private signals.[4] As ρ gets closer to one the information received by informed traders gets increasingly shared between themselves; when $\rho = 1$ private signals are perfectly correlated and the private signal ξ_i is just the public signal $\tilde{\xi}$. In contrast, as ρ goes down then the noise components become increasingly uncorrelated across traders, and we recover the usual specification with uncorrelated informational noise with $\rho = 0$. We may rewrite the private signal x_i as follows:

$$x_i = \theta + \alpha_x^{-1/2}\tilde{\xi} + \alpha_x^{-1/2}\eta_i, \tag{3.1}$$

which implies that x_i provides a noisy signal about $\theta + \alpha_x^{-1/2}\tilde{\xi}$ (with neither θ nor $\tilde{\xi}$ being individually observed).

Aside from informed traders, there are *noise traders* in the market, which place a net asset demand for the risky asset $\varepsilon \sim \mathcal{N}\left(0, \alpha_\varepsilon^{-1}\right)$, with $\alpha_\varepsilon > 0$. Following Vives [15,16], we consider a competitive market microstructure wherein (a) all or some traders place *market orders* (rather than full demand schedules), and (b) a (competitive, risk-neutral) market-making sector sets the price p. In contrast to a demand schedule, a market order is conditional on the private signal x_i but not on the execution price p; once placed, it is executed irrevocably at whatever value of p is set by market makers. The market-making sector observes the order book $L(\cdot)$ emanating from informed and noise traders and sets the price p; competition among risk-neutral market makers then causes them to undercut

[4] See Vives [14] for a model with a similar form of cross-correlation between agent types (formulated in terms of the marginal utility of a good, rather than a direct signal about an asset payoff).

each other until $p = \mathbb{E}[\theta | L(\cdot)]$. Note that $L(\cdot)$ is itself a function of p whenever a positive mass of informed traders places demand schedules.

We introduce a general correlation structure for the signals received by informed traders for the following reason. In our model placing a full demand schedule (as opposed to a market order) allows a trader to trade conditionally on the price, which is valuable for two very distinct reasons. First, the price aggregates dispersed private information and thus provides additional information about the fundamental relative to one's own signal. Second, it provides information about the realised amount of noise trading; that is, it effectively allows a trader to partly hedge noise trader risk (to which a market-order trader is exposed). For $\rho \in [0,1)$, these two informational roles of price are present, but when $\rho = 1$ only the second one is. Hence, considering the full correlation spectrum $[0,1]$ will shed light on which role of the price really matters for our results.

Let $M \subset I$ be the set of market-order traders and $I \backslash M$ the complementary set of demand-schedules traders, of measure $\nu \equiv \int_{I\backslash M} \mathrm{d}i \in [0,1]$. We will consider both the case where M and $I\backslash M$ are exogenous (Section 3) and that where they are endogenous (Section 4). All informed traders have zero initial wealth (this is without loss of generality) and preferences $V(w_i;\gamma_i) = -e^{-\gamma_i w_i}$, where $\gamma_i > 0$ and $w_i = (\theta - p)\,k_i$ are the risk aversion coefficient and terminal wealth of trader i, respectively. We denote by $\gamma_I^{-1} \equiv \int_I \gamma_i^{-1}\mathrm{d}i$ the average risk tolerance of informed traders. We rank informed traders in nondecreasing order of risk aversion and define the nondecreasing function $\gamma : [0,1] \to \mathbb{R}_+$. Finally, we assume that (i) γ_i is increasing and continuous in i and such that $\gamma_0 > 0$; and (ii) γ_i^{-1} is independent of $\xi_i - \tilde{\xi}$, i.e., $\forall J \subset I$, $\int_J \gamma_i^{-1}(\xi_i - \tilde{\xi})\mathrm{d}i = 0$ a.s.

Definition. A *Bayesian equilibrium* of the trading game is a pair of investment functions for demand-schedule $(k_{I\backslash M}(x_i, p; \gamma_i))$ and market-order $(k_M(x_i;\gamma_i))$ traders as well as a price function $p(\theta, \tilde{\xi}, \varepsilon)$ such that (i) $k_{I\backslash M}(\cdot)$ and $k_M(\cdot)$ maximise informed traders' expected utility:

$$\forall i \in I\backslash M,\ k_{I\backslash M}(x_i,p;\gamma_i) \in \arg\max_{k\in\mathbb{R}} \mathbb{E}[V((\theta-p)k;\gamma_i)|x_i,p], \tag{3.2}$$

$$\forall i \in M,\ k_M(x_i;\gamma_i) \in \arg\max_{k\in\mathbb{R}} \mathbb{E}[V((\theta-p)k;\gamma_i)|x_i]; \tag{3.3}$$

and (ii) the market-making sector sets the price $p = \mathbb{E}[\theta|L(\cdot)]$, where

$$L(p) = \int_{I\backslash M} k_{I\backslash M}(x_i,p;\gamma_i)\mathrm{d}i + \int_M k_M(x_i;\gamma_i)\mathrm{d}i + \varepsilon. \tag{3.4}$$

We then have the following lemma:

Lemma 1. The trading game has a unique linear Bayesian equilibrium, which is characterised by:

- The investment functions

$$k_{I\backslash M}(x_i,p;\gamma_i) = \frac{\alpha_\theta\rho\bar{\theta} + \alpha_x x_i - (\alpha_x + \alpha_\theta\rho)\,p}{\gamma_i\,(1+(1-\rho)\rho B^2\alpha_\varepsilon/\alpha_x)} \quad \text{and} \quad k_M(x_i;\gamma_i) = \frac{\beta(x_i-\bar{\theta})}{\gamma_i}, \tag{3.5}$$

with

$$\beta = \left[\left(\alpha_x^{-1}+\alpha_\theta^{-1}\right)\left(1+B^2\rho\alpha_x^{-1}\alpha_\varepsilon\right) - \left(B^2\alpha_\varepsilon\left(1+\rho\alpha_x^{-1}\alpha_\theta\right)+\alpha_\theta\right)^{-1}\right]^{-1}; \tag{3.6}$$

- The price function

$$p(\theta,\tilde{\xi},\varepsilon) = (1-\lambda B)\,\bar{\theta} + \lambda B(\theta + \alpha_x^{-1/2}\tilde{\xi} + B^{-1}\varepsilon), \tag{3.7}$$

with

$$\lambda = B\alpha_\varepsilon\left[B^2\alpha_\varepsilon\,(1+\rho\alpha_\theta/\alpha_x)+\alpha_\theta\right]^{-1}. \tag{3.8}$$

In those functions, $B > 0$ is the unique real solution to the cubic equation:

$$B = \frac{\alpha_x \nu \gamma_{I\backslash M}^{-1}}{1 + (1-\rho)\rho B^2 \alpha_\varepsilon/\alpha_x} + \frac{(1-\nu)\,\gamma_M^{-1}}{\left(\alpha_x^{-1} + \alpha_\theta^{-1}\right)(1 + B^2\rho\alpha_\varepsilon/\alpha_x) - \left[B^2\alpha_\varepsilon\left(1 + \rho\alpha_\theta/\alpha_x\right) + \alpha_\theta\right]^{-1}}, \quad (3.9)$$

where $\gamma_{I\backslash M}^{-1} \equiv \nu^{-1}\int_{I\backslash M}\gamma_i^{-1}\mathrm{d}i$ and $\gamma_M^{-1} \equiv (1-\nu)^{-1}\int_M \gamma_i^{-1}\mathrm{d}i$ are the average risk tolerance coefficients of demand-schedule and market-order traders, respectively.

Lemma 1 generalises the trading game in Vives [15, Sec. 4.3] in two directions: i) heterogenous risk aversion, and ii) correlated noise in the private signals. Heterogeneity in risk aversion is the dimension along which informed traders sort themselves into demand-schedule versus market-order traders in Section 4. The possibility that private informational noise be correlated will imply that our results do not depend on whether the signal x_i is private ($\rho < 1$) or public ($\rho = 1$).

Equation (4.8) implies that observing p is equivalent to observing $\theta + \alpha_x^{-1/2}\tilde{\xi} + B^{-1}\varepsilon$. Hence p provides a public signal that is jointly informative of θ, $\tilde{\xi}$ and ε. Note that the signal extraction problem faced by demand-schedule traders is more involved when $\rho \in (0,1)$ than when $\rho \in \{0,1\}$. When $\rho = 0$ we have $\tilde{\xi} = 0$ a.s., hence p provides a signal about θ with noise $B^{-1}\varepsilon$. When $\rho = 1$, the public signal $\theta + \alpha_x^{-1/2}\tilde{\xi}$ is observed (i.e., $\eta_i = 0$ $\forall i$) jointly with p, hence $B^{-1}\varepsilon$ can be perfectly inferred (see (4.8) again). In contrast, when $\rho \in (0,1)$ then $\tilde{\xi}$ must be inferred together with θ from the observation of p. This joint signal extraction problem manifests itself by a greater residual uncertainty about θ (conditional on a given signal x_i) when $\rho \in (0,1)$ than when $\rho \in \{0,1\}$, which lowers the responsiveness of the demand for assets by demand-schedule traders to their signal.[5]

3.3 Exogenous trader types

We first analyse price informativeness at high signal precision when the distribution of informed traders across types is exogenous. We then have the following proposition:

[5]This effect shows up in the fact that the multiplier $\alpha_x/\left[\gamma_i\left(1 + (1-\rho)\rho B^2\alpha_\varepsilon/\alpha_x\right)\right]$ in the investment function of demand-schedule traders (see (3.5)) has the term $\rho(1-\rho)$ in the denominator. This product is equal to zero for $\rho \in \{0,1\}$ but is positive for $\rho \in (0,1)$ and is maximal at $\rho = 1/2$.

Proposition 1. (a) In a pure market-order market ($\nu = 0$), the informativeness of the price signal is bounded above; formally, $\alpha_p \underset{\alpha_x \to \infty}{\to} B_0^2 \alpha_\varepsilon < +\infty$, where $B_0 > 0$ uniquely solves $\gamma_M B_0 \left(\alpha_\theta^{-1} - \left(\alpha_\theta + \alpha_\varepsilon B_0^2\right)^{-1}\right) = 1$. (b) Whenever there is a positive mass of demand-schedule traders ($\nu > 0$), then the precision of the price signal is unbounded as $\alpha_x \to \infty$; more specifically,

$$\alpha_p \underset{\alpha_x \to +\infty}{\sim} \mathbf{1}_{\rho>0} \left(\alpha_x/\rho\right) + \mathbf{1}_{\rho=0} \left(\nu \alpha_x / \gamma_{I \setminus M}\right)^2 \alpha_\varepsilon$$

Proposition 1 shows that the speed of information aggregation as $\alpha_x \to +\infty$ depends on both the cross-correlation of informed traders' signals and the share of market-order traders among them. First, whenever $\nu > 0$, then information aggregation is less effective when $\rho > 0$ than when $\rho = 0$ (the informativeness of the price α_p grows at the rate of α_x in the former case but at the rate of α_x^2 in the latter). Second, information aggregation is less effective when $\nu = 0$ than otherwise (the informativeness of p is bounded above as $\alpha_x \to \infty$ in the former case, not in the latter).

The intuition for the second result follows from our assumed information structure and its implications for the Bayesian updating problem of demand-schedule traders. As stressed above, this problem is more involved when $\rho > 0$ than when $\rho = 0$: in the latter case the quality of the price signal p is only blurred by the extent of noise trading ε, while in the former it is also blurred by the common informational noise component $\tilde{\xi}$. As α_x increases and private signals become more and more aligned with θ, the impact of ε on p diminishes (since α_ε is constant) but that of $\tilde{\xi}$ does not (since its precision $1/\mathbb{V}(\alpha_x^{-1/2}\tilde{\xi}) = \alpha_x/\rho$ increases at the rate of α_x). In the special case where $\rho = 0$ the informativeness of p grows very rapidly (at the rate of α_x^2) because the common noise component effect is absent; whenever it is present, the quality of the price signal *cannot* grow at a rate faster than α_x. This suggests that the usual specification where $\rho = 0$ is somewhat special and that the conclusions drawn from it are not necessarily robust. Here it implies that as α_x grows large then x_i (whose precision grows at rate α_x) looses value relative to p (whose precision grows at the rate of α_x^2); eventually, demand-schedule traders only base their Bayesian estimate of θ on p. In contrast, when $\rho > 0$ and the informativeness of p grows at the same rate as α_x, then x_i and p keep constant weights in the computation of the posterior mean of θ as $\alpha_x \to +\infty$.

The intuition for the first result in the proposition (i.e., that price informativeness is bounded when $\nu = 0$) is as follows. In a pure *demand-schedule* market ($\nu = 1$), informed traders can condition their trades on p, so the only source of risk they face concerns the true value of θ. As the precision of private signals increases, informed traders collectively trade more and more aggressively against any discrepancy between p and θ. Formally, from (3.1) and Lemma 1 the total asset demand by informed traders in a pure demand-schedule market can be written as:

$$\int_I \left(\frac{\alpha_\theta \rho \bar{\theta} + \alpha_x (\theta + \alpha_x^{-1/2} \tilde{\xi} + \alpha_x^{-1/2} \eta_i) - (\alpha_x + \alpha_\theta \rho) p}{\gamma_i \left(1 + (1-\rho)\rho B^2 \alpha_\varepsilon / \alpha_x\right)} \right) \mathrm{d}i = B \left(\theta - p + \frac{\alpha_\theta \rho}{\alpha_x} \left(\bar{\theta} - p\right) + \alpha_x^{-1/2} \tilde{\xi} \right),$$

where B uniquely solves $B\gamma_I \left(1 + (1-\rho)\rho B^2 \alpha_\varepsilon / \alpha_x\right) = \alpha_x$ (since $I \backslash M = I$ and hence $\nu = 1$). The latter expression implies that $B \to +\infty$ as $\alpha_x \to +\infty$, and thus, by equations (4.8)–(3.8), that $p \to \theta$ as $\alpha_x \to +\infty$ –i.e., in the limit p becomes perfectly informative of θ. In contrast, in a pure *market-order* market ($\nu = 0$) informed traders do *not* condition their trades on p and hence face a residual payoff risk even as the x_is get more and more informative of θ. This risk leads market-order traders to trade relatively less aggressively on the basis of their private signal, which limits the amount of information that is transmitted into the price. Formally, from Lemma 1 again the total asset demand by informed traders in a pure market-order market is:

$$\int_I \left(\frac{\beta(\theta + \alpha_x^{-1/2} \tilde{\xi} + \alpha_x^{-1/2} \eta_i - \bar{\theta})}{\gamma_i} \right) \mathrm{d}i = \frac{\beta(\theta + \alpha_x^{-1/2} \tilde{\xi} - \bar{\theta})}{\gamma_M},$$

where β is given by (3.6) B solves (since $\nu = 0$):

$$\left(\alpha_x^{-1} + \alpha_\theta^{-1}\right) \left(1 + B^2 \rho \alpha_\varepsilon / \alpha_x\right) - \left(B^2 \alpha_\varepsilon \left(1 + \rho \alpha_\theta / \alpha_x\right) + \alpha_\theta\right)^{-1} = 1/(\gamma_M B)$$

In this situation, as $\alpha_x \to +\infty$ we have $B \to B_0$ (> 0), where B_0 is the unique solution to $\left(\alpha_\theta^{-1} - 1/\left(B^2 \alpha_\varepsilon + \alpha_\theta\right)\right) \gamma_M B = 1$. By implication, the trading intensity of market-order traders is bounded above as $\alpha_x \to +\infty$. From (4.8), as $\nu = 0$ and $\alpha_x \to +\infty$ we have:

$$p \to \frac{\alpha_\theta \bar{\theta}/\alpha_\varepsilon + B_0^2 \left(\theta + B_0^{-1}\varepsilon\right)}{B_0^2 + \alpha_\theta/\alpha_\varepsilon}.$$

Thus, asymptotically observing the price is equivalent to observing $\theta + B_0^{-1}\varepsilon$, i.e., p provides a signal about θ with precision $B_0^2\alpha_\varepsilon < +\infty$. The intermediate case $\nu \in (0,1)$ retains the main properties of the pure demand-schedule case, because any positive measure of demand-schedule traders is sufficient for their trading aggressiveness (which is unbounded as $\alpha_x \to +\infty$) to eliminate the impact of ε on p. As we show next, this intermediate case is that which necessarily arises in equilibrium when traders are free to choose their order type and the quality of information is high.

3.4 Endogenous trader types

We now analyse traders' choice of order type and determine the equilibrium sets M and $I\backslash M$. The basic tradeoff is that a demand schedule isolates a trader from price risk, but requires the trader to place a large (infinite) number of limit orders to generate complete conditionality of trades on the execution price. We capture this tradeoff by normalising the cost of a market order to zero and setting that of a full demand schedule to $c > 0$. We know from the CARA-Normal model that the value function associated with the information set G_i is:

$$W\left(G_i;\gamma_i\right) \equiv \max_k \mathbb{E}[V(w_i - \kappa c)|G_i;\gamma_i] = -\exp\left(-\frac{\mathbb{E}[\theta - p|G_i]^2}{2\mathbb{V}[\theta - p|G_i]} + \kappa c \gamma_i\right),$$

where $\kappa = 1$ if $G_i = (x_i, p)$ (i.e., the trader places a full demand schedule) or $\kappa = 0$ if $G_i = x_i$ (i.e., the trader places a market order). Let $W_{I\backslash M}\left(x_i, p;\gamma_i\right)$ and $W_M\left(x_i;\gamma_i\right)$ denote the expected utilities of a demand-schedule and a market order trader, respectively, with preferences γ_i and conditional on their full information set (i.e., x_i or (x_i, p)). There are two possible timing assumptions here, depending on whether we allow informed traders to choose their order type after ("timing 1") or before ("timing 2") observing x_i. Under timing 1 traders compare expected utilities *conditional on* x_i (i.e., $W_M(x_i;\gamma_i)$ and $\mathbb{E}\left[W_{I\setminus M}(x_i,p;\gamma_i)\middle|x_i\right]$), while under timing 2 they compare the same expected utilities *integrated over* x_i (i.e., $\mathbb{E}[W_M(x_i;\gamma_i)]$ and $\mathbb{E}\left[W_{I\setminus M}(x_i,p;\gamma_i)\right]$).[6] The following Lemma shows that the expected utility ratios are the same under the two timing assumptions, hence both lead to the same discrete choice of order type.

Lemma 2. The ratios of conditional and unconditional expected utilities are given by:

$$\underbrace{\frac{W_M(x_i;\gamma_i)}{\mathbb{E}\left[W_{I\setminus M}(x_i,p;\gamma_i)\middle|x_i\right]}}_{\text{timing 1}} = \underbrace{\frac{\mathbb{E}[W_M(x_i;\gamma_i)]}{\mathbb{E}\left[W_{I\setminus M}(x_i,p;\gamma_i)\right]}}_{\text{timing 2}} = \underbrace{\mathrm{e}^{-c\gamma_i}\sqrt{\frac{\mathbb{V}[\theta-p|x_i]}{\mathbb{V}[\theta-p|x_i,p]}}}_{\text{common value of the ratio}},$$

where $\mathbb{V}[\theta-p|x_i,p]$ and $\mathbb{V}[\theta-p|x_i]$ are given by equations (A2) and (A4) in Appendix A.

In other words, when we move from timing 1 to timing 2, both expected utilities change but in the same proportion, leaving the basic tradeoff between order types unchanged. It follows that under either timing informed trader i places a demand schedule if and only if the relevant ratio is below or equal to one. Given the value of $\mathbb{V}[\theta-p|x_i]/\mathbb{V}[\theta-p|x_i,p]$, computed from equations (A2) and (A4) in the appendix, this is equivalent to:

$$\begin{aligned}\gamma_i \le \bar{\gamma} = \frac{1}{2c}\Bigg[&\ln\left(\frac{1}{\alpha_x+\alpha_\theta}\left(\rho+\frac{\rho\alpha_\theta}{\alpha_x}+\frac{\alpha_\theta}{B^2\alpha_\varepsilon}\right)^2+\frac{\rho(1-\rho)}{\alpha_x}+\frac{1}{B^2\alpha_\varepsilon}\right)\\ &-\ln\left(\frac{1}{\alpha_\theta}\left(1-\left(1+\frac{\rho\alpha_\theta}{\alpha_x}+\frac{(1-\rho)\,\alpha_\theta}{\alpha_x+(1-\rho)\,B^2\alpha_\varepsilon}\right)^{-1}\right)\left(1+\frac{\rho\alpha_\theta}{\alpha_x}+\frac{\alpha_\theta}{B^2\alpha_\varepsilon}\right)^2\right)\Bigg],\end{aligned} \tag{3.10}$$

[6] The equilibria that we focus on under timing 1 are linear Bayesian equilibria with linear price functionals. In these equilibria informed traders choose their type on the basis of their risk aversion only (and not, say, on the level of their signal). Consequently, (i) the distribution of signals remains independent of that of risk aversion within each set M and $I\setminus M$, even though these are determined after the signals are observed; and hence (ii) the equilibrium measure ν is uninformative of the dividend. Note that even in this timing $W_{I\setminus M}(x_i,p;\gamma_i)$ is not known because it is a function of p, a random variable at the time the order type is chosen. In contrast $W_M(x_i;\gamma_i)$ is known, since it is *not* conditional on the yet unknown value of p (by the mere definition of a market order). This is why traders must compute $\mathbb{E}\left[W_{I\setminus M}(x_i,p;\gamma_i)\middle|x_i\right]$ and compare it with $W_M(x_i;\gamma_i)$.

where, from Lemma 1,

$$B = \left(\frac{\alpha_x}{(1+(1-\rho)\rho B^2\alpha_\varepsilon/\alpha_x)}\right)\int_0^{\gamma^{-1}(\bar\gamma)}\gamma_i^{-1}\mathrm{d}i$$
$$+\left(\left(\frac{1}{\alpha_x}+\frac{1}{\alpha_\theta}\right)\left(1+\frac{B^2\rho\alpha_\varepsilon}{\alpha_x}\right)-\frac{1}{B^2\alpha_\varepsilon\left(1+\rho\alpha_\theta/\alpha_x\right)+\alpha_\theta}\right)^{-1}\int_{\gamma^{-1}(\bar\gamma)}^{1}\gamma_i^{-1}\mathrm{d}i. \qquad (3.11)$$

with $\gamma^{-1}(\bar\gamma)=0$ if $\bar\gamma<\gamma(0)$ and $\gamma^{-1}(\bar\gamma)=1$ if $\bar\gamma>\gamma(1)$. For $(\alpha_x,\alpha_\theta,\alpha_\varepsilon,\rho)\in\mathbb{R}^{+3}\times[0,1]$ given, the properties of the γ function imply that the solution $(\bar\gamma,B)$ to (4.15)–(4.16), if it exists, can be of three types: it is either such that $\bar\gamma\in[\gamma(0),\gamma(1)]$, in which case the solution is interior (i.e., $M\neq\varnothing$ and $I\backslash M\neq\varnothing$); or $\bar\gamma<\gamma(0)$, so that the solution is corner and all traders place market orders (i.e., $(M,I\backslash M)=(I,\varnothing)$); or $\bar\gamma>\gamma(1)$ and all traders place demand schedules (i.e., $(M,I\backslash M)=(\varnothing,I)$). The intuition for the sorting of traders along the degree of risk aversion is that greater risk aversion lowers trading aggressiveness, hence the expected benefit from expanding the information set from x_i to (x_i,p). Proposition 2 states our main results under endogenous order types:[7]

Proposition 2. For any $(\alpha_\theta,\alpha_\varepsilon,\rho)\in\mathbb{R}^{+2}\times[0,1]$, and as $\alpha_x\to+\infty$, (a) the solution $(\bar\gamma,B)$ to (4.15)–(4.16) is unique; (b) both M and $I\backslash M$ have strictly positive measure (i.e., the equilibrium is interior); (c) $\bar\gamma\to\gamma_0$ (i.e., market-order traders eventually overwhelm the market); (d) α_p goes to infinity as the same rate as α_x; formally, defining the bijection $h_\rho:\mathbb{R}^+\to(0,1)$, $h_\rho(x)=\left(\rho+\frac{(1-\rho)}{1+(1-\rho)x}\right)\left(\rho+x^{-1}\right)^{-1}$ and h_ρ^{-1} its inverse, we have

$$\alpha_p \underset{\alpha_x\to\infty}{\sim} \frac{\alpha_x}{\rho+\left[h_\rho^{-1}\left(\mathrm{e}^{-2\gamma_0 c}\right)\right]^{-1}}.$$

Our information structure gives us some intuition about why demand-schedule traders vanish as $\alpha_x\to+\infty$ (point (c)). In our analysis p plays two distinct informational roles: it provides information about the distribution of signals received by the other informed traders, and about the net asset demand of noise traders. In the special case where

[7] Note from (4.15) that heterogeneity in c is formally equivalent to heterogeneity in γ. To encompass both cases, rank traders in nondecreasing orders of $g(i)\equiv c(i)\gamma(i)$, assume that $g(i)$ is continuous, strictly increasing, that its reciprocal is continuous, and that $0<g(0)<g(1)<+\infty$; then solve for the marginal trader exactly in the same way as in the case where c is homogenous.

information about θ is entirely public (i.e., $\rho = 1$), there is nothing to learn from the other informed traders by observing p. However, trading conditional on p is still valuable because this provides insurance against noise trader risk–to which market-order traders are exposed. That the crowding out of demand-schedule traders by market-order traders as $\alpha_x \to +\infty$ also takes place when $\rho = 1$ suggests that when the quality of information is high the primary value of a demand schedule relative to a market order comes from its hedging role against noise trader risk, rather than its role at extracting dispersed information. It follows that for the share of demand-schedule traders to vanish when signals become increasingly precise, it must be that the gain from hedging noise trader risk itself vanishes. But the reason for this is immediate: as the precision of information increases, demand-schedule traders trade more and more aggressively on their information. In so doing, they take p closer and closer to their own signal x_i, which is itself closer and closer to θ. Eventually, they completely eliminate noise trader risk, and thereby the relative benefit of a demand schedule.

This feature also explains why the equilibrium must necessarily be interior, i.e., why $I \backslash M$, whilst asymptotically vanishing, must always keep positive measure (point (b)). If it were not the case, then the market would be a full market-order market similar to that examined in Section 3. In this situation, the trading intensity of informed traders would be bounded above, hence the uncertainty about the dividend would be bounded below (see Proposition 1). But then noise trader risk would no longer be eliminated even at high levels of precision of the signals, and thus knowing the price (in addition to the signal about the dividend) to insure against noise trader risk would become valuable again. Demand schedules thus display a form of substituability: when information about θ is precise, then a positive but small measure of demand-schedule traders deters all the other traders from placing a demand schedule (however small c is).

Finally, the informativeness of p (point (d)) is closely related to the composition of the market (point (c)). As discussed in Section 3, market-order traders tend to reduce information aggregation. Consequently, the gradual crowding out of demand-schedule traders as $\alpha_x \to +\infty$ tends to reduce the pace of information aggregation, relative to the case with constant, exogenous shares of each type. For example, in the case where $\rho = 0$ the precision of the price signal grows at the rate of α_x, instead of α_x^2 when the sets M and $I \backslash M$ are exogenous.

3.5 Concluding remarks

This paper has analysed the joint determination of price informativeness and the composition of the market in a large market with dispersed information. By allowing market-order and demand-schedule traders to coexist, and by letting traders choose their preferred order type, the microstructure considered here is richer and more realistic than the pure demand schedule/Walrasian auctioneer specification. Our main result that the set of demand-schedule traders vanishes when signals become highly informative follows logically from the structure of the trading game, so we expect it to hold under more general assumptions than those we have assumed. For example, we have adopted the usual CARA-Normal framework, which is the only tractable specification under our information structure. However, nothing in our results seem to depend on a particular feature of preferences, at least in an obvious way; we thus conjecture that they would remain valid under much more general (risk averse) preferences. Similarly, while our information structure allows for the presence of both idiosyncratic and common informational noise components, it remains restrictive in the sense that both components are constrained to vanish at the same rate when the precision of private information goes to infinity (since the relation between the two is parameterised by the correlation coefficient ρ). We show formally in the following technical appendix that our results can be generalised to an information structure allowing each noise component to vanish at different rates.

3.6 Appendix

A. Proof of Lemma 1

There are three aggregate shocks $(\theta,\, \tilde{\xi},\, \varepsilon)$, hence three random variables that may affect p. Equation (3.1), implies that the effects of θ and $\alpha_x^{-1/2}\tilde{\xi}$ on private signals are indistinguishable. Hence we define $\tilde{\theta} \equiv \theta + \alpha_x^{-1/2}\tilde{\xi}$ and restrict our attention to equilibrium price functions $p(\tilde{\theta},\varepsilon)$ that are linear in $(\tilde{\theta},\varepsilon)$ (so that p is normally distributed). A trader i with risk aversion coefficient γ_i and information set G_i has a demand for assets $k_i(G_i) = \gamma_i^{-1}\mathbb{E}[\theta - p|G_i]/\mathbb{V}[\theta - p|G_i]$. We may thus write the demands by demand-schedule and market-order traders as $k^i_{I\backslash M}(x_i,p) = \gamma_i^{-1} f_{I\backslash M}(x_i,p)$ and $k^i_M(x_i) = \gamma_i^{-1} f_M(x_i)$, respectively, with

$$f_{I\backslash M}(x_i,p) = \mathbb{E}[\theta - p|x_i,p]/\mathbb{V}[\theta|x_i,p], \quad f_M(x_i) = \mathbb{E}[\theta - p|x_i]/\mathbb{V}[\theta - p|x_i].$$

A.1. Price function

We conjecture that $f_{I\backslash M}$, f_M have the form $f_{I\backslash M}(x_i,p) = a(x_i - \bar{\theta}) + \zeta(p)$ and $f_M(x_i) = c(x_i - \overline{\theta})$, where a, c are normalised trading intensities (for a trader with $\gamma_i = 1$) and $\zeta(\cdot)$ is a linear function. Using the convention that the average signal equals $\tilde{\theta}$ a.s., and recalling that γ_i is independent from $\xi_i - \tilde{\xi}$, the order book is given by

$$\begin{aligned} L(p) &= \int_{I\backslash M} k^i_{I\backslash M}(x_i,p)\mathrm{d}i + \int_M k^i_M(x_i)\mathrm{d}i + \varepsilon = \int_{I\backslash M} \frac{a(x_i - \bar{\theta}) + \zeta(p)}{\gamma_i}\mathrm{d}i + \int_M \frac{c(x_i - \bar{\theta})}{\gamma_i}\mathrm{d}i + \varepsilon \\ &= \left[a(\theta + \alpha_x^{-1/2}\tilde{\xi} - \bar{\theta})\right]\int_{I\backslash M}\gamma_i^{-1}\mathrm{d}i + c(\theta + \alpha_x^{-1/2}\tilde{\xi} - \bar{\theta})\int_M \gamma_i^{-1}\mathrm{d}i + \varepsilon + \zeta(p)\int_{I\backslash M}\gamma_i^{-1}\mathrm{d}i \\ &= B(\theta + \alpha_x^{-1/2}\tilde{\xi} + B^{-1}\varepsilon) - B\bar{\theta} + \gamma^{-1}_{I\backslash M}\nu\zeta(p), \quad \text{with} \quad B = a\nu\gamma^{-1}_{I\backslash M} + c\,(1-\nu)\,\gamma_M^{-1}. \end{aligned}$$

The market making sector observes $L(.)$, a linear function of p, and sets $p = \mathbb{E}[\theta|L(.)] = \mathbb{E}[\theta|z]$, where $z = \theta + \alpha_x^{-1/2}\tilde{\xi} + B^{-1}\varepsilon$ summarises the information provided by the order book. Since z provides a signal about θ with noise $\tilde{\varepsilon} \equiv \alpha_x^{-1/2}\tilde{\xi} + B^{-1}\varepsilon$ we have:

$$p = \mathbb{E}[\theta|z] = \frac{\alpha_\theta\bar{\theta} + \alpha_{\tilde{\varepsilon}} z}{\alpha_\theta + \alpha_{\tilde{\varepsilon}}} \sim \mathcal{N}\left(\bar{\theta}, \mathbb{V}\left(p\right)\right), \text{ where } \alpha_{\tilde{\varepsilon}} = \frac{1}{\mathbb{V}\left(\tilde{\varepsilon}\right)} = \frac{B^2\alpha_\varepsilon}{1 + \rho B^2\alpha_\varepsilon/\alpha_x}.$$

Rearranging the latter expression gives the price function (4.8) in Lemma 1.

A.2. Investment functions for $\rho \in (0,1)$

We now need to identify the parameters a and c in the demand functions, which requires computing the conditional moments of $\theta - p|G_i$, for $G_i = (x_i, p)$ (demand schedules) or $G_i = x_i$ (market orders). We start with the former and specifically focus on the moments of $\theta|x_i, p$ which is without loss of generality. To this purpose define $\mathbf{\Theta} \equiv \begin{bmatrix} \theta & \tilde{\theta} & p & x_i \end{bmatrix}' \sim \mathcal{N}_4 \left(\mathbb{E}\left(\mathbf{\Theta}\right), \mathbb{V}\left(\mathbf{\Theta}\right)\right)$. From (3.1), (4.8) and the fact that $\tilde{\theta} = \theta + \alpha_x^{-1/2}\tilde{\xi}$, we have, for $\rho \in (0,1)$, $\mathbf{\Theta} = \begin{bmatrix} \bar{\theta} & \bar{\theta} & \bar{\theta} & \bar{\theta} \end{bmatrix}' + \mathbf{MS}$, with

$$\mathbf{M} = \begin{bmatrix} \alpha_\theta^{-1/2} & 0 & 0 & 0 \\ \alpha_\theta^{-1/2} & \rho^{1/2}\alpha_x^{-1/2} & 0 & 0 \\ \lambda B \alpha_\theta^{-1/2} & \lambda B \rho^{1/2}\alpha_x^{-1/2} & \lambda \alpha_\varepsilon^{-1/2} & 0 \\ \alpha_\theta^{-1/2} & \rho^{1/2}\alpha_x^{-1/2} & 0 & (1-\rho)^{1/2}\alpha_x^{-1/2} \end{bmatrix}, \ \mathbf{S} = \begin{bmatrix} \alpha_\theta^{1/2}\left(\theta - \bar{\theta}\right) \\ \rho^{-1/2}\tilde{\xi} \\ \alpha_\varepsilon^{1/2}\varepsilon \\ (1-\rho)^{-1/2}\eta_i \end{bmatrix} \sim \mathcal{N}_4\left(\mathbf{0}, \mathbf{I}\right).$$

Next, we compute $\mathbb{V}\left(\mathbf{\Theta}\right) = \mathbf{MM}'$ and then partition $\mathbb{V}\left(\mathbf{\Theta}\right)$ as $\mathbb{V}\left(\mathbf{\Theta}\right) = \left[\Sigma_{km}\right], k = 1, 2$, with all four Σ_{km} being 2×2 matrices. Then, from standard multivariate normal theory we know that $\begin{bmatrix} \theta & \tilde{\theta} \end{bmatrix}'|p, x_i$ has distribution $\mathcal{N}_2\left(\begin{bmatrix} \bar{\theta} & \bar{\theta} \end{bmatrix}' + \Sigma_{12}\Sigma_{22}^{-1}\begin{bmatrix} p - \bar{\theta} & x_i - \bar{\theta} \end{bmatrix}', \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\right)$, from which we infer the following conditional moments:

$$\mathbb{E}\left[\theta|p, x_i\right] = \bar{\theta} + \frac{\alpha_x\left(x_i - \bar{\theta}\right) + (1-\rho)\, B^2\alpha_\varepsilon\left(\lambda B\right)^{-1}\left(p - \bar{\theta}\right)}{(1-\rho)\left(1 + \rho\alpha_\theta/\alpha_x\right) B^2\alpha_\varepsilon + \alpha_x + \alpha_\theta}, \tag{A1}$$

$$\mathbb{V}[\theta|p, x_i] = \frac{1 + (1-\rho)\,\rho B^2\alpha_\varepsilon/\alpha_x}{(1-\rho)\left(1 + \rho\alpha_\theta/\alpha_x\right) B^2\alpha_\varepsilon + \alpha_x + \alpha_\theta}. \tag{A2}$$

Substituting these values into $k^i_{I\backslash M}(x_i, p) = \gamma_i^{-1}\mathbb{E}[\theta - p|x_i, p]/\mathbb{V}[\theta|x_i, p]$ and rearranging gives the corresponding asset demand in Lemma 1.

We now turn to the computation of $\theta - p|x_i$. We define, still for $\rho \in (0,1)$, $\mathbf{\Omega} \equiv [\begin{array}{cccc} \theta & \tilde{\theta} & \theta - p & x_i \end{array}]'$ $\sim \mathcal{N}_4\left(\mathbb{E}\left(\mathbf{\Omega}\right), \mathbb{V}\left(\mathbf{\Omega}\right)\right)$ and note that $\mathbf{\Omega} = [\begin{array}{cccc} \bar{\theta} & \bar{\theta} & 0 & \bar{\theta} \end{array}]' + \mathbf{NT}$, where

$$\mathbf{N} = \left[\begin{array}{cccc} \alpha_\theta^{-1/2} & 0 & 0 & 0 \\ \alpha_\theta^{-1/2} & \rho^{1/2}\alpha_x^{-1/2} & 0 & 0 \\ (1-\lambda B)\alpha_\theta^{-1/2} & -\lambda B \rho^{1/2}\alpha_x^{-1/2} & -\lambda \alpha_\varepsilon^{-1/2} & 0 \\ \alpha_\theta^{-1/2} & \rho^{1/2}\alpha_x^{-1/2} & 0 & (1-\rho)^{1/2}\alpha_x^{-1/2} \end{array}\right], \ \mathbf{T} = \left[\begin{array}{c} \alpha_\theta^{1/2}\left(\theta - \bar{\theta}\right) \\ \rho^{-1/2}\tilde{\xi} \\ \alpha_\varepsilon^{1/2}\varepsilon \\ (1-\rho)^{-1/2}\eta_i \end{array}\right],$$

so that $\mathbf{T} \sim \mathcal{N}_4\left(\mathbf{0}, \mathbf{I}\right)$. We partition $\mathbb{V}\left(\mathbf{\Omega}\right) = \mathbf{NN}'$ as follows: $\mathbb{V}\left(\mathbf{\Omega}\right) = \left[\bar{\Sigma}_{km}\right], k, m = 1, 2,$ where $\bar{\Sigma}_{11}$ is 3×3, $\bar{\Sigma}_{12}$ is 3×1, $\bar{\Sigma}_{21}$ is 1×3 and $\bar{\Sigma}_{22}$ is 1×1. It follows that $[\begin{array}{ccc} \theta & \tilde{\theta} & \theta - p \end{array}]'|x_i$ has distribution $\mathcal{N}_3\left([\begin{array}{ccc} \bar{\theta} & \bar{\theta} & 0 \end{array}]' + \bar{\Sigma}_{12}\bar{\Sigma}_{22}^{-1}\left(x_i - \bar{\theta}\right), \bar{\Sigma}_{11} - \bar{\Sigma}_{12}\bar{\Sigma}_{22}^{-1}\bar{\Sigma}_{21}\right)$. After some calculations, we infer from this joint distribution that:

$$\mathbb{E}\left[\theta - p|x_i\right] = \frac{\left(\left(1-\lambda B\right)\alpha_\theta^{-1} - \lambda B \rho \alpha_x^{-1}\right)\left(x_i - \bar{\theta}\right)}{\alpha_\theta^{-1} + \alpha_x^{-1}}, \tag{A3}$$

$$\mathbb{V}[\theta - p|x_i] = (\lambda B)^2\left(\frac{\left(\rho + \rho\alpha_\theta/\alpha_x + B^{-2}\alpha_\theta/\alpha_\varepsilon\right)^2}{\alpha_x + \alpha_\theta} + \frac{(1-\rho)\rho}{\alpha_x} + \frac{1}{B^2\alpha_\varepsilon}\right). \tag{A4}$$

Substituting (A3)-(A4) into $k_M^i(x_i) = \gamma_i^{-1}\mathbb{E}[\theta - p|x_i]/\mathbb{V}[\theta - p|x_i]$ and rearranging gives the asset demand of market-order traders in Lemma 1.

A.3. Investment functions for $\rho \in \{0,1\}$

The expressions for $f_{I\setminus M}(x_i, p)$, $f_M(x_i)$, which have been derived for $\rho \in (0,1)$, can be extended by continuity to $\rho \in \{0,1\}$. For example, for $\rho = 0$ we have $z = \theta + B^{-1}\varepsilon$, and computing the joint distribution of (p, x_i, θ) gives the same conditional moments as those in (A1)–(A4) when setting $\rho = 0$. Similarly, when $\rho = 1$ all informed traders receive the same signal $x = \theta + \alpha_x^{-1/2}\tilde{\xi}$, $\tilde{\xi} \sim \mathcal{N}(0,1)$, hence observing p does not provide any more information about θ than observing x. It follows that:

$$\mathbb{E}[\theta|p, x] = \mathbb{E}[\theta|x] = \left(\alpha_\theta\bar{\theta} + \alpha_x x\right)/\left(\alpha_\theta + \alpha_x\right), \ \mathbb{V}[\theta|p, x] = \mathbb{V}[\theta|x] = \left(\alpha_x + \alpha_\theta\right)^{-1},$$

which is recovered by setting $\rho = 1$ in (A1)–(A4). Hence the expressions for $k_{I\setminus M}(x_i, p; \gamma_i)$, $k_M(x_i; \gamma_i)$ in Lemma 1 are valid for $\rho \in (0,1) \cup \{0,1\} = [0,1]$.

A.4. Uniqueness of linear equilibrium

We finally show that B is unique, positive and finite. To do this, define the function $f : \mathbb{R} \to \mathbb{R}$ as

$$f(B) = \frac{\nu\alpha_x}{\gamma_{I\setminus M}\left(1 + \frac{(1-\rho)\rho B^2\alpha_\varepsilon}{\alpha_x}\right)} + \frac{1-\nu}{\gamma_M\left(\left(\alpha_x^{-1} + \alpha_\theta^{-1}\right)\left(1 + B^2\frac{\rho\alpha_\varepsilon}{\alpha_x}\right) - \left(\alpha_\theta + B^2\alpha_\varepsilon\left(1 + \frac{\rho\alpha_\theta}{\alpha_x}\right)\right)^{-1}\right)} - B,$$

so that a root of $f(B)$ solves (4.9). f is continuous and strictly decreasing over $[0, +\infty)$ and such that $f(0) = \frac{\nu\alpha_x}{\gamma_{I\setminus M}} + (1-\nu)\frac{\alpha_x}{\gamma_M} > 0$ and $f(+\infty) = -\infty$. Hence f is a bijection with a unique root $B_0 > 0$ over $[0, +\infty)$. Since f is strictly positive on $\mathbb{R}^-$, B_0 is the unique root of f in $\mathbb{R}$.

B. Proof of Proposition 1

From (4.8), p is observationally equivalent to $z = \theta + \alpha_x^{-1/2}\tilde{\xi} + B^{-1}\varepsilon$, so both provide the same information about θ. It follows that the precision of the price signal is:

$$\alpha_p = \alpha_z = 1/\mathbb{V}(\alpha_x^{-1/2}\tilde{\xi} + B^{-1}\varepsilon) = 1/\left(\rho\alpha_x^{-1} + B^{-2}\alpha_\varepsilon^{-1}\right) \tag{B1}$$

B.1. Full market-order case

We know from Lemma 1 that $B > 0$ uniquely solves (4.9). Now define the function $g : \mathbb{R}_+^* \times \mathbb{R}_+^* \to \mathbb{R}$ as follows:

$$g(B, \alpha_x) = \gamma_I^{-1}\left[\left(\alpha_x^{-1} + \alpha_\theta^{-1}\right)\left(1 + B^2\rho\alpha_\varepsilon/\alpha_x\right) - \left(\alpha_\theta + B^2\alpha_\varepsilon\left(1 + \rho\alpha_\theta/\alpha_x\right)\right)^{-1}\right]^{-1} - B,$$

When $\nu = 0$, B is the unique solution to $g(B, \alpha_x) = 0$. Since g is continuously differentiable, increasing in a_x and decreasing in B on $\mathbb{R}_+^* \times \mathbb{R}_+^*$, the implicit function $B(\alpha_x)$ defined by $g(B, \alpha_x) = 0$ is continuously differentiable and increasing over $\mathbb{R}_+^*$. Moreover, we have:

$$\left(\alpha_x^{-1}+\alpha_\theta^{-1}\right)\left(1+B^2\rho\alpha_\varepsilon/\alpha_x\right)-\left(\alpha_\theta+B^2\alpha_\varepsilon\left(1+\rho\alpha_\theta/\alpha_x\right)\right)^{-1}$$
$$\geq \alpha_\theta^{-1}-\left(B^2\alpha_\varepsilon\left(1+\rho\alpha_\theta/\alpha_x\right)+\alpha_\theta\right)^{-1}\geq \alpha_\theta^{-1}-\left(B^2\alpha_\varepsilon+\alpha_\theta\right)^{-1},$$
$$\text{so that}\quad B=\gamma_I^{-1}\left[\left(\alpha_x^{-1}+\alpha_\theta^{-1}\right)\left(1+B^2\rho\alpha_\varepsilon/\alpha_x\right)-\left(\alpha_\theta+B^2\alpha_\varepsilon\left(1+\rho\alpha_\theta/\alpha_x\right)\right)^{-1}\right]^{-1}$$
$$\leq \gamma_I^{-1}\left[\alpha_\theta^{-1}-\left(B^2\alpha_\varepsilon+\alpha_\theta\right)^{-1}\right]^{-1}=\gamma_I^{-1}\alpha_\theta\left(B^2\alpha_\varepsilon+\alpha_\theta\right)\leq\gamma_I^{-1}\alpha_\theta\left(1+\alpha_\theta B^{-2}\alpha_\varepsilon^{-1}\right)$$

The function $h_+ : B \to B-\gamma_I^{-1}\alpha_\theta\left(1+\alpha_\theta/B^2\alpha_\varepsilon\right)$ is continuous and strictly increasing over $\mathbb{R}_+^*$, and such that $h_+(0)=-\infty$ and $h_+(+\infty)=+\infty$. It is thus bijective and we denote its inverse by h_+^{-1}. Then $B\in h_+^{-1}\left(]-\infty;0]\right)$ is bounded above by $h_+^{-1}(0)$, which is positive and independent of α_x. Hence, $B_0\equiv\lim_{\alpha_x\to\infty}B(\alpha_x)$ is defined and, by continuity, is the unique solution to $\gamma_I B=\left(\alpha_\theta^{-1}-\left(\alpha_\theta+\alpha_\varepsilon B^2\right)^{-1}\right)^{-1}$. From (4.8) we then infer that $\lim_{\alpha_x\to\infty}\alpha_p=B_0^2\alpha_\varepsilon$.

B.2. Other cases

When $\nu\in(0,1]$, $B(\alpha_x)$ is implicitly defined as the unique solution to (4.9). When $\rho\in\{0,1\}$, we have

$$\frac{B(\alpha_x)}{\frac{\nu\alpha_x}{\gamma_{I\backslash M}}}=1+\frac{(1-\nu)/\gamma_M}{\left(\alpha_x^{-1}+\alpha_\theta^{-1}\right)(1+B^2\rho\alpha_\varepsilon/\alpha_x)-\left(B(\alpha_x)^2\alpha_\varepsilon(1+\rho\alpha_\theta/\alpha_x)+\alpha_\theta\right)^{-1}}\frac{1}{\frac{\nu\alpha_x}{\gamma_{I\backslash M}}},$$

so that $B(\alpha_x)\geq\nu\alpha_x/\gamma_{I\backslash M}$. Hence, whenever $\rho\in\{0,1\}$ and $\nu>0$ we have $\lim_{\alpha_x\to+\infty}B(\alpha_x)=+\infty$ and $\lim_{\alpha_x\to+\infty}[B(\alpha_x)]^2/\alpha_x=+\infty$.

It follows that $\lim_{\alpha_x\to+\infty}B(\alpha_x)/\left(\frac{\nu\alpha_x}{\gamma_{I\backslash M}}\right)$ is equal to

$$1+\lim_{\alpha_x\to+\infty}\frac{(1-\nu)/\gamma_M}{\left(\alpha_x^{-1}+\alpha_\theta^{-1}\right)\left(1+B(\alpha_x)^2\rho\alpha_\varepsilon/\alpha_x\right)-\left(B(\alpha_x)^2\alpha_\varepsilon(1+\rho\alpha_\theta/\alpha_x)+\alpha_\theta\right)^{-1}}\frac{1}{\frac{\nu\alpha_x}{\gamma_{I\backslash M}}}=1,$$

so that $B(\alpha_x)\underset{\alpha_x\to+\infty}{\sim}\nu\alpha_x/\gamma_{I\backslash M}$. Now, recall from (4.8) that $\alpha_p=\left(\rho/\alpha_x+B^{-2}/\alpha_\varepsilon\right)^{-1}$. Hence, for $\rho=0$ we have $\lim_{\alpha_x\to+\infty}\alpha_p=\lim_{\alpha_x\to+\infty}[B(\alpha_x)]^2\alpha_\varepsilon=(\nu\alpha_x/\gamma_{I\backslash M})^2\alpha_\varepsilon$, that is $\alpha_p\underset{\alpha_x\to+\infty}{\sim}(\nu/\gamma_{I\backslash M})^2\alpha_\varepsilon\alpha_x^2$. However, for $\rho=1$ we have $\alpha_p\underset{\alpha_x\to+\infty}{\sim}\alpha_x$. Indeed, in this case we have $\alpha_p^{-1} = \alpha_x^{-1} + B^{-2}\alpha_\varepsilon^{-1}$, and we know that $\lim_{\alpha_x\to+\infty} B^{-2}\alpha_\varepsilon^{-1}/\alpha_x^{-1} = \alpha_\varepsilon^{-1}\lim_{\alpha_x\to+\infty}\alpha_x/B^2 = 0$, from which it follows that $\alpha_p^{-1} \underset{\alpha_x\to+\infty}{\sim} \alpha_x^{-1}$.

Let us now turn to the case where $\rho \in (0,1)$. In that situation (4.9) implies that

$$B(\alpha_x) \geq \frac{\nu\alpha_x}{\gamma_{I\setminus M}}\frac{1}{1+(1-\rho)\rho B(\alpha_x)^2\alpha_\varepsilon/\alpha_x} \Leftrightarrow \frac{B(\alpha_x)}{\alpha_x^{1/2}}\left(1+(1-\rho)\rho B(\alpha_x)^2\alpha_\varepsilon/\alpha_x\right) \geq \frac{\nu\alpha_x^{1/2}}{\gamma_{I\setminus M}},$$

so that $\lim_{\alpha_x\to\infty} B(\alpha_x)/\alpha_x^{1/2} = \lim_{\alpha_x\to\infty} B(\alpha_x)^2/\alpha_x = +\infty$. Moreover, again from (4.9) we have

$$\frac{B(\alpha_x)^3}{\alpha_x^2} = \frac{\nu/\gamma_{I\setminus M}}{\alpha_x/B(\alpha_x)^2+(1-\rho)\rho\alpha_\varepsilon} + \frac{[(1-\nu)/\gamma_M]\times B(\alpha_x)^2/\alpha_x^2}{\left(\frac{1}{\alpha_x}+\frac{1}{\alpha_\theta}\right)\left(1+\frac{B(\alpha_x)^2\rho\alpha_\varepsilon}{\alpha_x}\right)-\left(B(\alpha_x)^2\alpha_\varepsilon\left(1+\frac{\rho\alpha_\theta}{\alpha_x}\right)+\alpha_\theta\right)^{-1}}$$

Now, since

$$\frac{[(1-\nu)/\gamma_M]\times B(\alpha_x)^2/\alpha_x^2}{\left(\frac{1}{\alpha_x}+\frac{1}{\alpha_\theta}\right)\left(1+\frac{B(\alpha_x)^2\rho\alpha_\varepsilon}{\alpha_x}\right)-\left(B(\alpha_x)^2\alpha_\varepsilon\left(1+\frac{\rho\alpha_\theta}{\alpha_x}\right)+\alpha_\theta\right)^{-1}} \underset{\alpha_x\to\infty}{\to} 0 \quad\text{and}\quad \frac{\alpha_x}{B(\alpha_x)^2} \underset{\alpha_x\to\infty}{\to} 0,$$

we get

$$\frac{B^3(\alpha_x)}{\alpha_x^2} \underset{\alpha_x\to\infty}{\to} \frac{\nu/\gamma_{I\setminus M}}{(1-\rho)\rho\alpha_\varepsilon} \Rightarrow B(\alpha_x) \underset{\alpha_x\to\infty}{\sim} \left(\frac{\nu\alpha_x^2/\gamma_{I\setminus M}}{(1-\rho)\rho\alpha_\varepsilon}\right)^{1/3}.$$

Recall that $\alpha_p^{-1} = \rho\alpha_x^{-1} + B^{-2}\alpha_\varepsilon^{-1}$, and we have shown that $\lim_{\alpha_x\to\infty} B(\alpha_x)^2/\alpha_x = +\infty$. Hence, $\lim_{\alpha_x\to\infty}\alpha_x/B(\alpha_x)^2 = 0$, so that $\alpha_p^{-1} \underset{\alpha_x\to\infty}{\sim} \rho\alpha_x^{-1}$.

C. Proof of Lemma 2

Let us first state the version of the law of total variance that is relevant in our context:

$$\mathbb{V}[\mathbb{E}[\theta-p|x_i,p]|x_i] = \mathbb{V}[\theta-p|x_i] - \mathbb{E}[\mathbb{V}[\theta-p|x_i,p]|x_i] = \mathbb{V}[\theta-p|x_i] - \mathbb{V}[\theta-p|x_i,p], \tag{C1}$$

$$\text{and}\quad \mathbb{V}[\mathbb{E}[\theta-p|x_i]] = \mathbb{V}[\theta-p] - \mathbb{E}[\mathbb{V}[\theta-p|x_i]] = \mathbb{V}[\theta-p] - \mathbb{V}[\theta-p|x_i]. \tag{C2}$$

C.1. Timing 1: Order type chosen after the signal is observed

Under timing 1 traders observe x_i, know that the price will be a linear function of normally distributed variables, and then compare $W_M(x_i;\gamma_i)$ and $\mathbb{E}\left[W_{I\setminus M}(x_i,p;\gamma_i)\middle|\, x_i\right]$. We first write $W_M(\cdot)$ as

$$W_M(x_i;\gamma_i) = -\mathrm{e}^{-y_{M,i}^2}, \text{ with } y_{M,i} \equiv \frac{\mathbb{E}[\theta - p|x_i]}{\sqrt{2\mathbb{V}[\theta - p|x_i]}}. \tag{C3}$$

Similarly, $W_{I\setminus M}(x_i,p;\gamma_i) = -\mathrm{e}^{-y_{I\setminus M,i}^2 + c\gamma_i}$, with $y_{I\setminus M,i} \equiv \mathbb{E}[\theta - p|x_i,p]/\sqrt{2\mathbb{V}[\theta - p|x_i,p]}$ and

$$y_{I\setminus M,i}|x_i \sim \mathcal{N}\left(\frac{\mathbb{E}[\theta - p|x_i]}{\sqrt{2\mathbb{V}[\theta - p|x_i,p]}}, \frac{\mathbb{V}[\mathbb{E}[\theta - p|x_i,p]|x_i]}{2\mathbb{V}[\theta - p|x_i,p]}\right).$$

Using the fact that $y_{I\setminus M,i}^2|x_i$ has a noncentral *chi*-square distribution (see, e.g., Grossman and Stiglitz, 1980), the moment-generating function yields:

$$\begin{aligned}
&\mathbb{E}\left[W_{I\setminus M}(x_i,p;\gamma_i)\middle|\, x_i\right] \\
&= -\mathbb{E}[\mathrm{e}^{-y_{I\setminus M,i}^2 + c\gamma_i}|x_i] = -\frac{\mathrm{e}^{c\gamma_i}}{\sqrt{1 + 2\mathbb{V}\left[y_{I\setminus M,i}|x_i\right]}} \exp\left(-\frac{\left(\mathbb{E}\left[y_{I\setminus M,i}|x_i\right]\right)^2}{1 + 2\mathbb{V}\left[y_{I\setminus M,i}|x_i\right]}\right) \\
&= -\mathrm{e}^{c\gamma_i}\left(\frac{\mathbb{V}[\theta - p|x_i,p]}{\mathbb{V}[\theta - p|x_i,p] + \mathbb{V}[\mathbb{E}[\theta - p|x_i,p]|x_i]}\right)^{1/2} \exp\left(-\frac{1}{2}\frac{(\mathbb{E}[\theta - p|x_i])^2}{\mathbb{V}[\theta - p|x_i,p] + \mathbb{V}[\mathbb{E}[\theta - p|x_i,p]|x_i]}\right) \\
&= -\mathrm{e}^{c\gamma_i}\sqrt{\frac{\mathbb{V}[\theta - p|x_i,p]}{\mathbb{V}[\theta - p|x_i]}} \exp\left(-\frac{1}{2}\frac{(\mathbb{E}[\theta - p|x_i])^2}{\mathbb{V}[\theta - p|x_i]}\right),
\end{aligned} \tag{C4}$$

where we have used (C1) to go from the second to the third line. Comparing (C3) and (C4) gives the first ratio in Lemma 2.

C2. Timing 2: Order type chosen before the signal is observed

Under timing 2, traders choose their order type before knowing x_i. We recover the relevant expected utility levels by integrating those under timing 1 over x_i. First, noting that $y_{M,i}^2$

has a noncentral *chi*-square distribution and making use of (C2), we obtain:

$$\mathbb{E}\left[W_M\left(x_i;\gamma_i\right)\right] = \mathbb{E}\left[-\mathrm{e}^{-y_{M,i}^2}\right] = -\left(1+\frac{\mathbb{V}[\mathbb{E}[\theta-p|x_i]]}{\mathbb{V}[\theta-p|x_i]}\right)^{-1/2} = -\left(\frac{\mathbb{V}[\theta-p]}{\mathbb{V}[\theta-p|x_i]}\right)^{-1/2}. \tag{C5}$$

Now, using the law of iterated expectations and integrating (C4) over x_i we get:

$$\begin{aligned}\mathbb{E}\left[W_{I\backslash M}\left(x_i,p;\gamma_i\right)\right] &= \mathbb{E}\left[\mathbb{E}\left[W_{I\backslash M}\left(x_i,p;\gamma_i\right)\middle| x_i\right]\right]\\ &= -\mathrm{e}^{c\gamma_i}\left(\frac{\mathbb{V}[\theta-p|x_i,p]}{\mathbb{V}[\theta-p|x_i]}\right)^{1/2}\mathbb{E}\left[\exp\left(-\frac{1}{2}\frac{\left(\mathbb{E}[\theta-p|x_i]\right)^2}{\mathbb{V}[\theta-p|x_i]}\right)\right]\end{aligned} \tag{C6}$$

But again, $\left(\mathbb{E}[\theta-p|x_i]\right)^2$ also has a noncentral *chi*-square distribution and we have:

$$\mathbb{E}\left[\exp\left(\frac{-\left(\mathbb{E}[\theta-p|x_i]\right)^2}{2\mathbb{V}[\theta-p|x_i]}\right)\right] = \left(1+\frac{\mathbb{V}[\mathbb{E}[\theta-p|x_i]]}{\mathbb{V}[\theta-p|x_i]}\right)^{-1/2}. \tag{C7}$$

Substituting (C7) into (C6), making use of (C2) and rearranging, we get:

$$\mathbb{E}\left[W_{I\backslash M}\left(x_i,p;\gamma_i\right)\right] = -\mathrm{e}^{c\gamma_i}\left(\frac{\mathbb{V}[\theta-p]}{\mathbb{V}[\theta-p|x_i,p]}\right)^{-1/2}. \tag{C8}$$

Comparing (C5) and (C8) gives the second ratio in Lemma 2.

D. Proof of Proposition 2

See the following technical appendix.

Technical Appendix

3.A Proof of Proposition 2

D.1. Proof of (b)

From Lemma 2, the risk aversion coefficient γ_i of the marginal informed trader satisfies $\mathrm{e}^{-2c\gamma_i}\mathbb{V}[\theta - p|x_i] = \mathbb{V}[\theta - p|x_i, p]$. Let us define the function $f : [0,1] \times \mathbb{R}^{+} \to \mathbb{R}$ as follows:

$$f(i, \alpha_x) = \mathrm{e}^{-2c\gamma_i} - \mathbb{V}[\theta - p|x_i, p] / \mathbb{V}[\theta - p|x_i],$$

so that a root $i^*(\alpha_x)$ of $f(\cdot, \alpha_x)$ defines the marginal trader. From (A2) and (A4) we get:

$$\frac{\mathbb{V}[\theta - p|x_i, p]}{\mathbb{V}[\theta - p|x_i]} = \frac{\frac{1}{\alpha_\theta}\left(1 - \left(1 + \frac{\rho\alpha_\theta}{\alpha_x} + \frac{\alpha_\theta}{\frac{\alpha_x}{1-\rho} + [B(i,\alpha_x)]^2\alpha_\varepsilon}\right)^{-1}\right)\left(1 + \frac{\rho\alpha_\theta}{\alpha_x} + \frac{\alpha_\theta}{[B(i,\alpha_x)]^2\alpha_\varepsilon}\right)^2}{\frac{1}{\alpha_x+\alpha_\theta}\left(\rho + \frac{\rho\alpha_\theta}{\alpha_x} + \frac{\alpha_\theta}{[B(i,\alpha_x)]^2\alpha_\varepsilon}\right)^2 + \frac{\rho(1-\rho)}{\alpha_x} + \frac{1}{[B(i,\alpha_x)]^2\alpha_\varepsilon}}, \tag{D1}$$

where $B(i, \alpha_x)$ is defined as the unique solution to:

$$\begin{aligned} B(i, \alpha_x) = {} & \left(\frac{\alpha_x}{1 + (1-\rho)\rho\,[B(i,\alpha_x)]^2\,\alpha_\varepsilon/\alpha_x}\right)\int_0^i \gamma_i^{-1}\mathrm{d}i \\ & + \left(\left(\frac{1}{\alpha_x} + \frac{1}{\alpha_\theta}\right)\left(1 + \frac{[B(i,\alpha_x)]^2\,\rho\alpha_\varepsilon}{\alpha_x}\right) - \frac{1}{[B(i,\alpha_x)]^2\,\alpha_\varepsilon\,(1 + \rho\alpha_\theta/\alpha_x) + \alpha_\theta}\right)^{-1}\int_i^1 \gamma_i^{-1}\mathrm{d}i. \end{aligned}$$

We show that for α_x sufficiently large the equilibrium cannot be such that $I = M$ or $I = I \backslash M$. We know from Appendix B that $\lim_{\alpha_x \to +\infty} B(0, \alpha_x)$ is positive and finite while

$$B\left(1, \alpha_x\right) \underset{\alpha_x \to +\infty}{\sim} \mathbf{1}_{\rho \in (0,1)} \left(\frac{\alpha_x^2}{\gamma_I (1-\rho) \rho \alpha_\varepsilon} \right)^{1/3} + \mathbf{1}_{\rho \in \{0,1\}} \left(\frac{\alpha_x}{\gamma_I} \right)$$

This in turn implies that $f(0, \alpha_x) \underset{\alpha_x \to \infty}{\to} \mathrm{e}^{-2c\gamma_0} > 0$ while $f(1, \alpha_x) \underset{\alpha_x \to \infty}{\to} \mathrm{e}^{-2c\gamma_1} - 1 < 0$. Thus, by the intermediate value theorem, for α_x sufficiently large, $f(\cdot, \alpha_x)$ has at least one interior root i^* (but none at the corners).

D.2. Proof of (c)

Let us first prove that the set of roots of $f(., \alpha_x)$ is compact when α_x is sufficiently large. To this purpose, we extend $f(., \alpha_x)$ to $\mathbb{R} \times \mathbb{R}_+^*$ and continuously extend $f(., \alpha_x)$ on $\mathbb{R}$ by means of the following $f_{\alpha_x}^{ext}$ function:

$$f_{\alpha_x}^{ext}(i) = \mathbf{1}_{i \in [0;1]} f(i, \alpha_x) + \mathbf{1}_{i > 1} f(1, \alpha_x) + \mathbf{1}_{i < 0} f(0, \alpha_x).$$

For α_x sufficiently large, the set of roots of $f_{\alpha_x}^{ext}$ is a compact set of $\mathbb{R}$: it is bounded, because it is included in $(0,1)$; it is also closed, since it is the reciprocal image of a closed subset of $\mathbb{R}$ (namely, $\{0\}$) by a continuous function. The set of roots being compact, for each α_x we can define a maximum and a minimum solution, which we denote $i^*_{\max}(\alpha_x)$ and $i^*_{\min}(\alpha_x)$.

Next, we show by contradiction that $\forall \epsilon > 0, \exists A, \forall \alpha_x, [\alpha_x \geq A \Rightarrow \forall i^*(\alpha_x) \in \left(f_{\alpha_x}^{ext}\right)^{-1}(\{0\}), 0 < i^*(\alpha_x) < \epsilon]$. In that part of the proof we need to distinguish $\rho \in (0,1)$ and $\rho \in \{0,1\}$.

Case 1: $\rho \in (0,1)$

For $i \in [\epsilon, 1]$ we have $\mathrm{e}^{-2\gamma_1 c} \leq \mathrm{e}^{-2\gamma_i c} \leq \mathrm{e}^{-2\gamma_\epsilon c}$ and, using (D1),

$$\frac{\mathbb{V}[\theta - p|x_i, p]}{\mathbb{V}[\theta - p|x_i]} \geq \frac{\frac{1}{\alpha_\theta}\left[1 - \left(1 + \frac{\rho\alpha_\theta}{\alpha_x} + \frac{\alpha_\theta}{\frac{\alpha_x}{1-\rho} + [B(i,\alpha_x)]^2\alpha_\varepsilon}\right)^{-1}\right]}{\frac{1}{\alpha_x+\alpha_\theta}\left(\rho + \rho\frac{\alpha_\theta}{\alpha_x} + \frac{\alpha_\theta}{[B(i,\alpha_x)]^2\alpha_\varepsilon}\right)^2 + \frac{(1-\rho)\rho}{\alpha_x} + \frac{\alpha_\varepsilon^{-1}}{[B(i,\alpha_x)]^2}}$$

$$\geq \frac{\frac{1}{\alpha_\theta}\left[1 - \left(1 + \frac{\rho\alpha_\theta}{\alpha_x} + \frac{\alpha_\theta}{\frac{\alpha_x}{1-\rho} + [B(1,\alpha_x)]^2\alpha_\varepsilon}\right)^{-1}\right]}{\frac{1}{\alpha_x+\alpha_\theta}\left(\rho + \rho\frac{\alpha_\theta}{\alpha_x} + \frac{\alpha_\theta}{B(\epsilon,\alpha_x)^2\alpha_\varepsilon}\right)^2 + \frac{(1-\rho)\rho}{\alpha_x} + \frac{\alpha_\varepsilon^{-1}}{B(\epsilon,\alpha_x)^2}} \equiv \frac{NUM}{DEN}$$

We know from Appendix B that for $\rho \in (0,1)$ we have:

$$B(\epsilon,\alpha_x) \underset{\alpha_x\to\infty}{\sim} \left(\frac{\left(\int_0^\epsilon \gamma_i^{-1}\mathrm{d}i\right)\alpha_x^2}{(1-\rho)\rho\alpha_\varepsilon}\right)^{1/3} \quad \text{and} \quad B(1,\alpha_x) \underset{\alpha_x\to\infty}{\sim} \left(\frac{\left(\int_0^1 \gamma_i^{-1}\mathrm{d}i\right)\alpha_x^2}{(1-\rho)\rho\alpha_\varepsilon}\right)^{1/3},$$

which implies that $NUM \underset{\alpha_x\to+\infty}{\sim} \rho/\alpha_x$ while $DEN \underset{\alpha_x\to+\infty}{\sim} \alpha_x/\rho$, and hence $NUM/DEN \underset{\alpha_x\to\infty}{\to} 1$. Hence, for any arbitrarily small $\chi > 0$, for α_x sufficiently large we have $|\mathbb{V}[\theta - p|x_i, p]/\mathbb{V}[\theta - p|x_i] - 1| \leq \chi$, so that $\mathbb{V}[\theta - p|x_i, p]/\mathbb{V}[\theta - p|x_i] \geq 1 - \chi$. Choosing $\chi = 1 - \mathrm{e}^{-\gamma_\epsilon c}$, we obtain:

$$\mathbb{V}[\theta - p|x_i, p]/\mathbb{V}[\theta - p|x_i] \geq 1 - (1 - \mathrm{e}^{-\gamma_\epsilon c}) = \mathrm{e}^{-\gamma_\epsilon c} > \mathrm{e}^{-2\gamma_\epsilon c} \geq \mathrm{e}^{-2\gamma_i c},$$

so that $f(i, +\infty) < 0$. Therefore, for α_x sufficiently large there cannot be any root of $f(\cdot, \alpha_x)$ in $[\epsilon, 1]$, i.e., any root must be in $[0, \epsilon)$.

Case 2: $\rho \in \{0,1\}$

In the case where $\rho = 0$, equation (D1) gives:

$$\frac{\mathbb{V}[\theta - p|x_i, p]}{\mathbb{V}[\theta - p|x_i]} = \frac{\left(\frac{1}{\alpha_\theta + \alpha_x + [B(i,\alpha_x)]^2 \alpha_\varepsilon}\right)\left(1 + \frac{\alpha_\theta}{[B(i,\alpha_x)]^2 \alpha_\varepsilon}\right)^2}{\frac{1}{\alpha_x + \alpha_\theta}\left(\frac{\alpha_\theta}{[B(i,\alpha_x)]^2 \alpha_\varepsilon}\right)^2 + \frac{1}{[B(i,\alpha_x)]^2 \alpha_\varepsilon}}$$

$$\geq \left(\alpha_\theta + \alpha_x + [B(i,\alpha_x)]^2 \alpha_\varepsilon\right)^{-1}\left(\frac{1}{\alpha_x + \alpha_\theta}\left(\frac{\alpha_\theta}{[B(i,\alpha_x)]^2 \alpha_\varepsilon}\right)^2 + \frac{1}{[B(i,\alpha_x)]^2 \alpha_\varepsilon}\right)^{-1}$$

$$\geq \left(\left(\frac{\alpha_\theta}{[B(i,\alpha_x)]^2 \alpha_\varepsilon}\right)^2 + \frac{\alpha_\theta^2}{\alpha_x + \alpha_\theta}\left(\frac{1}{[B(i,\alpha_x)]^2 \alpha_\varepsilon}\right) + \frac{\alpha_\theta + \alpha_x}{[B(i,\alpha_x)]^2 \alpha_\varepsilon} + 1\right)^{-1}$$

$$\geq \left(\left(\frac{\alpha_\theta}{[B(\epsilon,\alpha_x)]^2 \alpha_\varepsilon}\right)^2 + \frac{\alpha_\theta^2}{\alpha_x + \alpha_\theta}\left(\frac{1}{[B(\epsilon,\alpha_x)]^2 \alpha_\varepsilon}\right) + \frac{\alpha_\theta + \alpha_x}{[B(\epsilon,\alpha_x)]^2 \alpha_\varepsilon} + 1\right)^{-1}.$$

We know from Appendix B that, for $\rho = 0$ we have $B(\epsilon,\alpha_x) \underset{\alpha_x \to \infty}{\sim} \alpha_x \left(\int_0^\epsilon \gamma_i^{-1} \mathrm{d}i\right)$, which in turn implies that

$$\left(\left(\frac{\alpha_\theta}{[B(\epsilon,\alpha_x)]^2 \alpha_\varepsilon}\right)^2 + \frac{\alpha_\theta^2}{\alpha_x + \alpha_\theta}\left(\frac{1}{[B(\epsilon,\alpha_x)]^2 \alpha_\varepsilon}\right) + \frac{\alpha_\theta + \alpha_x}{[B(\epsilon,\alpha_x)]^2 \alpha_\varepsilon} + 1\right)^{-1} \underset{\alpha_x \to \infty}{\to} 1,$$

so we can apply the same reasoning as when $\rho \in (0,1)$ to show that $\mathbb{V}[\theta - p|x_i, p]/\mathbb{V}[\theta - p|x_i] > \mathrm{e}^{-2\gamma_i c}$ for α_x sufficiently large.

When $\rho = 1$ we have

$$\frac{\mathbb{V}[\theta - p|x_i, p]}{\mathbb{V}[\theta - p|x_i]} = \frac{\frac{1}{\alpha_\theta}\left(1 - \left(1 + \frac{\alpha_\theta}{\alpha_x}\right)^{-1}\right)\left(1 + \frac{\alpha_\theta}{\alpha_x} + \frac{\alpha_\theta}{[B(i,\alpha_x)]^2 \alpha_\varepsilon}\right)^2}{\frac{1}{\alpha_x + \alpha_\theta}\left(1 + \frac{\alpha_\theta}{\alpha_x} + \frac{\alpha_\theta}{[B(i,\alpha_x)]^2 \alpha_\varepsilon}\right)^2 + \frac{1}{[B(i,\alpha_x)]^2 \alpha_\varepsilon}}$$

$$\geq \left[\left(1 + \frac{\alpha_\theta}{\alpha_x} + \frac{\alpha_\theta}{[B(i,\alpha_x)]^2 \alpha_\varepsilon}\right)^2 + \frac{\alpha_x + \alpha_\theta}{[B(i,\alpha_x)]^2 \alpha_\varepsilon}\right]^{-1}$$

$$\geq \left[\left(1 + \frac{\alpha_\theta}{\alpha_x} + \frac{\alpha_\theta}{[B(\epsilon,\alpha_x)]^2 \alpha_\varepsilon}\right)^2 + \frac{\alpha_x + \alpha_\theta}{[B(\epsilon,\alpha_x)]^2 \alpha_\varepsilon}\right]^{-1}.$$

Again, in this case $B(\epsilon, \alpha_x) \underset{\alpha_x \to \infty}{\sim} \alpha_x \left(\int_0^\epsilon \gamma_i^{-1} \mathrm{d}i\right)$ and thus:

$$\left[\left(1 + \frac{\alpha_\theta}{\alpha_x} + \frac{\alpha_\theta}{[B(\epsilon,\alpha_x)]^2 \alpha_\varepsilon}\right)^2 + \frac{\alpha_x + \alpha_\theta}{[B(\epsilon,\alpha_x)]^2 \alpha_\varepsilon}\right]^{-1} \underset{\alpha_x \to \infty}{\to} 1,$$

so that $\mathbb{V}[\theta - p|x_i, p]/\mathbb{V}[\theta - p|x_i] > \mathrm{e}^{-2\gamma_i c}$ for α_x sufficiently large. It follows that, $\forall \rho \in [0,1]$ and for α_x sufficiently large, $\mathbb{V}[\theta - p|x_i, p]/\mathbb{V}[\theta - p|x_i] > \mathrm{e}^{-2\gamma_i c}$, so that $f(i, \alpha_x) < 0$. This implies that, for any choice function $\hat{\imath} : \alpha_x \to \hat{\imath}(\alpha_x) \in \left(f^{ext}_{\alpha_x}\right)^{-1}(\{0\})$, we have $\hat{\imath}(\alpha_x) \underset{\alpha_x \to \infty}{\to} 0$; in particular, $i^*_{\max}(\alpha_x) \underset{\alpha_x \to \infty}{\to} 0$ and $i^*_{\min}(\alpha_x) \underset{\alpha_x \to \infty}{\to} 0$.

D.3. Proof of (d)

For any choice function $\hat{\imath}(\alpha_x)$, since $f(\hat{\imath}(\alpha_x), \alpha_x) = 0$ and $\hat{\imath}(\alpha_x) \underset{\alpha_x \to \infty}{\to} 0$ we have, by continuity,

$$\Xi \equiv \frac{\frac{1}{\alpha_\theta}\left(1 - \left(1 + \frac{\rho\alpha_\theta}{\alpha_x} + \frac{\alpha_\theta}{\frac{\alpha_x}{1-\rho} + [B(\hat{\imath}(\alpha_x),\alpha_x)]^2\alpha_\varepsilon}\right)^{-1}\right)\left(1 + \frac{\rho\alpha_\theta}{\alpha_x} + \frac{\alpha_\theta}{[B(\hat{\imath}(\alpha_x),\alpha_x)]^2\alpha_\varepsilon}\right)^2}{\left(\frac{1}{\alpha_x+\alpha_\theta}\left(\rho + \frac{\rho\alpha_\theta}{\alpha_x} + \frac{\alpha_\theta}{B(\hat{\imath}(\alpha_x),\alpha_x)^2\alpha_\varepsilon}\right)^2 + \frac{(1-\rho)\rho}{\alpha_x} + \frac{\alpha_\varepsilon^{-1}}{[B(\hat{\imath}(\alpha_x),\alpha_x)]^2}\right)} \underset{\alpha_x \to \infty}{\to} \mathrm{e}^{-2\gamma_0 c} \tag{D1}$$

We first show by contradiction that, for $\rho \in (0,1)$ we have $B(\hat{\imath}(\alpha_x), \alpha_x) \underset{\alpha_x \to \infty}{\to} \infty$, that is, $\forall M > 0, \exists A > 0, \forall \alpha_x, [\alpha_x \geq A \Rightarrow B(\hat{\imath}(\alpha_x), \alpha_x) > M]$. If it were not the case, we could find $M > 0$ such that, $\forall A > 0$, there exists an $\alpha_x(A)$ verifying $\alpha_x(A) \geq A$ and $B(\hat{\imath}(\alpha_x)) \leq M$. Choose any such M and consider the serie $\alpha_n = (\alpha_x(n))_{n \geq n_0}$ with n_0 large enough to ensure the existence of $(\hat{\imath}(\alpha_n))_{n \geq n_0}$. In this case,

$$\frac{1}{\alpha_\theta}\left(1 - \left(1 + \frac{\rho\alpha_\theta}{\alpha_n} + \frac{\alpha_\theta}{\frac{\alpha_n}{1-\rho} + [B(\hat{\imath}(\alpha_n), \alpha_n)]^2 \alpha_\varepsilon}\right)^{-1}\right) \underset{n \to \infty}{\to} 0$$

while

$$0 \leq \left(1 + \frac{\rho\alpha_\theta}{\alpha_n} + \frac{\alpha_\theta}{B(\hat{\imath}(\alpha_n), \alpha_n)^2 \alpha_\varepsilon}\right)^2 \leq \left(1 + \frac{\rho\alpha_\theta}{\alpha_n} + \frac{\alpha_\theta}{B(0, \alpha_n)^2 \alpha_\varepsilon}\right)^2,$$

where the right hand side is bounded above as $n \to \infty$. Moreover,

$$\left(\frac{1}{\alpha_n+\alpha_\theta}\left(\rho+\frac{\rho\alpha_\theta}{\alpha_n}+\frac{\alpha_\theta}{B\left(\hat{\imath}(\alpha_n)\right)^2\alpha_\varepsilon}\right)^2+\frac{(1-\rho)\rho}{\alpha_n}+\frac{\alpha_\varepsilon^{-1}}{B\left(\hat{\imath}(\alpha_n)\right)^2}\right) \geq \left(\frac{1}{\alpha_n+\alpha_\theta}\left(\rho+\frac{\rho\alpha_\theta}{\alpha_n}+\frac{\alpha_\theta}{M^2\alpha_\varepsilon}\right)^2+\frac{(1-\rho)\rho}{\alpha_n}+\frac{1}{M^2\alpha_\varepsilon}\right)$$

which is bounded below as $n \to \infty$. Hence, we would have $\Xi \underset{n\to\infty}{\to} 0$. However, since the limit is unique this would require $\mathrm{e}^{-2\gamma_0 c}=0$, a contradiction. It must thus be that $B(\hat{\imath}(\alpha_x)) \underset{\alpha_x\to\infty}{\to} \infty$, which in turn implies:

$$\left(1-\left(1+\frac{\rho\alpha_\theta}{\alpha_x}+\frac{\alpha_\theta}{\frac{\alpha_x}{1-\rho}+\left[B(\hat{\imath}(\alpha_x),\alpha_x)\right]^2\alpha_\varepsilon}\right)^{-1}\right) \underset{\alpha_x\to\infty}{\sim} \frac{\rho\alpha_\theta}{\alpha_x}+\frac{(1-\rho)\,\alpha_\theta}{\alpha_x+(1-\rho)\,B(\hat{\imath}(\alpha_x),\alpha_x)^2\alpha_\varepsilon},$$

and
$$\left(\frac{1}{\alpha_x+\alpha_\theta}\left(\rho+\frac{\rho\alpha_\theta}{\alpha_x}+\frac{\alpha_\theta}{\left[B(\hat{\imath}(\alpha_x),\alpha_x)\right]^2\alpha_\varepsilon}\right)^2+\frac{(1-\rho)\rho}{\alpha_x}+\frac{1}{\left[B(\hat{\imath}(\alpha_x),\alpha_x)\right]^2\alpha_\varepsilon}\right) \underset{\alpha_x\to\infty}{\sim} \frac{\rho^2}{\alpha_x}+\frac{(1-\rho)\rho}{\alpha_x}+\frac{1}{\left[B(\hat{\imath}(\alpha_x),\alpha_x)\right]^2\alpha_\varepsilon} = \frac{\rho}{\alpha_x}+\frac{1}{\left[B(\hat{\imath}(\alpha_x),\alpha_x)\right]^2\alpha_\varepsilon}$$

so that

$$\Xi \underset{\alpha_x\to\infty}{\sim} \frac{\frac{1}{\alpha_\theta}\left(\frac{\rho\alpha_\theta}{\alpha_x}+\frac{(1-\rho)\alpha_\theta}{\alpha_x+(1-\rho)\left[B(\hat{\imath}(\alpha_x),\alpha_x)\right]^2\alpha_\varepsilon}\right)}{\frac{\rho}{\alpha_x}+\frac{1}{\left[B(\hat{\imath}(\alpha_x),\alpha_x)\right]^2\alpha_\varepsilon}} = \frac{\frac{\rho}{\alpha_x}+\frac{1-\rho}{\alpha_x+(1-\rho)\left[B(\hat{\imath}(\alpha_x),\alpha_x)\right]^2\alpha_\varepsilon}}{\frac{\rho}{\alpha_x}+\frac{1}{\left[B(\hat{\imath}(\alpha_x),\alpha_x)\right]^2\alpha_\varepsilon}} = \frac{\rho+\frac{1-\rho}{1+(1-\rho)\frac{\left[B(\hat{\imath}(\alpha_x),\alpha_x)\right]^2\alpha_\varepsilon}{\alpha_x}}}{\rho+\frac{\alpha_x}{\left[B(\hat{\imath}(\alpha_x),\alpha_x)\right]^2\alpha_\varepsilon}}.$$

We know from (D1) that $\Xi \underset{\alpha_x\to\infty}{\to} \mathrm{e}^{-2\gamma_0 c}$, hence

$$\left(\rho+\frac{1-\rho}{1+(1-\rho)\frac{\left[B(\hat{\imath}(\alpha_x),\alpha_x)\right]^2\alpha_\varepsilon}{\alpha_x}}\right)\left(\rho+\frac{\alpha_x}{\left[B(\hat{\imath}(\alpha_x),\alpha_x)\right]^2\alpha_\varepsilon}\right)^{-1} \underset{\alpha_x\to\infty}{\to} \mathrm{e}^{-2\gamma_0 c}.$$

Now, define the following function:

$$h_\rho : x \to \left(\rho+\frac{1-\rho}{1+(1-\rho)\,x}\right)\left(\rho+x^{-1}\right)^{-1},$$

which is continuous, increasing on $\mathbb{R}_+^*$ and maps $\mathbb{R}_+^*$ onto $(0,1)$. Then by continuity we have:

$$\frac{\left[B(\hat{\imath}(\alpha_x),\alpha_x)\right]^2\alpha_\varepsilon}{\alpha_x} \underset{\alpha_x\to\infty}{\to} h_\rho^{-1}(\mathrm{e}^{-2\gamma_0 c}),$$

and thus

$$B(\hat{\imath}(\alpha_x),\alpha_x) \underset{\alpha_x\to\infty}{\sim} \left(\frac{\alpha_x}{\alpha_\varepsilon}h_\rho^{-1}(\exp\left[-2\gamma_0 c\right])\right)^{1/2}.$$

We infer from the latter expression as well as (B1) above that

$$\alpha_p = \frac{\alpha_x}{\rho+\frac{\alpha_x}{\left[B(\hat{\imath}(\alpha_x),\alpha_x)\right]^2\alpha_\varepsilon}} \underset{\alpha_x\to\infty}{\sim} \frac{\alpha_x}{\rho+\left[h_\rho^{-1}(\mathrm{e}^{-2\gamma_0 c})\right]^{-1}}.$$

D.4. Proof of (a)

We now show that i^* (the marginal trader) is unique for α_x sufficiently large. To this purpose, consider the differentiable functions of B defined as follows:

$$u(B) = \frac{1}{\alpha_\theta}\left(1-\left(1+\rho\frac{\alpha_\theta}{\alpha_x}+\frac{\alpha_\theta}{\frac{\alpha_x}{1-\rho}+B^2\alpha_\varepsilon}\right)^{-1}\right)\left(1+\frac{\rho\alpha_\theta}{\alpha_x}+\frac{\alpha_\theta}{B^2\alpha_\varepsilon}\right)^2,$$

$$d(B) = \left(\frac{1}{\alpha_x+\alpha_\theta}\left(\rho+\frac{\rho\alpha_\theta}{\alpha_x}+\frac{\alpha_\theta}{B^2\alpha_\varepsilon}\right)^2+\frac{(1-\rho)\rho}{\alpha_x}+\frac{1}{B^2\alpha_\varepsilon}\right),\; r(B) = \frac{u(B)}{d(B)}.$$

For all $i \in [i^*_{\min}(\alpha_x), i^*_{\max}(\alpha_x)]$ we have, after some manipulations:

$$r'(B(i)) = \frac{u'(B(i))}{d(B(i))} - \frac{d'(B(i))}{d^2(B(i))}u(B(i)) \geq \lambda_1(\alpha_x)+\lambda_2(\alpha_x),$$

where $\lambda_1(\alpha_x)$ and $\lambda_2(\alpha_x)$ are given, if $\rho \in [0;1)$, by:

$$\lambda_1(\alpha_x) \equiv -\frac{2\overline{B}\alpha_\varepsilon\left(\frac{\left(1+\rho\frac{\alpha_\theta}{\alpha_x}+\frac{\alpha_\theta}{\underline{B}^2\alpha_\varepsilon}\right)^2}{\left(\left(1+\frac{\rho\alpha_\theta}{\alpha_x}\right)\left(\frac{\alpha_x}{1-\rho}+\underline{B}^2\alpha_\varepsilon\right)+\alpha_\theta\right)^2}+\frac{2\left(1+\frac{\rho\alpha_\theta}{\alpha_x}+\frac{\alpha_\theta}{\underline{B}^2\alpha_\varepsilon}\right)}{\left(\underline{B}^2\alpha_\varepsilon\right)^2}\left(1-\left(1+\frac{\rho\alpha_\theta}{\alpha_x}+\frac{\alpha_\theta}{\frac{\alpha_x}{1-\rho}+\underline{B}^2\alpha_\varepsilon}\right)^{-1}\right)\right)}{\frac{1}{\alpha_x+\alpha_\theta}\left(\rho+\rho\frac{\alpha_\theta}{\alpha_x}+\frac{\alpha_\theta}{\overline{B}^2\alpha_\varepsilon}\right)^2+(1-\rho)\rho\frac{1}{\alpha_x}+\frac{\alpha_\varepsilon^{-1}}{\overline{B}^2}}$$

$$\underset{\alpha_x\to\infty}{\sim} -\frac{2\left(\alpha_\varepsilon h_\rho^{-1}(\mathrm{e}^{-2\gamma_0 c})\right)^{1/2}}{\left(\left(\frac{1}{(1-\rho)}+h_\rho^{-1}(\mathrm{e}^{-2\gamma_0 c})\right)\right)^2\left(\rho+\frac{1}{h_\rho^{-1}(\mathrm{e}^{-2\gamma_0 c})}\right)}\alpha_x^{-1/2},$$

and

$$\lambda_2(\alpha_x) \equiv$$

$$\frac{\left(\frac{2}{\alpha_x+\alpha_\theta}\left(\rho+\rho\frac{\alpha_\theta}{\alpha_x}+\frac{\alpha_\theta}{\overline{B}^2\alpha_\varepsilon}\right)\left(\frac{2}{\overline{B}^3\alpha_\varepsilon}\right)+\frac{2}{\overline{B}^3\alpha_\varepsilon\alpha_\theta}\right)\left(1-\left(1+\frac{\rho\alpha_\theta}{\alpha_x}+\frac{\alpha_\theta}{\frac{\alpha_x}{(1-\rho)}+\overline{B}^2\alpha_\varepsilon}\right)^{-1}\right)\left(1+\frac{\rho\alpha_\theta}{\alpha_x}+\frac{\alpha_\theta}{\overline{B}^2\alpha_\varepsilon}\right)^2}{\left(\frac{1}{\alpha_x+\alpha_\theta}\left(\rho+\frac{\rho\alpha_\theta}{\alpha_x}+\frac{\alpha_\theta}{\underline{B}^2\alpha_\varepsilon}\right)^2+\frac{(1-\rho)\rho}{\alpha_x}+\frac{1}{\underline{B}^2\alpha_\varepsilon}\right)^2}$$

$$\underset{\alpha_x\to\infty}{\sim} \frac{2}{\left(h_\rho^{-1}(\mathrm{e}^{-2\gamma_0 c})\right)^{3/2}\alpha_\varepsilon^{-1/2}}\frac{1}{\left(\rho+\frac{1}{h_\rho^{-1}(\mathrm{e}^{-2\gamma_0 c})}\right)^2}\left(\rho+\frac{1}{\frac{1}{1-\rho}+h_\rho^{-1}(\mathrm{e}^{-2\gamma_0 c})}\right)\alpha_x^{-1/2}$$

and where we have defined $\underline{B} \equiv B(i^*_{\min}(\alpha_x),\alpha_x)$ and $\overline{B} \equiv B(i^*_{\max}(\alpha_x),\alpha_x)$. Next, define:

$$\mu_1 \equiv \frac{2\left(\alpha_\varepsilon h_\rho^{-1}(\mathrm{e}^{-2\gamma_0 c})\right)^{1/2}}{\left(\left(\frac{1}{(1-\rho)}+h_\rho^{-1}(\mathrm{e}^{-2\gamma_0 c})\right)\right)^2\left(\rho+\frac{1}{h_\rho^{-1}(\mathrm{e}^{-2\gamma_0 c})}\right)},\ \mu_2 \equiv \frac{2\left(\rho+\frac{1}{\frac{1}{1-\rho}+h_\rho^{-1}(\mathrm{e}^{-2\gamma_0 c})}\right)}{\left(h_\rho^{-1}(\mathrm{e}^{-2\gamma_0 c})\right)^{3/2}\alpha_\varepsilon^{-1/2}\left(\rho+\frac{1}{h_\rho^{-1}(\mathrm{e}^{-2\gamma_0 c})}\right)^2}$$

Since $\mu_2 > \mu_1$, it follows that for α_x sufficiently large, and for any $i \in [i^*_{\min}(\alpha_x), i^*_{\max}(\alpha_x)]$ we have $r'(B(i)) \geq (\mu_2 - \mu_1)\,\alpha_x^{-1/2}/2 > 0$. It can be shown that $\partial B(i,\alpha_x)/\partial i > 0$, which in turn implies that $\partial r(B)/\partial i = r'(B(i))\partial B/\partial i > 0$. Then, on this range we have $\partial f(i,\alpha_x)/\partial i < 0$, so that $f(i,\alpha_x)$ is a bijection on $[i^*_{\min}(\alpha_x), i^*_{\max}(\alpha_x)]$. As $f(i^*_{\min}(\alpha_x),\alpha_x) = f(i^*_{\max}(\alpha_x),\alpha_x) = 0$, we have $i^*_{\min}(\alpha_x) = i^*_{\max}(\alpha_x)$, i.e., the solution is unique (for α_x sufficiently large).

If $\rho = 1$, we can show that $\lambda_2(\alpha_x) + \lambda_1(\alpha_x) \underset{\alpha_x\to\infty}{\sim} \lambda_2(\alpha_x)$ and the argument holds.

3.B Generalising the information structure

Our baseline information structure allows private noise to be correlated across informed traders. In so doing, we have made the natural simplifying assumption that the correlation coefficient ρ governing the correlation between any to pairs of draws for the private noise is constant. This directly implies that the idiosyncratic and aggregate components of the noise in the private signals vary at the same rate when we let α_x vary. Consequently, it is not possible in our baseline specification to look at the limiting cases where the two noise components vanish separately. In this section we show that all our results carry over to generalisation of the the information structure whereby the two noise components are allowed to vanish separately. Technically, this generalisation is done by relaxing the assumption that ρ is constant and by letting it instead vary with α_x. Intuitively, the reasons why our results can be straightforwardly generalised to this situation are as follows. First, Lemma 1 and Lemma 2 are valid irrespective of the amount of informational noise in the model, and for values of ρ over the full closed interval $[0,1]$. Hence they hold no matter how we let ρ vary with α_x and need not be extended. Second, Proposition 2 is also valid for $\rho \in [0,1]$, and there is no discontinuity between $\rho = 0$ and $\rho \to 0$, or between $\rho = 1$ and $\rho \to 1$. This implies that if one wants, for example, to examine the case where private noise vanishes quicker than common noise as $\alpha_x \to +\infty$, it is sufficient to look at the case where $\rho = 1$ and $\alpha_x \to +\infty$ in the statement of Proposition 2. Third, Proposition 1 (our results under *exogenous* order types) is also valid for $\rho \in [0,1]$, except that there is a discontinuity between $\rho = 0$ and $\rho \to 0$; we thus extend our analysis to study what happens when $\rho \to 0$ for private noise to vanish quicker than common noise.

Let us now formally substantiate these claims and extend Proposition 1 and 2 accordingly. Assume that $\alpha_x \to \rho(\alpha_x)$ is continuous in α_x and has a limit as $\alpha_x \to +\infty$. In particular, we will consider two case: $\rho(\alpha_x) \underset{\alpha_x \to \infty}{\to} 0$ (in which case the common noise component ultimately overwhelm the idiosyncratic noise component), and $\rho(\alpha_x) \underset{\alpha_x \to \infty}{\to} 1$ (in which case the opposite occurs).

3.B.1 Propositions 1 and 2 revisited

Proposition 1 and 2 can be generalised as follows:

Proposition 1b. (a) In a pure market-order market ($\nu = 0$), the informativeness of the price signal is bounded above; formally, $\alpha_p \underset{\alpha_x\to\infty}{\to} B_0^2\alpha_\varepsilon < +\infty$, where $B_0 > 0$ uniquely solves $\gamma_M B_0(\alpha_\theta^{-1} - \left(\alpha_\theta + \alpha_\varepsilon B_0^2\right)^{-1}) = 1$. (b) Whenever there is a positive mass of demand-schedule traders ($\nu > 0$), then the precision of the price signal is unbounded as $\alpha_x \to \infty$; more specifically,

1. If $\lim_{\alpha_x\to\infty} \rho = l \in (0;1]$, then $\alpha_p \underset{\alpha_x\to+\infty}{\sim} \alpha_x/\rho$;
2. If $\lim_{\alpha_x\to\infty} \rho = 0$ then $\alpha_x \underset{\alpha_x\to\infty}{=} o(\alpha_p)$ and $\alpha_p \underset{\alpha_x\to\infty}{=} O(\alpha_x^2)$; furthermore

   a) if $\lim_{\alpha_x\to\infty} \rho\alpha_x = 0$ then $\alpha_p \underset{\alpha_x\to\infty}{\sim} \left(\frac{\nu\alpha_x}{\gamma_{I\setminus M}}\right)^2 \alpha_\varepsilon$

   b) if $\lim_{\alpha_x\to\infty} \rho\alpha_x = M > 0$ then $\alpha_p \underset{\alpha_x\to\infty}{\sim} \frac{(\alpha_x)^2}{M}\frac{1}{1+(M')^{-2}}$, where M' is the unique positive solution to $X\left(1+X^2\right) = (\nu\sqrt{M\alpha_\varepsilon})/\gamma_{I\setminus M}$;

   c) if $\lim_{\alpha_x\to\infty} \rho\alpha_x = \infty$ then $\alpha_p \underset{\alpha_x\to+\infty}{\sim} \alpha_x/\rho$

Proposition 2b. For any $(\alpha_\theta, \alpha_\varepsilon) \in \mathbb{R}^{+2}$, and $\alpha_x \to +\infty$, (a) the solution $(\bar\gamma, B)$ to (10)–(11) is unique; (b) both M and $I\setminus M$ have strictly positive measure (i.e., the equilibrium is interior); (c) $\bar\gamma \to \gamma_0$ (i.e., market-order traders eventually overwhelm the market); (d) α_p goes to infinity as the same rate as α_x; formally, defining $\lim_{\alpha_x\to\infty} \rho = l$, the bijection $h_l : \mathbb{R}^+ \to (0,1)$, $h_l(x) = \left(l + \frac{(1-l)}{1+x(1-l)}\right)\left(l + x^{-1}\right)^{-1}$ and h_l^{-1} its inverse, we have

$$\alpha_p \underset{\alpha_x\to\infty}{\sim} \frac{\alpha_x}{l + \left[h_l^{-1}(\mathrm{e}^{-2\gamma_0 c})\right]^{-1}}.$$

Notice that

$$h_l^{-1} : x \to \frac{\sqrt{1 + 4l\,(1-l)\,\frac{x}{1-x}} - 1}{2l\,(1-l)}$$

if $l \in (0,1)$ while $h_0^{-1} = h_1^{-1} : x \to x/\,(1-x)$.

3.B.2 Proof of Proposition 1b

Full market order case

The first thing to note is that the full market order case ($\nu = 0$) remains unchanged. Indeed, B and B_0 solve, respectively:

$$\gamma_I B = \left[\left(\alpha_x^{-1} + \alpha_\theta^{-1}\right)\left(1 + B^2\rho\alpha_\varepsilon/\alpha_x\right) - \left(\alpha_\theta + B^2\alpha_\varepsilon\left(1 + \rho\alpha_\theta/\alpha_x\right)\right)^{-1}\right]^{-1}$$

and

$$\gamma_I B_0 = \left(\alpha_\theta^{-1} - \left(\alpha_\theta + \alpha_\varepsilon B_0^2\right)^{-1}\right)^{-1}$$

After some algebraic manipulations, we get:

$$(B - B_0) = \\ -\alpha_x^{-1}\left(1 + \frac{B^2\rho\alpha_\varepsilon}{\alpha_x} + \frac{B^2\rho\alpha_\varepsilon}{\alpha_\theta} + \frac{B^2\alpha_\varepsilon\rho\alpha_\theta}{\left(\alpha_\theta + B^2\alpha_\varepsilon\left(1 + \frac{\rho\alpha_\theta}{\alpha_x}\right)\right)\left(\alpha_\theta + \alpha_\varepsilon B_0^2\right)}\right)\frac{\left(\gamma_I\right)^2 BB_0}{\left(\gamma_I + \frac{(B+B_0)\alpha_\varepsilon}{\left(\alpha_\theta + B^2\alpha_\varepsilon\left(1+\frac{\rho\alpha_\theta}{\alpha_x}\right)\right)\left(\alpha_\theta + \alpha_\varepsilon B_0^2\right)}\right)}$$

The same argument as before holds, implying that B is bounded above as $\alpha_x \to \infty$. Thus, the right hand side of the latter equation goes to 0 as $\alpha_x \to \infty$. This implies that $B \underset{\alpha_x\to\infty}{\to} B_0$, as before.

Other cases

Consider now the other cases, we have:

$$B = \frac{\nu}{\gamma_{I\backslash M}}\left(\frac{\alpha_x}{(1 + (1-\rho)\rho B^2\alpha_\varepsilon/\alpha_x)}\right) \\ + \frac{1-\nu}{\gamma_M}\left(\left(\frac{1}{\alpha_x} + \frac{1}{\alpha_\theta}\right)\left(1 + \frac{B^2\rho\alpha_\varepsilon}{\alpha_x}\right) - \frac{1}{B^2\alpha_\varepsilon\left(1 + \rho\alpha_\theta/\alpha_x\right) + \alpha_\theta}\right)^{-1} \tag{3.12}$$

Then,

$$B \geq \frac{\nu}{\gamma_{I\backslash M}} \left(\frac{\alpha_x}{(1 + (1-\rho)\rho B^2 \alpha_\varepsilon / \alpha_x)} \right)$$

which implies that $B \underset{\alpha_x \to \infty}{\to} \infty$ and $\frac{B}{\sqrt{\alpha_x}} \underset{\alpha_x \to \infty}{\to} \infty$.

Then, as $\frac{1}{\alpha_p} = \frac{\rho}{\alpha_x} + \frac{1}{B^2\alpha_\varepsilon}$, we have

$$\frac{\alpha_x}{\alpha_p} = \rho + \frac{\alpha_x}{B^2 \alpha_\varepsilon}$$

If $\lim\limits_{\alpha_x \to \infty} \rho = l \in (0;1]$, this implies $\lim\limits_{\alpha_x \to \infty} \rho + \frac{\alpha_x}{B^2\alpha_\varepsilon} = l \in (0;1]$ and $\alpha_p \underset{\alpha_x \to +\infty}{\sim} \alpha_x/\rho$;

If $\lim\limits_{\alpha_x \to \infty} \rho = 0$, this implies $\lim\limits_{\alpha_x \to \infty} \rho + \frac{\alpha_x}{B^2\alpha_\varepsilon} = 0 \in (0;1]$ and $\alpha_x \underset{\alpha_x \to \infty}{=} o(\alpha_p)$;

Moreover, some calculations show

$$\left(\left(\frac{1}{\alpha_x} + \frac{1}{\alpha_\theta} \right) \left(1 + \frac{B^2 \rho \alpha_\varepsilon}{\alpha_x} \right) - \frac{1}{B^2 \alpha_\varepsilon \left(1 + \rho\alpha_\theta/\alpha_x\right) + \alpha_\theta} \right)^{-1} \underset{\alpha_x \to \infty}{=} o\left(\frac{\alpha_x}{(1 + (1-\rho)\rho B^2 \alpha_\varepsilon / \alpha_x)} \right)$$

which implies

$$B \underset{\alpha_x \to \infty}{\sim} \frac{\nu}{\gamma_{I\backslash M}} \left(\frac{\alpha_x}{(1 + (1-\rho)\rho B^2 \alpha_\varepsilon / \alpha_x)} \right).$$

Hence, for α_x large enough

$$\begin{aligned} B &\leq 2 \frac{\nu}{\gamma_{I\backslash M}} \left(\frac{\alpha_x}{(1 + (1-\rho)\rho B^2 \alpha_\varepsilon / \alpha_x)} \right) \\ &\leq 2 \frac{\nu \alpha_x}{\gamma_{I\backslash M}} \end{aligned}$$

Then, for α_x large enough

$$
\begin{aligned}
\frac{\alpha_x^2}{\alpha_p} &= \rho\alpha_x + \frac{\alpha_x^2}{B^2\alpha_\varepsilon} \\
&\geq \frac{\alpha_x^2}{B^2\alpha_\varepsilon} \\
&\geq \frac{\gamma_{I\setminus M}^2}{(2\nu)^2\,\alpha_\varepsilon}
\end{aligned}
$$

and $\alpha_p \underset{\alpha_x\to\infty}{=} O(\alpha_x^2)$.

In cases where $\lim\limits_{\alpha_x\to\infty}\rho\alpha_x$ is well defined we can go a little bit further.

We also have:

$$
B\sqrt{\frac{(1-\rho)\rho\alpha_\varepsilon}{\alpha_x}}\left(1+\frac{(1-\rho)\rho B^2\alpha_\varepsilon}{\alpha_x}\right) \underset{\alpha_x\to\infty}{\sim} \frac{\nu\sqrt{(1-\rho)\rho\alpha_x\alpha_\varepsilon}}{\gamma_{I\setminus M}}
$$

1. If $\rho\alpha_x \underset{\alpha_x\to\infty}{\to} \infty$, we have

$$
B\sqrt{\frac{(1-\rho)\rho\alpha_\varepsilon}{\alpha_x}} \underset{\alpha_x\to\infty}{\to} \infty
$$

and

$$
B(\alpha_x) \underset{\alpha_x\to\infty}{\sim} \left(\frac{\nu\alpha_x^2/\gamma_{I\setminus M}}{(1-\rho)\rho\alpha_\varepsilon}\right)^{1/3}
$$

In this case, as $B\sqrt{\frac{(1-\rho)\rho\alpha_\varepsilon}{\alpha_x}} \underset{\alpha_x\to\infty}{\to} \infty$, we have $\frac{B^2\rho\alpha_\varepsilon}{\alpha_x} \underset{\alpha_x\to\infty}{\to} \infty$ and $\alpha_p^{-1} \underset{\alpha_x\to\infty}{\sim} \rho\alpha_x^{-1}$.

2. If $\rho\alpha_x \underset{\alpha_x\to\infty}{\to} 0$ then $B\sqrt{\frac{(1-\rho)\rho\alpha_\varepsilon}{\alpha_x}} \underset{\alpha_x\to\infty}{\to} 0$ and $B(\alpha_x) \underset{\alpha_x\to\infty}{\sim} \frac{\nu\alpha_x}{\gamma_{I\setminus M}}$, so that

$$
\frac{\rho B^2\alpha_\varepsilon}{\alpha_x} \underset{\alpha_x\to\infty}{\sim} \rho\alpha_x\alpha_\varepsilon\left(\frac{\nu}{\gamma_{I\setminus M}}\right)^2.
$$

Then

$$
\frac{\rho B^2\alpha_\varepsilon}{\alpha_x} \underset{\alpha_x\to\infty}{\to} 0
$$

and

$$\alpha_p \underset{\alpha_x \to \infty}{\sim} \left(\frac{\nu \alpha_x}{\gamma_{I \setminus M}}\right)^2 \alpha_\varepsilon$$

3. If

$$\rho \alpha_x \underset{\alpha_x \to \infty}{\rightarrow} M$$

with $M \neq 0$

In this case

$$B\sqrt{\frac{(1-\rho)\rho\alpha_\varepsilon}{\alpha_x}}\left(1 + \frac{(1-\rho)\rho B^2 \alpha_\varepsilon}{\alpha_x}\right) \underset{\alpha_x \to \infty}{\rightarrow} \frac{\nu\sqrt{M\alpha_\varepsilon}}{\gamma_{I \setminus M}}$$

Then $B\sqrt{\frac{(1-\rho)\rho\alpha_\varepsilon}{\alpha_x}}$ has a non zero limit, denoted M'.

And

$$B(\alpha_x) \underset{\alpha_x \to \infty}{\sim} \sqrt{\frac{\alpha_x}{(1-\rho)\rho\alpha_\varepsilon}} M'$$

Then

$$\frac{\rho B^2 \alpha_\varepsilon}{\alpha_x} \underset{\alpha_x \to \infty}{\sim} \left(M'\right)^2$$

and

$$\alpha_p \underset{\alpha_x \to \infty}{\sim} \frac{(\alpha_x)^2}{M} \frac{1}{1 + (M')^{-2}}$$

3.C Proof of Proposition 2b

Proof of (b)

From Lemma 2 the risk aversion coefficient γ_i of the marginal informed trader satisfies $\mathrm{e}^{-2c\gamma_i}\mathbb{V}[\theta-p|x_i]=\mathbb{V}[\theta-p|x_i,p]$. Let us define the function $f:[0,1]\times\mathbb{R}_+^*\to\mathbb{R}$ as follows:

$$f(i,\alpha_x)=\mathrm{e}^{-2c\gamma_i}-\mathbb{V}[\theta-p|x_i,p]/\mathbb{V}[\theta-p|x_i],$$

so that a root $i^*(\alpha_x)$ of $f(\cdot,\alpha_x)$ defines the marginal trader. From (A2) and (A4) we get:

$$\frac{\mathbb{V}[\theta-p|x_i,p]}{\mathbb{V}[\theta-p|x_i]}=\frac{\frac{1}{\alpha_\theta}\left(1-\left(1+\frac{\rho\alpha_\theta}{\alpha_x}+\frac{\alpha_\theta}{\frac{\alpha_x}{1-\rho}+[B(i,\alpha_x)]^2\alpha_\varepsilon}\right)^{-1}\right)\left(1+\frac{\rho\alpha_\theta}{\alpha_x}+\frac{\alpha_\theta}{[B(i,\alpha_x)]^2\alpha_\varepsilon}\right)^2}{\frac{1}{\alpha_x+\alpha_\theta}\left(\rho+\frac{\rho\alpha_\theta}{\alpha_x}+\frac{\alpha_\theta}{[B(i,\alpha_x)]^2\alpha_\varepsilon}\right)^2+\frac{\rho(1-\rho)}{\alpha_x}+\frac{1}{[B(i,\alpha_x)]^2\alpha_\varepsilon}},\tag{D1}$$

where $B(i,\alpha_x)$ is defined as the unique solution to:

$$\begin{aligned}B(i,\alpha_x)=&\left(\frac{\alpha_x}{1+(1-\rho)\rho\,[B(i,\alpha_x)]^2\,\alpha_\varepsilon/\alpha_x}\right)\int_0^i\gamma_i^{-1}\mathrm{d}i\\&+\left(\left(\frac{1}{\alpha_x}+\frac{1}{\alpha_\theta}\right)\left(1+\frac{[B(i,\alpha_x)]^2\rho\alpha_\varepsilon}{\alpha_x}\right)-\frac{1}{[B(i,\alpha_x)]^2\alpha_\varepsilon(1+\rho\alpha_\theta/\alpha_x)+\alpha_\theta}\right)^{-1}\int_i^1\gamma_i^{-1}\mathrm{d}i.\end{aligned}$$

We show that for α_x sufficiently large the equilibrium cannot be such that $I=M$ or $I=I\backslash M$. We know that $\lim_{\alpha_x\to+\infty}B(0,\alpha_x)$ is positive and finite while $\lim_{\alpha_x\to+\infty}B(1,\alpha_x)=+\infty$. In this case,

$$\begin{aligned}\frac{\mathbb{V}[\theta-p|x_i,p]}{\mathbb{V}[\theta-p|x_i]}&=\frac{\frac{1}{\alpha_\theta}\left(1-\left(1+\frac{\rho\alpha_\theta}{\alpha_x}+\frac{\alpha_\theta}{\frac{\alpha_x}{1-\rho}+[B(1,\alpha_x)]^2\alpha_\varepsilon}\right)^{-1}\right)\left(1+\frac{\rho\alpha_\theta}{\alpha_x}+\frac{\alpha_\theta}{[B(1,\alpha_x)]^2\alpha_\varepsilon}\right)^2}{\frac{1}{\alpha_x+\alpha_\theta}\left(\rho+\frac{\rho\alpha_\theta}{\alpha_x}+\frac{\alpha_\theta}{[B(1,\alpha_x)]^2\alpha_\varepsilon}\right)^2+\frac{\rho(1-\rho)}{\alpha_x}+\frac{1}{[B(1,\alpha_x)]^2\alpha_\varepsilon}}\\&\underset{\alpha_x\to\infty}{\sim}\frac{\frac{\rho}{\alpha_x}+\frac{1}{\frac{\alpha_x}{1-\rho}+[B(1,\alpha_x)]^2\alpha_\varepsilon}}{\frac{\rho}{\alpha_x}+\frac{1}{[B(1,\alpha_x)]^2\alpha_\varepsilon}}\end{aligned}$$

With

$$\frac{\frac{\rho}{\alpha_x}+\frac{1}{\frac{\alpha_x}{1-\rho}+[B(1,\alpha_x)]^2\alpha_\varepsilon}}{\frac{\rho}{\alpha_x}+\frac{1}{[B(1,\alpha_x)]^2\alpha_\varepsilon}}=1-\frac{1}{\left(1+\frac{[B(1,\alpha_x)]^2\alpha_\varepsilon\rho}{\alpha_x}\right)\left(1+\frac{[B(1,\alpha_x)]^2\alpha_\varepsilon(1-\rho)}{\alpha_x}\right)}$$

We have

$$0\leq\frac{1}{\left(1+\frac{[B(1,\alpha_x)]^2\alpha_\varepsilon\rho}{\alpha_x}\right)\left(1+\frac{[B(1,\alpha_x)]^2\alpha_\varepsilon(1-\rho)}{\alpha_x}\right)}\leq\frac{1}{\frac{[B(1,\alpha_x)]^2\alpha_\varepsilon}{\alpha_x}}\underset{\alpha_x\to\infty}{\to}0$$

Finally,

$$\frac{\mathbb{V}[\theta-p|x_i,p]}{\mathbb{V}[\theta-p|x_i]}\underset{\alpha_x\to\infty}{\to}1$$

when $I\backslash M=I$. This in turn implies that $f(0,\alpha_x)\underset{\alpha_x\to\infty}{\to}\mathrm{e}^{-2c\gamma_0}>0$ while $f(1,\alpha_x)\underset{\alpha_x\to\infty}{\to}\mathrm{e}^{-2c\gamma_1}-1<0$. Thus, by the intermediate value theorem, for α_x sufficiently large $f(\cdot,\alpha_x)$ has at least one interior root i^* (but none at the corners).

Proof of (c)

Let us first prove that the set of roots of $f(.,\alpha_x)$ is compact when α_x is sufficiently large. To this purpose, we extend $f(.,\alpha_x)$ to $\mathbb{R}\times\mathbb{R}_+^*$ and continuously extend $f(.,\alpha_x)$ on $\mathbb{R}$ by means of the following $f_{\alpha_x}^{ext}$ function:

$$f_{\alpha_x}^{ext}(i)=\mathbf{1}_{i\in[0;1]}f(i,\alpha_x)+\mathbf{1}_{i>1}f(1,\alpha_x)+\mathbf{1}_{i<0}f(0,\alpha_x).$$

For α_x sufficiently large, the set of roots of $f_{\alpha_x}^{ext}$ is a compact set of $\mathbb{R}$: it is bounded, because it is included in $(0,1)$; it is also closed, since it is the reciprocal image of a closed subset of $\mathbb{R}$ (namely, $\{0\}$) by a continuous function. The set of roots being compact, for each α_x we can define a maximum and a minimum solution, which we denote $i^*_{\max}(\alpha_x)$ and $i^*_{\min}(\alpha_x)$ Next, we show by contradiction that $\forall\epsilon>0,\exists A,\forall\alpha_x,[\alpha_x\geq A\Rightarrow\forall i^*(\alpha_x)\in\left(f_{\alpha_x}^{ext}\right)^{-1}(\{0\}),0<i^*(\alpha_x)<\epsilon]$.

Assume that, $\exists \epsilon > 0, \forall A, \exists \alpha_x, \left[\alpha_x \geq A \wedge \exists i^*(\alpha_x) \in \left(f_{\alpha_x}^{ext}\right)^{-1}(\{0\}), \epsilon \leq i^*(\alpha_x) < 1\right]$. Take such an ϵ and consider an associated series $\alpha_x(n) = \alpha_n$ such that $\forall n$, $\alpha_x(n) \geq n$, and the series $i^*(\alpha_n) \in \left(f_{\alpha_x}^{ext}\right)^{-1}(\{0\})$, with $\epsilon \leq i^*(\alpha_x) < 1$.

We will start showing that

$$B_n = B\left(i^*(\alpha_n), \alpha_n\right) \underset{n\to\infty}{\to} \infty$$

First, we have

$$B_n \geq \left(\frac{\alpha_n}{1 + (1-\rho)\rho\left[B_n\right]^2 \alpha_\varepsilon/\alpha_n}\right)\int_0^{i^*(\alpha_n)} \gamma_i^{-1}\mathrm{d}i \geq \left(\frac{\alpha_n}{1 + (1-\rho)\rho\left[B_n\right]^2 \alpha_\varepsilon/\alpha_n}\right)\int_0^{\epsilon} \gamma_i^{-1}\mathrm{d}i$$

Now, since

$$\frac{B_n}{\sqrt{\alpha_n}}\left(1 + \frac{\alpha_\varepsilon}{4}\frac{\left[B_n\right]^2}{\alpha_n}\right) \geq \int_0^{\epsilon} \gamma_i^{-1}\mathrm{d}i\sqrt{\alpha_n},$$

this implies

$$\frac{B_n}{\sqrt{\alpha_n}} \underset{n\to\infty}{\to} \infty$$

and, in particular,

$$B_n \underset{n\to\infty}{\to} \infty$$

This implies once again that:

$$\begin{aligned}\frac{\mathbb{V}[\theta - p|x_i, p]}{\mathbb{V}[\theta - p|x_i]} &= \frac{\frac{1}{\alpha_\theta}\left(1 - \left(1 + \frac{\rho\alpha_\theta}{\alpha_n} + \frac{\alpha_\theta}{\frac{\alpha_n}{1-\rho} + B_n^2\alpha_\varepsilon}\right)^{-1}\right)\left(1 + \frac{\rho\alpha_\theta}{\alpha_n} + \frac{\alpha_\theta}{B_n^2\alpha_\varepsilon}\right)^2}{\frac{1}{\alpha_n + \alpha_\theta}\left(\rho + \frac{\rho\alpha_\theta}{\alpha_n} + \frac{\alpha_\theta}{B_n^2\alpha_\varepsilon}\right)^2 + \frac{\rho(1-\rho)}{\alpha_n} + \frac{1}{B_n^2\alpha_\varepsilon}} \\ &\underset{n\to\infty}{\sim} 1 - \frac{1}{\left(1 + \frac{B_n^2\alpha_\varepsilon\rho}{\alpha_n}\right)\left(1 + \frac{B_n^2\alpha_\varepsilon(1-\rho)}{\alpha_n}\right)}.\end{aligned}$$

Again, given that $\frac{B_n}{\sqrt{\alpha_n}} \underset{n\to\infty}{\to} \infty$, we have

$$1 - \frac{1}{\left(1+\frac{B_n^2\alpha_\varepsilon\rho}{\alpha_n}\right)\left(1+\frac{B_n^2\alpha_\varepsilon(1-\rho)}{\alpha_n}\right)} \underset{n\to\infty}{\to} 1.$$

However, by the definition of $i^*(\alpha_n)$ we also have

$$\exp[-2c\gamma_{i^*(\alpha_n)}] \underset{n\to\infty}{\to} 1,$$

which is impossible given that for all n $\exp[-2c\gamma_{i^*(\alpha_n)}] \leq \exp[-2c\gamma_\epsilon] < 1$.

We have showed that $\forall\epsilon > 0, \exists A, \forall\alpha_x, [\alpha_x \geq A \Rightarrow \forall i^*(\alpha_x) \in \left(f_{\alpha_x}^{ext}\right)^{-1}(\{0\}), 0 < i^*(\alpha_x) < \epsilon]$. This implies that, for any choice function $\hat{\imath} : \alpha_x \to \hat{\imath}(\alpha_x) \in \left(f_{\alpha_x}^{ext}\right)^{-1}(\{0\})$, we have $\hat{\imath}(\alpha_x) \underset{\alpha_x\to\infty}{\to} 0$; in particular, $i^*_{\max}(\alpha_x) \underset{\alpha_x\to\infty}{\to} 0$ and $i^*_{\min}(\alpha_x) \underset{\alpha_x\to\infty}{\to} 0$.

Proof of (d)

For any choice function $\hat{\imath}(\alpha_x)$, since $f(\hat{\imath}(\alpha_x), \alpha_x) = 0$ and $\hat{\imath}(\alpha_x) \underset{\alpha_x\to\infty}{\to} 0$ we have, by continuity,

$$\Xi \equiv \frac{\frac{1}{\alpha_\theta}\left(1-\left(1+\frac{\rho\alpha_\theta}{\alpha_x}+\frac{\alpha_\theta(1-\rho)}{\alpha_x+[B(\hat{\imath}(\alpha_x),\alpha_x)]^2\alpha_\varepsilon(1-\rho)}\right)^{-1}\right)\left(1+\frac{\rho\alpha_\theta}{\alpha_x}+\frac{\alpha_\theta}{[B(\hat{\imath}(\alpha_x),\alpha_x)]^2\alpha_\varepsilon}\right)^2}{\left(\frac{1}{\alpha_x+\alpha_\theta}\left(\rho+\frac{\rho\alpha_\theta}{\alpha_x}+\frac{\alpha_\theta}{B(\hat{\imath}(\alpha_x),\alpha_x)^2\alpha_\varepsilon}\right)^2+\frac{(1-\rho)\rho}{\alpha_x}+\frac{\alpha_\varepsilon^{-1}}{[B(\hat{\imath}(\alpha_x),\alpha_x)]^2}\right)} \underset{\alpha_x\to\infty}{\to} \mathrm{e}^{-2\gamma_0 c} \tag{D1}$$

We first show by contradiction that $B(\hat{\imath}(\alpha_x), \alpha_x) \underset{\alpha_x\to\infty}{\to} \infty$, that is, $\forall M > 0, \exists A > 0, \forall\alpha_x, [\alpha_x \geq A \Rightarrow B(\hat{\imath}(\alpha_x), \alpha_x) > M]$. If it were not the case, we could find $M > 0$ such that, $\forall A > 0$, there would exist an α_x satisfying $\alpha_x \geq A$ and $B(\hat{\imath}(\alpha_x)) \leq M$. Choose any such M and consider the serie $\alpha_n = (\alpha_x(n))_{n\geq n_0}$ with n_0 large enough to ensure the existence of $(\hat{\imath}(\alpha_n))_{n\geq n_0}$. In this case,

$$\frac{1}{\alpha_\theta}\left(1-\left(1+\frac{\rho\alpha_\theta}{\alpha_n}+\frac{\alpha_\theta(1-\rho)}{\alpha_n+[B(\hat{\imath}(\alpha_n),\alpha_n)]^2\alpha_\varepsilon(1-\rho)}\right)^{-1}\right) \underset{n\to\infty}{\to} 0,$$

while

$$0 \leq \left(1+\frac{\rho\alpha_\theta}{\alpha_n}+\frac{\alpha_\theta}{B\left(\hat{\imath}(\alpha_n),\alpha_n\right)^2\alpha_\varepsilon}\right)^2 \leq \left(1+\frac{\rho\alpha_\theta}{\alpha_n}+\frac{\alpha_\theta}{B(0,\alpha_n)^2\alpha_\varepsilon}\right)^2,$$

where the right hand side is bounded above as $n \to \infty$. Moreover,

$$\left(\frac{1}{\alpha_n+\alpha_\theta}\left(\rho+\frac{\rho\alpha_\theta}{\alpha_n}+\frac{\alpha_\theta}{B\left(\hat{\imath}(\alpha_n)\right)^2\alpha_\varepsilon}\right)^2+\frac{(1-\rho)\rho}{\alpha_n}+\frac{\alpha_\varepsilon^{-1}}{B\left(\hat{\imath}(\alpha_n)\right)^2}\right)$$
$$\geq \left(\frac{1}{\alpha_n+\alpha_\theta}\left(\rho+\frac{\rho\alpha_\theta}{\alpha_n}+\frac{\alpha_\theta}{M^2\alpha_\varepsilon}\right)^2+\frac{(1-\rho)\rho}{\alpha_n}+\frac{1}{M^2\alpha_\varepsilon}\right),$$

which is bounded below as $n \to \infty$. Hence, we would have $\Xi \underset{n\to\infty}{\to} 0$. However, since the limit is unique this would require $e^{-2\gamma_0 c} = 0$, a contradiction. It must thus be that $B(\hat{\imath}(\alpha_x)) \underset{\alpha_x\to\infty}{\to} \infty$, which in turn implies:

$$\left(1-\left(1+\frac{\rho\alpha_\theta}{\alpha_x}+\frac{\alpha_\theta(1-\rho)}{\alpha_x+\left[B(\hat{\imath}(\alpha_x),\alpha_x)\right]^2\alpha_\varepsilon(1-\rho)}\right)^{-1}\right) \underset{\alpha_x\to\infty}{\sim} \frac{\rho\alpha_\theta}{\alpha_x}+\frac{\alpha_\theta(1-\rho)}{\alpha_x+\left[B(\hat{\imath}(\alpha_x),\alpha_x)\right]^2\alpha_\varepsilon(1-\rho)},$$

and
$$\left(\frac{1}{\alpha_x+\alpha_\theta}\left(\rho+\frac{\rho\alpha_\theta}{\alpha_x}+\frac{\alpha_\theta}{\left[B(\hat{\imath}(\alpha_x),\alpha_x)\right]^2\alpha_\varepsilon}\right)^2+\frac{(1-\rho)\rho}{\alpha_x}+\frac{1}{\left[B(\hat{\imath}(\alpha_x),\alpha_x)\right]^2\alpha_\varepsilon}\right)$$
$$\underset{\alpha_x\to\infty}{\sim} \frac{\rho^2}{\alpha_x}+\frac{(1-\rho)\rho}{\alpha_x}+\frac{1}{\left[B(\hat{\imath}(\alpha_x),\alpha_x)\right]^2\alpha_\varepsilon} = \frac{\rho}{\alpha_x}+\frac{1}{\left[B(\hat{\imath}(\alpha_x),\alpha_x)\right]^2\alpha_\varepsilon}$$

so that:

$$\Xi \underset{\alpha_x\to\infty}{\sim} \frac{\frac{1}{\alpha_\theta}\left(\frac{\rho\alpha_\theta}{\alpha_x}+\frac{\alpha_\theta(1-\rho)}{\alpha_x+\left[B(\hat{\imath}(\alpha_x),\alpha_x)\right]^2\alpha_\varepsilon(1-\rho)}\right)}{\frac{\rho}{\alpha_x}+\frac{1}{\left[B(\hat{\imath}(\alpha_x),\alpha_x)\right]^2\alpha_\varepsilon}} = \frac{\frac{\rho}{\alpha_x}+\frac{(1-\rho)}{\alpha_x+\left[B(\hat{\imath}(\alpha_x),\alpha_x)\right]^2\alpha_\varepsilon(1-\rho)}}{\frac{\rho}{\alpha_x}+\frac{1}{\left[B(\hat{\imath}(\alpha_x),\alpha_x)\right]^2\alpha_\varepsilon}} = \frac{\rho+\frac{1-\rho}{1+\frac{\left[B(\hat{\imath}(\alpha_x),\alpha_x)\right]^2\alpha_\varepsilon}{\alpha_x}(1-\rho)}}{\rho+\frac{\alpha_x}{\left[B(\hat{\imath}(\alpha_x),\alpha_x)\right]^2\alpha_\varepsilon}}.$$

We know from (D1) that $\Xi \underset{\alpha_x\to\infty}{\to} e^{-2\gamma_0 c}$, hence

$$\left(\rho+\frac{1-\rho}{1+\frac{\left[B(\hat{\imath}(\alpha_x),\alpha_x)\right]^2\alpha_\varepsilon}{\alpha_x}(1-\rho)}\right)\left(\rho+\frac{\alpha_x}{\left[B(\hat{\imath}(\alpha_x),\alpha_x)\right]^2\alpha_\varepsilon}\right)^{-1} \underset{\alpha_x\to\infty}{\to} e^{-2\gamma_0 c}.$$

Now, define the following function:

$$h_\rho : x \to \left(\rho + \frac{(1-\rho)}{1+x(1-\rho)}\right)\left(\rho + x^{-1}\right)^{-1},$$

which is continuous, increasing on $\mathbb{R}^+$ and maps $\mathbb{R}^+$ onto $(0,1)$. This function is then a bijection, and its reciprocal is defined as follows on $(0,1)$:

$$h_\rho^{-1} : x \to \frac{\sqrt{1+4\rho(1-\rho)\frac{x}{1-x}}-1}{2\rho(1-\rho)}$$

if $\rho \neq 0,1$ and

$$h_0^{-1} = h_1^{-1} : x \to \frac{x}{1-x}$$

We then have for all α_x

$$\frac{\left[B(\hat{\imath}(\alpha_x),\alpha_x)\right]^2\alpha_\varepsilon}{\alpha_x} = h_\rho^{-1}\left(h_\rho\left(\frac{\left[B(\hat{\imath}(\alpha_x),\alpha_x)\right]^2\alpha_\varepsilon}{\alpha_x}\right)\right),$$

with

$$h_\rho\left(\frac{\left[B(\hat{\imath}(\alpha_x),\alpha_x)\right]^2\alpha_\varepsilon}{\alpha_x}\right) \underset{\alpha_x\to\infty}{\to} \mathrm{e}^{-2\gamma_0 c}.$$

Then, the function

$$\frac{h_\rho\left(\frac{\left[B(\hat{\imath}(\alpha_x),\alpha_x)\right]^2\alpha_\varepsilon}{\alpha_x}\right)}{1-h_\rho\left(\frac{\left[B(\hat{\imath}(\alpha_x),\alpha_x)\right]^2\alpha_\varepsilon}{\alpha_x}\right)}$$

is well defined for α_x large enough and

$$\frac{h_\rho\left(\frac{\left[B(\hat{\imath}(\alpha_x),\alpha_x)\right]^2\alpha_\varepsilon}{\alpha_x}\right)}{1-h_\rho\left(\frac{\left[B(\hat{\imath}(\alpha_x),\alpha_x)\right]^2\alpha_\varepsilon}{\alpha_x}\right)} \underset{\alpha_x\to\infty}{\to} \frac{\mathrm{e}^{-2\gamma_0 c}}{1-\mathrm{e}^{-2\gamma_0 c}}.$$

Let us now denote l the limit of ρ as $\alpha_x \to \infty$. If $l \in (0;1)$, then

$$h_\rho^{-1}\left(h_\rho\left(\frac{[B(\hat{\imath}(\alpha_x),\alpha_x)]^2\,\alpha_\varepsilon}{\alpha_x}\right)\right) \underset{\alpha_x\to\infty}{\to} \frac{\sqrt{1+4l\,(1-l)\,\frac{\mathrm{e}^{-2\gamma_0 c}}{1-\mathrm{e}^{-2\gamma_0 c}}}-1}{2l\,(1-l)} = h_l^{-1}\left(\mathrm{e}^{-2\gamma_0 c}\right),$$

and

$$\frac{[B(\hat{\imath}(\alpha_x),\alpha_x)]^2\,\alpha_\varepsilon}{\alpha_x} \underset{\alpha_x\to\infty}{\to} h_l^{-1}(\mathrm{e}^{-2\gamma_0 c}).$$

If $l \in \{0;1\}$, then

$$4\rho\,(1-\rho)\,\frac{h_\rho\left(\frac{[B(\hat{\imath}(\alpha_x),\alpha_x)]^2\alpha_\varepsilon}{\alpha_x}\right)}{1-h_\rho\left(\frac{[B(\hat{\imath}(\alpha_x),\alpha_x)]^2\alpha_\varepsilon}{\alpha_x}\right)} \underset{\alpha_x\to\infty}{\to} 0,$$

and hence:

$$\frac{\sqrt{1+4\rho\,(1-\rho)\,\frac{h_\rho\left(\frac{[B(\hat{\imath}(\alpha_x),\alpha_x)]^2\alpha_\varepsilon}{\alpha_x}\right)}{1-h_\rho\left(\frac{[B(\hat{\imath}(\alpha_x),\alpha_x)]^2\alpha_\varepsilon}{\alpha_x}\right)}}-1}{2\rho\,(1-\rho)} \underset{\alpha_x\to\infty}{\sim} \frac{\mathrm{e}^{-2\gamma_0 c}}{1-\mathrm{e}^{-2\gamma_0 c}} = h_l^{-1}(\mathrm{e}^{-2\gamma_0 c}).$$

Finally,

$$\frac{[B(\hat{\imath}(\alpha_x),\alpha_x)]^2\,\alpha_\varepsilon}{\alpha_x} \underset{\alpha_x\to\infty}{\to} h_l^{-1}(\mathrm{e}^{-2\gamma_0 c}),$$

and thus

$$B(\hat{\imath}(\alpha_x),\alpha_x) \underset{\alpha_x\to\infty}{\sim} \left(\frac{\alpha_x}{\alpha_\varepsilon}h_l^{-1}(\exp[-2\gamma_0 c])\right)^{1/2}.$$

We infer from the latter expression as well as (B1) above that

$$\alpha_p = \frac{\alpha_x}{\rho+\frac{\alpha_x}{[B(\hat{\imath}(\alpha_x),\alpha_x)]^2\alpha_\varepsilon}} \underset{\alpha_x\to\infty}{\sim} \frac{\alpha_x}{\lim\limits_{\alpha_x\to\infty}\rho+\left[h^{-1}_{\lim\limits_{\alpha_x\to\infty}\rho}(\mathrm{e}^{-2\gamma_0 c})\right]^{-1}}.$$

Proof of (a)

We now show that i^* (the marginal trader) is unique for α_x sufficiently large. To this purpose, consider the differentiable functions of B defined as follows:

$$u(B) = \frac{1}{\alpha_\theta}\left(1 - \left(1 + \rho\frac{\alpha_\theta}{\alpha_x} + \frac{(1-\rho)\,\alpha_\theta}{\alpha_x + B^2\alpha_\varepsilon(1-\rho)}\right)^{-1}\right)\left(1 + \frac{\rho\alpha_\theta}{\alpha_x} + \frac{\alpha_\theta}{B^2\alpha_\varepsilon}\right)^2,$$
$$d(B) = \left(\frac{1}{\alpha_x + \alpha_\theta}\left(\rho + \frac{\rho\alpha_\theta}{\alpha_x} + \frac{\alpha_\theta}{B^2\alpha_\varepsilon}\right)^2 + \frac{(1-\rho)\rho}{\alpha_x} + \frac{1}{B^2\alpha_\varepsilon}\right),\; r(B) = \frac{u(B)}{d(B)}.$$

For all $i \in [i^*_{\min}(\alpha_x), i^*_{\max}(\alpha_x)]$ we have, after some manipulations:

$$r'(B(i)) = \frac{u'(B(i))}{d(B(i))} - \frac{d'(B(i))}{d^2(B(i))}u(B(i)) \geq \lambda_1(\alpha_x) + \lambda_2(\alpha_x),$$

where $\lambda_1(\alpha_x)$ and $\lambda_2(\alpha_x)$ are given by:

$$\lambda_1(\alpha_x) \equiv -\frac{2\overline{B}\alpha_\varepsilon\left(\frac{(1-\rho)^2\left(1+\rho\frac{\alpha_\theta}{\alpha_x}+\frac{\alpha_\theta}{\underline{B}^2\alpha_\varepsilon}\right)^2}{\left(\left(1+\frac{\rho\alpha_\theta}{\alpha_x}\right)\left(\alpha_x+(1-\rho)\underline{B}^2\alpha_\varepsilon\right)+(1-\rho)\alpha_\theta\right)^2} + \frac{2\left(1+\frac{\rho\alpha_\theta}{\alpha_x}+\frac{\alpha_\theta}{\underline{B}^2\alpha_\varepsilon}\right)}{\left(\underline{B}^2\alpha_\varepsilon\right)^2}\left(1-\left(1+\frac{\rho\alpha_\theta}{\alpha_x}+\frac{(1-\rho)\alpha_\theta}{\alpha_x+\underline{B}^2\alpha_\varepsilon(1-\rho)}\right)^{-1}\right)\right)}{\frac{1}{\alpha_x+\alpha_\theta}\left(\rho+\rho\frac{\alpha_\theta}{\alpha_x}+\frac{\alpha_\theta}{\overline{B}^2\alpha_\varepsilon}\right)^2+(1-\rho)\rho\frac{1}{\alpha_x}+\frac{\alpha_\varepsilon^{-1}}{\overline{B}^2}}$$

and

$$\lambda_2(\alpha_x) \equiv$$
$$\frac{\left(\frac{2}{\alpha_x+\alpha_\theta}\left(\rho+\rho\frac{\alpha_\theta}{\alpha_x}+\frac{\alpha_\theta}{\overline{B}^2\alpha_\varepsilon}\right)\left(\frac{2}{\overline{B}^3\alpha_\varepsilon}\right)+\frac{2}{\overline{B}^3\alpha_\varepsilon\alpha_\theta}\right)\left(1-\left(1+\frac{\rho\alpha_\theta}{\alpha_x}+\frac{(1-\rho)\alpha_\theta}{\alpha_x+\overline{B}^2\alpha_\varepsilon(1-\rho)}\right)^{-1}\right)\left(1+\frac{\rho\alpha_\theta}{\alpha_x}+\frac{\alpha_\theta}{\overline{B}^2\alpha_\varepsilon}\right)^2}{\left(\frac{1}{\alpha_x+\alpha_\theta}\left(\rho+\frac{\rho\alpha_\theta}{\alpha_x}+\frac{\alpha_\theta}{\underline{B}^2\alpha_\varepsilon}\right)^2+\frac{(1-\rho)\rho}{\alpha_x}+\frac{1}{\underline{B}^2\alpha_\varepsilon}\right)^2}$$

and where we have defined $\underline{B} \equiv B(i^*_{\min}(\alpha_x), \alpha_x)$ and $\overline{B} \equiv B(i^*_{\max}(\alpha_x), \alpha_x)$.

Two cases must be considered. In the first case, we assume that $\rho \underset{\alpha_x\to\infty}{\to} 1$. Then, it can be shown that:

$$-\frac{\lambda_1(\alpha_x)}{\lambda_2(\alpha_x)} \underset{\alpha_x\to\infty}{\to} 0$$

This yields

$$\lambda_1(\alpha_x) + \lambda_2(\alpha_x) \underset{\alpha_x \to \infty}{\sim} \lambda_2(\alpha_x).$$

It follows that for α_x sufficiently large, and for any $i \in [i^*_{\min}(\alpha_x), i^*_{\max}(\alpha_x)]$ we have $r'(B(i)) \geq \lambda_2(\alpha_x)/2 > 0$. It can be shown that $\partial B(i, \alpha_x)/\partial i > 0$, which in turn implies that $\partial r(B)/\partial i = r'(B(i))\partial B/\partial i > 0$. Then, on this range we have $\partial f(i, \alpha_x)/\partial i < 0$, so that $f(i, \alpha_x)$ is a bijection on $[i^*_{\min}(\alpha_x), i^*_{\max}(\alpha_x)]$. As $f(i^*_{\min}(\alpha_x), \alpha_x) = f(i^*_{\max}(\alpha_x), \alpha_x) = 0$, we have $i^*_{\min}(\alpha_x) = i^*_{\max}(\alpha_x)$, i.e., the solution is unique (for α_x sufficiently large).

Let us now consider the case where $\lim_{\alpha_x \to \infty} \rho = l \in [0; 1)$. As for the proof of proposition 2, it can be shown that

$$\lambda_1(\alpha_x) \underset{\alpha_x \to \infty}{\sim} -\frac{2\left(\alpha_\varepsilon h_l^{-1}(\mathrm{e}^{-2\gamma_0 c})\right)^{1/2}}{\left(\frac{1}{(1-l)} + h_l^{-1}(\mathrm{e}^{-2\gamma_0 c})\right)^2 \left(l + \frac{1}{h_l^{-1}(\mathrm{e}^{-2\gamma_0 c})}\right)} \alpha_x^{-1/2},$$

Similarly,

$$\lambda_2(\alpha_x) \underset{\alpha_x \to \infty}{\sim} \frac{2}{\left(h_l^{-1}(\mathrm{e}^{-2\gamma_0 c})\right)^{3/2} \alpha_\varepsilon^{-1/2}} \frac{1}{\left(l + \frac{1}{h_l^{-1}(\mathrm{e}^{-2\gamma_0 c})}\right)^2} \left(l + \frac{1}{\frac{1}{1-l} + h_l^{-1}(\mathrm{e}^{-2\gamma_0 c})}\right) \alpha_x^{-1/2}.$$

Next, define:

$$\mu_1 \equiv \frac{2\left(\alpha_\varepsilon h_l^{-1}(\mathrm{e}^{-2\gamma_0 c})\right)^{1/2}}{\left(\frac{1}{(1-l)} + h_l^{-1}(\mathrm{e}^{-2\gamma_0 c})\right)^2 \left(l + \frac{1}{h_l^{-1}(\mathrm{e}^{-2\gamma_0 c})}\right)}, \; \mu_2 \equiv \frac{2\left(l + \frac{1}{\frac{1}{1-l} + h_\rho^{-1}(\mathrm{e}^{-2\gamma_0 c})}\right)}{\left(h_l^{-1}(\mathrm{e}^{-2\gamma_0 c})\right)^{3/2} \alpha_\varepsilon^{-1/2} \left(l + \frac{1}{h_l^{-1}(\mathrm{e}^{-2\gamma_0 c})}\right)^2}.$$

Since $\mu_2 > \mu_1$, it follows that for α_x sufficiently large, and for any $i \in [i^*_{\min}(\alpha_x), i^*_{\max}(\alpha_x)]$ we have $r'(B(i)) \geq (\mu_2 - \mu_1)\alpha_x^{-1/2}/2 > 0$. It can be shown that $\partial B(i, \alpha_x)/\partial i > 0$, which in turn implies that $\partial r(B)/\partial i = r'(B(i))\partial B/\partial i > 0$. Then, on this range we have $\partial f(i, \alpha_x)/\partial i < 0$, so that $f(i, \alpha_x)$ is a bijection on $[i^*_{\min}(\alpha_x), i^*_{\max}(\alpha_x)]$. As $f(i^*_{\min}(\alpha_x), \alpha_x) =$

$f(i^*_{\max}(\alpha_x), \alpha_x) = 0$, we have $i^*_{\min}(\alpha_x) = i^*_{\max}(\alpha_x)$, i.e., the solution is unique (for α_x sufficiently large).

Bibliography

[1] G.-M., Angeletos, I. Werning, Crises and prices: information aggregation, multiplicity, and volatility, Am. Econ. Rev. 96(5) (2006) 1720–1736.

[2] Y. Bai, E. Chang, J. Wang, Asset prices under short-sale constraints, MIT Working Paper (2006).

[3] G. Barlevy, P. Veronesi, Information acquisition in financial markets, Rev. Econ. Stud. 67 (2000) 79–90.

[4] D.P. Brown, Z.M. Zhang, Market orders and market efficiency, J. Finance 52(1) (1997) 277–308.

[5] D.W. Diamond, R.E. Verrecchia, Information aggregation in a noisy rational expectations economy, J. Financ. Econ. 9 (1981) 231–235.

[6] D.W. Diamond, R.E. Verrecchia, Constraints on short-selling and asset price adjustment to private information, J. of Financ. Econ. 18 (1987) 277–311

[7] S.J. Grossman, J. Stiglitz, On the impossibility of informationally efficient markets, Am. Econ. Rev., 70(3) (1980) 393–408.

[8] S.J. Grossman, J. Stiglitz, Information and competitive price system, Am. Econ. Rev. 66(2) (1976) 246–253.

[9] A.S. Kyle, Informed speculation under imperfect competition, Rev. Econ. Stud., 56(3) (1989) 317–355.

[10] L.A. Medrano, Market versus limit orders in an imperfectly competitive security market, UPF Working paper 165 (1996).

[11] E. Miller, Risk, uncertainty and divergence of opinion, J. Finance, 32 (1977) 1151–1168.

[12] J. Peress, The tradeoff between risk sharing and information production in financial markets, J. Econ. Theory, 145(1) (2004) 124–155.

[13] R.E. Verrecchia, Information acquisition in a noisy rationial expectation economy, Econometrica, 50(6) (1982) 1415–1430.

[14] X. Vives, On the possibility of informationally efficient markets, J. Eur. Econ. Assoc. 12(5) (2014) 1200–1239.

[15] X. Vives, Information and Learning in Markets: The Impact of Market Microstructure, Princeton University Press, Princeton, 2008.

[16] X. Vives, The speed of information revelation in a financial market mechanism, J. Econ. Theory 67(1) (1995) 178–204.

[17] J.K. Wald, H.T. Horrigan, Optimal limit order choice, J. Bus. 78(2) (2005) 597–619.

Chapter

4

Market microstructure, information aggregation and equilibrium uniqueness in a global game

Speculators contemplating an attack (e.g., on a currency peg) must guess the beliefs of other speculators, which they can do by looking at the stock market. This paper examines whether this information-gathering process is stabilizing – by better anchoring expectations – or destabilizing – by creating multiple self-fulfilling equilibria. To do so, we study the outcome of a two-stage global game wherein an asset price determined at the trading stage of the game provides an endogenous public signal about the fundamental that affects traders' decision to attack in the coordination stage of the game. In the trading stage, placing a full demand schedule (i.e., a continuum of limit orders) is costly, but traders may use riskier (and cheaper) market orders, i.e., order to sell or buy a fixed quantity of assets unconditional on the execution price. Price execution risk reduces traders aggressiveness and hence slows down information aggregation, which ultimately makes multiple equilibria in the coordination stage less likely. In this sense, microstructure frictions that lead to greater individual exposure (to price execution risk) may reduce aggregate uncertainty (by pinning down a unique equilibrium outcome).

JEL-code : *C72, D82, G14.*

Keywords: *Market microstructure; Information aggregation;Global game.*

4.1 Introduction

Consider a situation where a country may be subject to a speculative attack – on its currency, or its public debt, its banking sector, etc. – that involves an element of strategic complementarity: the attack is all the more likely to be successful, and hence the prevailing state of affairs to collapse, when the number of speculators who challenge it (i.e., bond traders, carry traders, short term lenders etc.) is large. In this situation, a speculator who is contemplating the option of attacking the prevailing regime must not only evaluate how strong the economy is but also, and even more importantly, how strong it is *perceived to be* by the other speculators. Making such an inference on the beliefs – and likely actions – of others is inherently more challenging than merely forecasting the economy's fundamentals. Crucially, it requires one to rely not only on one's own idiosyncratic assessment of the economic outook, but also on the kind of public information that is visible by all and may guide their actions. The stock market is one the first sources of public information that speculators scrutinise, for the very reason that it encodes information about how the market as a whole perceives the economy's soundness. Does this source of information contribute to stabilise the market, by helping to anchor speculators' expectations, or does it destabilise it, by easing their coordination on *a priori* indeterminate, but ultimately self-fulfilling, outcomes?

To answer this question, we study the equilibrium of a two-stage global game wherein a market-based asset price determined at the trading stage of the game provides an endogenous public signal about the fundamental that affects traders' decision in the coordination stage of the game. One motivation for doing so is to examine the concern, first raised by Atkeson (2001) and then made formal by Angeletos and Werning (2006), that a publicly observed market price may aggregate dispersed information so effectively as to crowd out private signals in traders' assessment of the fundamental, and in so doing facilitate their coordination on a self-fulfilling outcome. As illustrated by Angeletos and Werning (2006), this may precisely occur as the noise in the private signal vanishes, a result that directly challenges Carlsson and van Damme (1993) and Morris and Shin (1998)'s argument that a small perturbation of the full-information coordination game restores equilibrium uniqueness.

The possibility that a small amount of private noise lead to multiplicity rather than uniqueness of equilibrium outcomes arises when the precision of the endogenous public signal grows faster than that of the underlying exogenous private signals at high levels of precision.[1]

This relationship relies heavily on the condition that the market for the asset is very liquid. However, the markets considered are likely to suffer from liquidity issues, which is the case, in particular, for currency markets. The FOREX market is an over-the-counter, decentralized and opaque market. Market microstructure and OTC literatures suggest related frictions that may cause low liquidity in the market; they include information asymetry or search costs (see e.g. Lyons(2001), Geromichalos and Jung(2016)). Moreover, traders in currency markets are likely to suffer from funding constraints, which translates into lower liquidity in the currency markets. Brunnermeier, Nagel and Pedersen (2008) advocate such liquidity spirals, and show for instance that financial crises are associated with unwinding carry trades, leading to liquidity drops. Mancini, Ranaldo and Wrampelmeyer (2013) show that, after Lehman's failure, even the most liquid FX markets suffered from such drops. These liquidity issues are likely to be more stronger in small currencies, that can be used in carry trade strategies. Burnside(2008) argues that liquidity frictions might play a crucial role in explaining the profitability of carry trades. Consistently, among the currencies studied by Karnaukh et al. (2015), they show indeed that AUD, which is involved in typical carry trade strategy, has the least liquid exchange rate with respect to USD.

To take into account potential frictions in the asset market, we introduce market microstructure considerations. We show that the property put forward by Angeletos and Werning (2006) crucially depends on the type of market microstructure and what this

[1] See also Hellwig et al. (2006). Hellwig (2002) emphasised the role of the relative precison of public versus private information in determining the outcome of the game. Angeletos et al. (2006) study global games wherein endogenous public information comes from policy choices rather than an asset price.

microstructure implies for the amount of private information that is aggregated into the asset price.

We substantiate this point by considering a market microstructure for the trading stage wherein informed traders may place either full demand schedules or more basic *market orders*, i.e., order to sell or buy a fixed quantity of assets unconditional on the execution price.[2] All orders (from informed and noise traders) are then aggregated into an asset price by a competitive market-making sector. This price provides the endogenous public signal that informed traders may use to coordinate a speculative attack in the second stage of the game.

To summarise, our results are as follows. In a pure market-order market (see Vives, 1995), the precision of the endogneous public signal provided by the asset price is bounded above, even when the precision of the underlying private signals is very (arbitrarily) large. This is due to the competition of two forces. On the one hand, greater precision leads informed traders to trade more aggressively on their private information by opening the possiblity of reaping large payoffs from trading. On the other hand, this very aggressiveness renders the asset price very volatile ex post (after all market orders have irreversibly been aggregated), which raises the conditional volatility of the net payoff, i.e., the terminal dividend minus the trading price of the asset. The first effect makes the informativeness of the price an increasing function of the precision of private signals. The second effect, however, runs counter the first effect: it deters market-order traders, which are exposed to price execution risk, from placing large orders. As the precision of private information increases the strength of the second effect gradually catches up with that of the first effect and the precision of the price signal increases more and more slowly. This boundedness of the information conveyed by the price overturns the result in Angeletos and Werning (2006), because (endogenous) public information can no longer crowd out (exogenous) private information in traders' Bayesian learning of the fundamental. As a consequence, a high level of precision of private information can again uniquely pin down the outcome

[2] See Brown and Zhang (1997), Wald and Horrigan (2005) and Vives (2008) for further disscussion of the importance of market orders in actual asset markets. In Challe and Chrétien (2015) we study the functionning of a similar asset market under a more general information structure but with no dimension of coordination.

of the coordination game – and we are back to Morris and Shin (1998). When the share of market-order traders is still exogenous but not necessarily equal to one, our result must be qualified in the following sense. While it is again true that as private information becomes infinitely precise then so does public information, just as in the pure demand schedule/Walrasian auctioneer case of Angeletos and Werning (2006), it is nevertheless the case that for large range degrees of precision the uniqueness region can be greatly expanded relative to pure demand schedule case.

We finally examine the case where informed traders can choose their order type ex ante, where the tradeoff is between placing expensive demand schedules or cheap market orders.[3] We notably study the impact of this choice on the equilibrium share of market-order traders and, by way of consequence, on the outcome of the coordination stage. We show that as private noise vanishes the equilibrium is always interior (i.e., market-order and demand-schedule traders are both in positive measure), but market-order traders ultimately overwhelm the market (i.e., their measure tends to one). As a result, the rate of convergence of the precision of the public signal under endogenous order type is half that under exogenous order types. This implies that the endogenous adjustment of the share of market-order traders further reduces the multiplicity region as private noise decreases, relative to the case where this share is exogenous.

The rest of the paper is organised as follows. Section 2 presents two stages of the game. Section 3 analyses the outcome of the game when the shares of market-order and demand-schedule traders are exogenous. Section 4 studies the endogenous determination of those shares, and how this affects the size of the multiplicity versus uniqueness regions. Section 5 concludes the paper. All the proofs appear in the Appendix.

[3] In as much as demand schedules allow full conditionality of trades on the realised trading price, they are much more (in fact, infinitely more) complex that market orders (which are not conditional on the price). Therefore, demand schedules should be more expensive, as we assume them to be.

4.2 The model

Following Angeletos and Werning (2006), we consider a two-stage global game wherein a continuum of informed traders $i \in I = [0,1]$ trades an asset in a *trading stage* before deciding whether to attack the regime in the *coordination stage* –see Figure 1. Before the game starts, an unobserved fundamental θ is drawn from the distribution $\mathcal{N}\left(\bar{\theta}, \alpha_\theta^{-1}\right)$ (which is also the common prior of informed traders) and affects both asset payoffs in the trading stage and the ability of the government to withstand a speculative attack in the coordination stage. Every informed trader gets two noisy signals about θ. First, it gets an exogenous private signal

$$x_i = \theta + \alpha_x^{-1/2}\xi_i,$$

where

$$\alpha_x > 0,\ \xi_i \sim \mathcal{N}(0,1) \text{ and } \operatorname{cov}(\theta,\xi) = \operatorname{cov}(\xi_i,\xi_{j\neq i}) = 0.$$

Second, it gets a public signal

$$z = \theta + \alpha_z^{-1/2}\tilde{\varepsilon},$$

which satisfies

$$\tilde{\varepsilon} \sim \mathcal{N}(0,1) \text{ and } \operatorname{cov}(\tilde{\varepsilon},\theta) = \operatorname{cov}(\tilde{\varepsilon},\xi) = 0.$$

The public signal is taken as given by informed traders in the coordination stage but is endogenously determined in the trading stage of the game (as we describe in the next section).

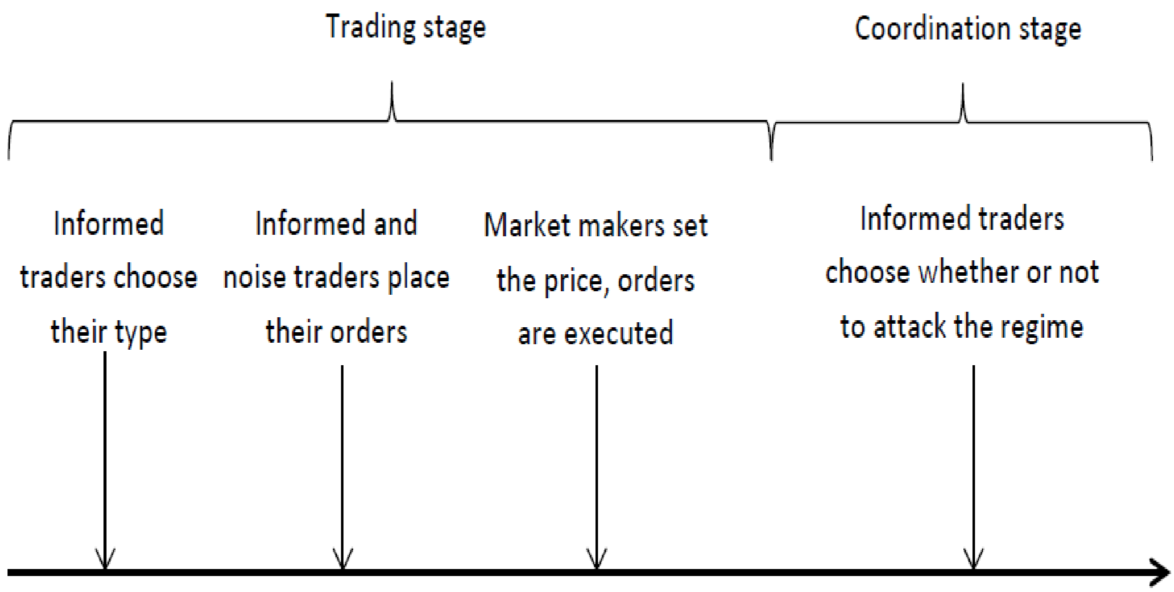


Figure 4.1: Sequence of events.

4.2.1 Coordination stage

In the coordination stage informed trader i chooses action $a_i \in \{0,1\}$, with $a_i = 1$ $(= 0)$ if the trader is attacking (not attacking) the regime.[4] The mass of attacking traders is thus $A = \int_0^1 a_i \mathrm{d}i$, and it is assumed that the regime collapses whenever $A > \theta$. Trader i's payoff at that stage is $U(a_i, A, \theta) = a_i (\mathbf{1}_{A>\theta} - c)$, where $c \in (0,1)$ is the cost of attacking the regime. Hence, the payoff for a trader who successfully (unsuccessfully) attacks the regime is $1 - c > 0$ $(-c < 0)$, while one who does not attack earns 0 for sure. In equilibrium A only depends on the aggregates (θ, z), i.e., $A = A(\theta, z)$. Trader i's policy function is $a(x_i, z) = \arg\max_{a \in \{0,1\}} \mathbb{E}[U(a, A(\theta,z), \theta) | x_i, z]$, with $A(\theta, z) = \int_{\mathbb{R}} a(x_i, z) f(x_i | \theta) \mathrm{d}x_i$, where $f(x|\theta)$ is the density of $x|\theta$ $(\sim \mathcal{N}(\theta, \alpha_x^{-1}))$.

We can restrict our attention to monotone equilibria, in which informed trader i chooses $a_i = 1$ (i.e., to attack) if and only if $x_i < x^*(z)$ (i.e., the trader is sufficiently

[4]This section parallels Angeletos and Werning (2006), except for the fact that we consider a nondiffuse prior, as is required for the asset demands of market-order traders to be well defined. For the sake of comparability we keep the same notations as theirs whenever this is possible.

pessimistic about θ, given (x_i, z)), where $x^*(z)$ is a strategy threshold common to all traders, to be determined as part of the equilibrium.[5] In such equilibria the mass of traders attacking the regime is $A(\theta, z) = \Pr(x_i < x^*(z) | \theta) = \Phi\left(\sqrt{\alpha_x}(x^*(z) - \theta)\right)$, where $\Phi(.)$ is the c.d.f. of the standard normal. The regime is abandoned whenever $A(\theta, z) > \theta$, or equivalently whenever $\theta < \theta^*(z)$, where $\theta^*(z)$ solves

$$\Phi\left(\sqrt{\alpha_x}(x^*(z) - \theta^*(z))\right) = \theta^*(z). \tag{4.1}$$

It directly follows from the properties of $\Phi(.)$ that the latter equation has a unique solution $\theta^*(z) \in (0,1)$ for all $x^*(z) \in \mathbb{R}$, and that $\theta^*(z)$ is continuous and strictly increasing in $x^*(z)$. This has the following interpretation. The threshold $x^*(z)$ summarises traders' *aggressiveness*, in that for any (θ, z) a greater value of $x^*(z)$ increases the attacking mass A. $\theta^*(z)$ represents the regime's *fragility*, in that for any z a greater value of $\theta^*(z)$ widens the range of realisations of θ leading to the regime's collapse. Hence equation (4.1) summarises the way in which a greater level of aggressiveness on the part of traders raises the fragility of the regime.

Since the regime collapses if and only if $\theta \leq \theta^*(z)$, trader i's expected payoff from attacking the regime is $\Pr(\theta \leq \theta^*(z) | x_i, z) - c$. In monotone equilibrium the threshold $x^*(z)$ corresponds to the signal received by the marginal trader (i.e. that indifferent between attacking or not) and hence must satisfy $\Pr(\theta \leq \theta^*(z) | x^*(z), z) = c$. Given the assumed information structure, $\theta | z, x$ is normally distributed with variance $\alpha^{-1} \equiv (\alpha_x + \alpha_z + \alpha_\theta)^{-1}$ and mean $\alpha^{-1}\left(\alpha_x x + \alpha_z z + \alpha_\theta \bar{\theta}\right)$. Hence, indifference of the marginal trader requires:

$$\Phi\left(\sqrt{\alpha_x + \alpha_z + \alpha_\theta}\left(\frac{\alpha_x x^*(z) + \alpha_z z + \alpha_\theta \bar{\theta}}{\alpha_x + \alpha_z + \alpha_\theta} - \theta^*(z)\right)\right) = 1 - c \tag{4.2}$$

The latter equality implicitly defines traders' aggressiveness $x^*(z) \in \mathbb{R}$ as a continuous, strictly increasing function of the regimes' fragility $\theta^*(z) \in (0,1)$ –i.e., a fragile regime makes it safer to bet on its collapse, thereby inducing a rightward shift in $x^*(z)$. Solving

[5] See, e.g., Morris and Shin (2004, Lemma 1).

both (4.1) and (4.2) for $x^*(z)$ and equating the two gives the equation $G(\theta^*) = \Gamma(z)$, where

$$G(\theta^*) \equiv \Phi^{-1}(\theta^*) - \frac{\alpha_z + \alpha_\theta}{\sqrt{\alpha_x}}\theta^*,\ \Gamma(z) = \sqrt{1 + \frac{\alpha_z + \alpha_\theta}{\alpha_x}}\Phi^{-1}(1-c) - \frac{\alpha_\theta}{\sqrt{\alpha_x}}\overline{\theta} - \frac{\alpha_z}{\sqrt{\alpha_x}}z,$$

so we have $\theta^*(z) \in G^{-1}(\Gamma(z))$. When $G : (0,1) \to \mathbb{R}$ is monotonically increasing, it necessarily crosses the $\Gamma(z)$ line exactly once whatever the value of z. When $G(.)$ is non-monotonic there are values of z such that $G(.)$ crosses the $\Gamma(z)$ more than once. It then follows from the minimal value of $\partial G/\partial\theta$ that there exists a unique Bayesian Nash equilibrium for all $z \in \mathbb{R}$ if and only if:

$$\sqrt{2\pi\alpha_x} \geq \alpha_z + \alpha_\theta. \tag{4.3}$$

4.2.2 Trading stage

Let us now turn to the trading stage, which will determine both the distribution (ex ante) and the realisation (ex post) of the public signal z. We assume that informed traders have access to two assets: (i) a riskless bond in perfectly elastic supply and paying out a constant interest rate (with gross value normalised to one); and (ii) a risky asset with trading price p and payoff θ. All informed traders have zero initial wealth (this is without generality), but may freely borrow at the riskless rate to purchase risky assets. It follows that the terminal wealth of informed trader $i \in I$, is given by:

$$w_i = (\theta - p)\,k_i,$$

where k_i is the number of units of risky assets the trader has purchased.

We consider a market microstructure wherein informed traders may place two types of orders. The first is a full *demand schedule*, i.e., a continuum of limit orders allowing full conditionality of the amount of trade on the trading price (as in, e.g., Grossman and Stiglitz, 1976, and much of the subsequent literature on information aggregation in asset markets). The second type of orders that traders may use are *market orders*, which are unconditional on the trading price: they are transmitted to the market maker-making sector before the actual trading price is known, and hence entail some price execution risk (as in, e.g., Vives, 1995, and Medrano, 1996). Aside from informed traders, who use demand schedules or market orders, *noise traders* place a net asset demand for the risky asset of $\varepsilon \sim \mathcal{N}\left(0, \alpha_\varepsilon^{-1}\right)$; this will prevent full revelation of dispersed information through asset prices. All orders are gathered by competitive, risk neutral market-makers. Because they are risk neutral and competitive, all market makers set p to the expected value of θ conditional on the information that they get, which is the total demand for risky assets or *order book* L. In other words we have :

$$p = \mathbb{E}\left(\theta | L\right), \text{ with } L = \int_{i \in I} k_i \mathrm{d}i + \varepsilon,$$

so that the expected profit of a market marker $(\mathbb{E}\left(p - \theta | L\right))$ is zero.

Let us call $M \subset I$ the set of informed traders who submit market orders and $I \backslash M$ the set of informed traders who use full demand schedules. We also define $\nu = \int_{I \backslash M} \mathrm{d}i \in [0,1]$ and $1 - \nu$ as the measures of those two sets sets. All informed traders have CARA preferences, i.e. $V(w_i; \gamma_i) = -\mathrm{e}^{-\gamma_i w_i}$, where γ_i is trader i's risk aversion coefficient. Finally, private signals are assumed to be independent of risk tolerance, i.e.,

$$\forall J \subset I, \int_J (\xi_i / \gamma_i) \, \mathrm{d}i = 0.$$

An *equilibrium* of the trading stage is a pair of investment functions for demand-schedule $(k_{I \backslash M}(x_i, p; \gamma_i))$ and market-order $(k_M(x_i; \gamma_i))$ traders and a price function $p\left(\theta, \varepsilon\right)$ such that:

- $k_{I \backslash M}(.)$ and $k_M(.)$ maximise traders' expected utility:

$$\forall i \in I \backslash M, \ k_{I \backslash M}(x_i, p; \gamma_i) \in \arg\max_{k \in \mathbb{R}} \mathbb{E}[V((\theta - p)k; \gamma_i) | x_i, p], \tag{4.4}$$

$$\forall i \in M, \ k_M(x_i; \gamma_i) \in \arg\max_{k \in \mathbb{R}} \mathbb{E}[V((\theta - p)k; \gamma_i) | x_i]; \tag{4.5}$$

- The market-making sector sets $p = \mathbb{E}[\theta | L(\cdot)]$, where

$$L(p) = \int_{I \setminus M} k_{I\setminus M}(x_i, p; \gamma_i)\mathrm{d}i + \int_M k_M(x_i; \gamma_i)\mathrm{d}i + \varepsilon. \tag{4.6}$$

We then have the following lemma.

Lemma 1. The trading stage has a unique linear Bayesian equilibrium, which is characterised by:

- the investment functions

$$k_{I\setminus M}(x_i, p; \gamma_i) = \frac{\alpha_x}{\gamma_i}(x_i - p) \quad \text{and} \quad k_M(x_i; \gamma_i) = \frac{\beta}{\gamma_i}(x_i - \overline{\theta}), \tag{4.7}$$

  with

$$\beta = \frac{1}{\alpha_x^{-1} + \alpha_\theta^{-1} - \left(\alpha_\theta + B^2 \alpha_\varepsilon\right)^{-1}}$$

- the price function

$$p(\theta, \varepsilon) = (1 - \lambda B)\,\overline{\theta} + \lambda B\left(\theta + B^{-1}\varepsilon\right), \text{ with } \lambda = \frac{B\alpha_\varepsilon}{B^2\alpha_\varepsilon + \alpha_\theta}. \tag{4.8}$$

In (4.7) and (4.8), the constant $B > 0$ is the unique (real) solution to the following equation:

$$B = \alpha_x \frac{\nu}{\gamma_{I\setminus M}} + \frac{1-\nu}{\gamma_M}\left(\frac{1}{\alpha_x} + \frac{1}{\alpha_\theta} - \frac{1}{\alpha_\theta + \alpha_\varepsilon B^2}\right)^{-1}, \tag{4.9}$$

where $\gamma_{I\setminus M}^{-1}$ and γ_M^{-1} are the average risk tolerance coefficients of demand-schedule and market-order traders:

$$\gamma_{I\setminus M}^{-1} = \frac{1}{\nu}\int_{I\setminus M}\gamma_i^{-1}\mathrm{d}i,\ \gamma_M^{-1} = \frac{1}{1-\nu}\int_M \gamma_i^{-1}\mathrm{d}i.$$

Equation (4.8) implies that observing p is equivalent to observing $\theta + B^{-1}\varepsilon$. Thus, the endogenous public signal z about θ, defined in the preceeding section, is $z = \theta + B^{-1}\varepsilon$ (i.e., $\tilde{\varepsilon} = B^{-1}\varepsilon$) and it has precision $\alpha_z = B^2\alpha_\varepsilon$. We then infer from (4.3) that equilibrium uniqueness in the coordination stage requires

$$\sqrt{2\pi\alpha_x} \geq B^2\alpha_\varepsilon + \alpha_\theta. \tag{4.10}$$

Note that when $\alpha_\theta \to 0$ (i.e., the prior is diffuse), $\nu = 1$ and $\gamma_i = \gamma\ \forall i \in [0,1]$ (i.e., all informed traders share the same preferences and place demand schedules), then equation (4.9) gives $B = \gamma^{-1}\alpha_x$, so that $p = \theta + \gamma\sigma_x^2\varepsilon$. Condition (4.10) then becomes $\sqrt{2\pi\alpha_x} \geq \gamma^{-2}\alpha_x^2\alpha_\varepsilon$, which is identical to that in Angeletos and Werning (2006).

4.3 Equilibrium uniqueness versus multiplicity

4.3.1 Markets with a single type

We first consider the case where all informed traders place market orders in the trading stage (as in Vives, 1995) and that where they all place full demand schedules. We then have the following proposition.

Proposition 1. If all informed traders place market orders in the trading stage, then the outcome of the coordination stage is unique as $\alpha_x \to +\infty$. If all informed traders place demand schedules, then there are multiple equilibrium outcomes in the coordination stage as $\alpha_x \to +\infty$.

Proposition 1 implies that when the market microstructure of the trading stage is such that traders place market orders and market makers set the price, then one recovers the original property in Morris and Shin (1998), according to which the outcome of the coordination stage is unique as the noise in the private signal vanishes. In contrast, in a pure demand-schedule market one recovers the basic result in Angeletos and Werning (2006), in which a Walrasian auctioneer (rather than a market-making sector) sets the price. The intuition for this difference is as follows. In a pure *demand-schedule* market ($\nu = 1$), informed traders are able to condition their trades on the trading price, so the only source of risk they face concerns the true value of the fundamental. As the precision of the private signals increases, traders collectively trade more aggressively against any discrepancy between the observed price p and the fundamental θ. Formally, from Lemma 1 the total asset demand by informed traders in a pure demand-schedule market is given by:

$$\int_{I\setminus M} \frac{\alpha_x}{\gamma_i}(\theta + \alpha_x^{-1/2}\xi_i - p)\mathrm{d}i = \alpha_x \left(\int_{I\setminus M} \gamma_i^{-1}\mathrm{d}i\right)(\theta - p) = \frac{\alpha_x}{\gamma_{I\setminus M}}(\theta - p),$$

which implies that $B = \gamma_{I\setminus M}^{-1}\alpha_x \to +\infty$, and thus $p \to \theta$, as $\alpha_x \to +\infty$. In the limit p becomes perfectly informative of θ (i.e. $\alpha_z \to +\infty$); this eventually causes every traders to choose a_i based exclusively on p (rather than x_i) in the second stage and thereby facilitates coordination on a self-fulfilling outcome. In contrast, in a pure *market-order* market ($\nu = 0$) informed traders do *not* condition their trades on p and hence face a residual payoff risk even as the x_is get more and more informative of θ. This payoff risk leads market-order traders to trade less aggressively on the basis of their private signal, which limits the amount of information that is aggregated into the price. Formally, from Lemma 1 again the total asset demand by informed traders in a pure market-order market is:

$$\int_{M} \frac{\beta}{\gamma_i}(\theta + \alpha_x^{-1/2}\xi_i - \bar{\theta})\mathrm{d}i = \beta \left(\int_{M} \gamma_i^{-1}\mathrm{d}i\right)(\theta - \bar{\theta}) = \frac{\beta}{\gamma_M}(\theta - \bar{\theta}),$$

In the limit as $\alpha_x \to +\infty$ we have $\alpha_z = B^2\alpha_\varepsilon < +\infty$, i.e. the precision of the public signal is bounded above. In this case *private* signals ultimately determine actions in the second stage of the game, which hinders coordination on a self-fulfilling outcome.

4.3.2 Market with both types

We now consider the case where both M and $I\setminus M$ have positive measure. For expositional clarity we assume here that traders share the same preferences, i.e., $\gamma_i = \gamma > 0$ $\forall i \in [0,1]$, but the result can straightforwardly be extended to the case of heterogenous γs. We first note that for all $\nu \in [0,1]$ it is necessarily the case that $0 < B \leq \alpha_x/\gamma$, with $B = \alpha_x/\gamma$ when $\nu = 1$ and $B < \alpha_x/\gamma$ when $\nu < 1$.[6] Moreover, the uniqueness condition (4.10) implies that, for $(\alpha_x, \alpha_\varepsilon, \alpha_\theta)$ given, the uniqueness region expands as B falls. Thus, if for a given set of parameters we are in the uniqueness region when $\nu = 1$ (i.e., the pure demand-schedule case), then we are also in the uniqueness region when $\nu < 1$ (and both types coexist). Total differencing (4.9) and using the fact that $0 < B \leq \alpha_x/\gamma$, we find that, for any $(\alpha_x, \alpha_\varepsilon, \alpha_\theta)$ given and for all $\nu \in [0,1]$ we have

$$\frac{\partial B}{\partial \nu} = \frac{\alpha_x - \beta}{\gamma + 2(1-\nu)B\alpha_\varepsilon\left(\alpha_\theta + B^2\alpha_\varepsilon\right)^{-2}\beta^2} > 0.$$

In short, the greater the fraction of market-order traders, the larger the uniqueness region. Again, this is because market order traders face price risk and hence trade less aggressively on their private information than demand-schedule traders do. This reduces the amount of private information that is aggregated into p, thereby reducing its weight in traders' assessment of θ and impeding traders' coordination.[7]

The role of ν in affecting the multiplicity region is illustrated in Figure 2. From the analysis in Section 2.2 we know that $\alpha_z = B^2\alpha_\varepsilon$. Total differencing equation (4.9), we find that for $\partial B/\partial \alpha_x > 0$, implying that a greater precision of the private signal tends to raise α_z. The dotted and dashed lines shows the monotone response of $\sigma_p \equiv \alpha_z^{-1/2}$ (i.e., the noise in the public price signal) to changes in $\sigma_x = \alpha_x^{-1/2}$ (i.e., the noise in the price signal) for different values of ν. The bold line represents the multiplicity versus uniqueness boundary (4.3), i.e. the $\sqrt{2\pi\sigma_x^{-2}} = \sigma_z^{-2} + \alpha_\theta$ line. A smaller value of ν is associated with a smaller uniqueness region as $\sigma_x \to 0$.

[6] Since $\alpha_\theta + B^2\alpha_\varepsilon \geq \alpha_\theta$, we have $\alpha_x^{-1} + \alpha_\theta^{-1} - \left(\alpha_\theta + B^2\alpha_\varepsilon\right)^{-1} \geq \sigma_x^2$ and hence $B \leq \alpha_x/\gamma$.

[7] Note the total effect of ν on B aggregates two effects. First, as ν increases, traders on average trade more aggressively and hence prices become more informative. Second, the aggressiveness of demand-schedule traders tends to increase the price risk faced by market-order traders, thereby pushing them to trade *less* aggressively on their private information as ν increases. The direct effect always dominate, implying that $\partial B/\partial \nu > 0$.

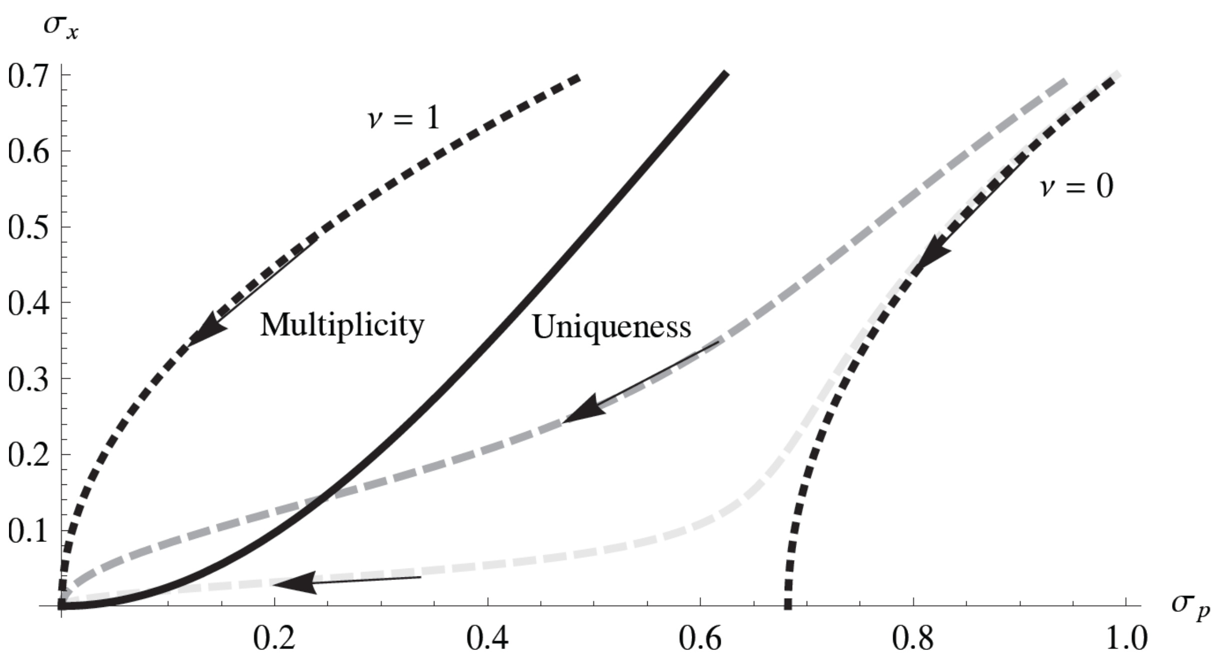


Figure 4.2: Multiplicity and uniqueness regions under exogenous order types. Note: $\sigma_p \equiv \alpha_z^{-1/2}$ and $\sigma_x \equiv \alpha_x^{-1/2}$ denote the noise in the public and private signals, respectively. The bold line is the uniqueness frontier, while the dotted lines shows how σ_p depends on σ_x for different values of ν.

4.4 Endogenous sorting

The analysis above shows that the presence of market-order traders tends to reduce the indeterminacy region by limiting the impact of the price signal on ex post beliefs about the fundamental. We now analyse the equilibrium when traders sort themselves into demand-schedule and market-order traders, so that the two sets are endogenous. What determines the choice of order type by a particular informed trader? The key tradeoff a trader faces is as follows. On the one hand, placing a demand schedule insulates the expected net payoff of a trader from price risk (since effective trades are conditional on the price). On the other hand, it is more costly than a market order, as it requires to place a large (in fact, infinite) number of limit orders in order to generate a complete conditionality of the quantity traded on the execution price. Following Vives (2008), we normalise the cost of a market order to zero and set that of a full demand schedule to $c > 0$. We work out the

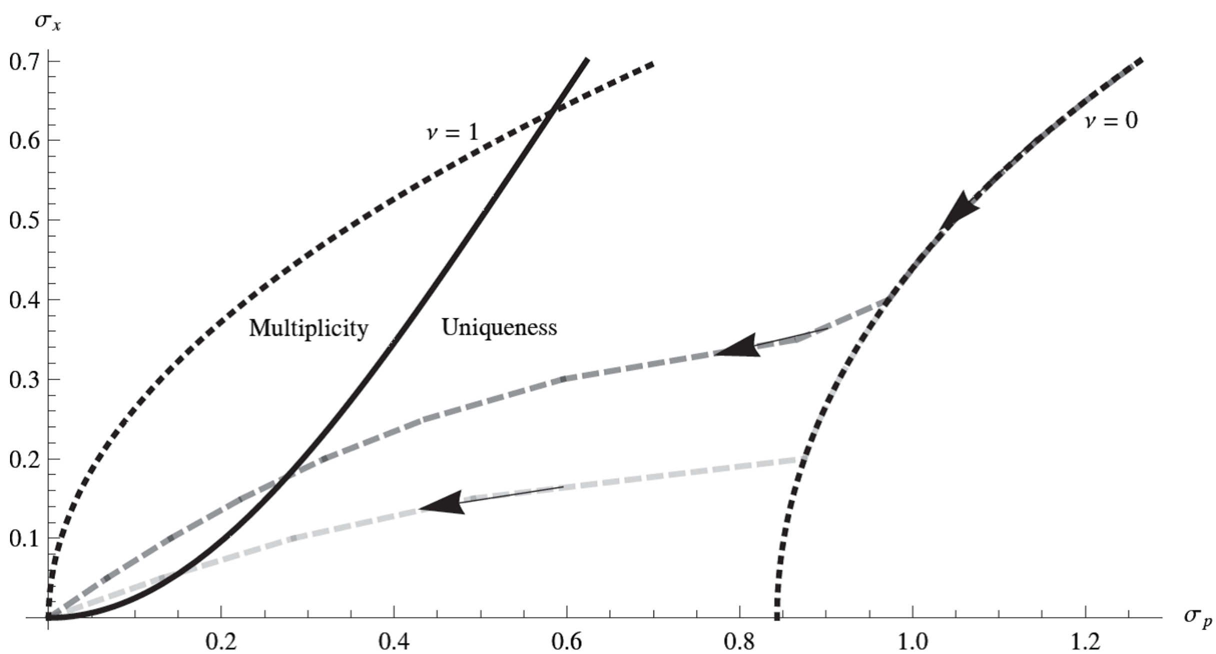


Figure 4.3: Multiplicity and uniqueness regions under endogenous order types. Note: $\sigma_p \equiv \alpha_z^{-1/2}$ and $\sigma_x \equiv \alpha_x^{-1/2}$ denote the noise in the public and private signals, respectively. The bold line is the uniqueness frontier, while the dotted lines shows how σ_p depends on σ_x for different values of c (the relative cost of demand schedules), taking into acount the endogenous adjustment of ν.

solution to this problem under the maintained assumption that the choice of order type must be made before the traders observe their private signal and place their orders–see Figure 1 again.[8]

We rank informed traders in nondecreasing order of risk aversion, define the nondecreasing function $\gamma : [0,1] \to \mathbb{R}_+$, and further assume that $\gamma(.)$ is an increasing homeomorphism and that $\gamma(0) > 0$. We solve for traders' choice of order backwards. First, we compute the expected utility of a trader of each type conditional on its information set (i.e. (x_i, p) $\forall i \in I \backslash M$, and x_i $\forall i \in M$). Second, we compute the unconditional ex ante utility of each type; and third, we compare the two ex ante utilities for a given risk aversion coefficient.

[8]This follows Medrano (1996) and Brown and Zhang (1997). If it were not the case, traders could potentially be willing to adjust their trades (both in terms of order type and amount of trades) depending on the observed shares of demand-schedule and market-order traders, and this would make the signal extraction problem intractable.

We know from the CARA-Normal framework that the value function associated with the information set G_i is:

$$W\left(G_i;\gamma_i\right) \equiv \max_k \mathbb{E}[V(w_i-\kappa c)|G_i;\gamma_i] = -\exp\left[-\frac{\mathbb{E}[\theta-p|G_i]^2}{2\mathbb{V}[\theta-p|G_i]}+\kappa c\gamma_i\right],$$

where $\kappa = 1$ if $G_i = (x_i, p)$ (i.e., the trader places a full demand schedule) or $\kappa = 0$ if $G_i = x_i$ (i.e., the trader places a market order). Using the conditional distributions of θ and $\theta - p$ for demand-schedule and market-order traders (see equations (A1)–(A2) in the Appendix A for details), we find the corresponding value functions to be:

$$W_{I\backslash M}\left(x_i,p;\gamma_i\right) = -\exp\left[-\frac{C}{2}(x_i-p)^2\right],\ C \equiv \frac{\alpha_x^2}{\alpha_x+\alpha_\theta+B^2\alpha_\varepsilon}, \tag{4.11}$$

$$W_M\left(x_i;\gamma_i\right) = -\exp\left[-\frac{D}{2}(x_i-\bar{\theta})^2\right],\ D \equiv \beta^2\left(\frac{(1-\lambda B)^2}{\alpha_x+\alpha_\theta}+\frac{\lambda^2}{\alpha_\varepsilon}\right), \tag{4.12}$$

where β and B are defined in Lemma 1. Let $f(x)$ denote the ex ante (i.e., unconditional) density of the signal x. From the distributions of θ and ξ we have $x \sim \mathcal{N}\left(\theta, \alpha_\theta^{-1}+\alpha_x^{-1}\right)$. Hence, using (4.12) and rearranging the ex ante utility from being a market-order trader is found to be

$$\mathbb{E}\left[W_M\left(x_i;\gamma_i\right)\right] = \int_{\mathbb{R}} W_M\left(x_i;\gamma_i\right) f\left(x_i\right) \mathrm{d}x_i = -\sqrt{\frac{\frac{\alpha_\theta^2}{\alpha_\theta+\alpha_x}+B^2\alpha_\varepsilon}{\alpha_\theta+B^2\alpha_\varepsilon}}. \tag{4.13}$$

The ex ante utility of demand-schedule traders is computed in a similar way, except that we must first condition their information set (x_i, p) on x_i before computing the unconditional expectation of $W_{I\backslash M}\left(x_i,p;\gamma_i\right)$.[9] Applying the law of iterated expectations

[9] Here the intermediate step is the computation of $\mathbb{E}\left[W_L\left(x_i,p;\gamma_i\right)|x_i\right]$. Using the price function (4.8) and the fact that $\theta|x_i \sim \mathcal{N}\left(\frac{\alpha_x x_i+\alpha_\theta\bar{\theta}}{\alpha_x+\alpha_\theta}, \frac{1}{\alpha_x+\alpha_\theta}\right)$ we find that

$$W_L\left(x_i,p;\gamma_i\right)|x_i \sim \mathcal{N}\left(\frac{\alpha_x[\alpha_x(1-\lambda B)+\alpha_\theta](x_i-\bar{\theta})}{\sqrt{2(\alpha_x+\alpha_\theta+B^2\alpha_\varepsilon)}(\alpha_x+\alpha_\theta)}, \left(\frac{\alpha_x\sqrt{(\lambda B)^2(\alpha_x+\alpha_\theta)^{-1}+\lambda^2\alpha_\varepsilon^{-1}}}{\sqrt{2(\alpha_x+\alpha_\theta+B^2\alpha_\varepsilon)}}\right)\right)^2.$$

and rearranging we get:

$$\begin{aligned}\mathbb{E}\left[\mathbb{E}\left[W_{I\backslash M}\left(x_i;\gamma_i\right)\middle|x_i\right]\right] &= \int_{\mathbb{R}} \mathbb{E}\left[W_{I\backslash M}\left(x_i;\gamma_i\right)\middle|x_i\right] f\left(x_i\right) \mathrm{d}x_i \\ &= -\mathrm{e}^{c\gamma(i)}\sqrt{\frac{\alpha_\theta + B^2\alpha_\varepsilon}{\alpha_\theta + B^2\alpha_\varepsilon + \alpha_x}} \qquad (4.14)\end{aligned}$$

Trader i chooses chooses to place a full demand schedule if and only if $\mathbb{E}\left[\mathbb{E}\left[W_{I\backslash M}\left(x_i;\gamma_i\right)\middle|x_i\right]\right] \geq \mathbb{E}\left[W_M\left(x_i;\gamma_i\right)\right]$, i.e., if and only if

$$\gamma\left(i\right) \leq \bar{\gamma} = \frac{1}{c}\ln\left(\frac{\sqrt{\left(\frac{\alpha_\theta^2}{\alpha_\theta+\alpha_x} + B^2\alpha_\varepsilon\right)\left(\alpha_\theta + B^2\alpha_\varepsilon + \alpha_x\right)}}{\alpha_\theta + B^2\alpha_\varepsilon}\right), \qquad (4.15)$$

where, from Lemma 1,

$$B = \alpha_x \int_0^{\gamma^{-1}(\bar{\gamma})} \gamma\left(i\right)^{-1} \mathrm{d}i + \left(\frac{1}{\alpha_x} + \frac{1}{\alpha_\theta} - \frac{1}{\alpha_\theta + \alpha_\varepsilon B^2}\right)^{-1} \int_{\gamma^{-1}(\bar{\gamma})}^1 \gamma\left(i\right)^{-1} \mathrm{d}i, \qquad (4.16)$$

with $\gamma^{-1}\left(\bar{\gamma}\right) = 0$ if $\bar{\gamma} < \gamma\left(0\right)$ and $\gamma^{-1}\left(\bar{\gamma}\right) = 1$ if $\bar{\gamma} > \gamma\left(1\right)$. For a given triplet $\left(\alpha_x, \alpha_\theta, \alpha_\varepsilon\right) \in \mathbb{R}_+^3$ given, the properties of the $\gamma\left(.\right)$ function imply that the solution $\left(\bar{\gamma}, B\right)$ to (4.15)–(4.16), if it exists, can be of three types. More specifically, it is either such that $\bar{\gamma} \in \left[\gamma\left(0\right), \gamma\left(1\right)\right]$, in which case the solution is interior (i.e., both M and $I\backslash M$ are nonempty); or $\bar{\gamma} < \gamma\left(0\right)$, so that the solution is corner and all traders placing market orders (i.e., $\left(M, I\backslash M\right) = \left(I, \varnothing\right)$); or $\bar{\gamma} > \gamma\left(1\right)$ and all traders place full demand schedules (i.e., $\left(M, I\backslash M\right) = \left(\varnothing, I\right)$). The intuition for this sorting of informed traders according to their degree of risk aversion is that a greater risk aversion lowers trading aggressiveness, and hence the expected benefit from expanding the information set from x_i to $\left(x_i, p\right)$.[10]

As before we are interested in the outcome of the coordination stage of the game as α_x becomes large (holding $\left(\alpha_\theta, \alpha_\varepsilon, c\right)$ fixed), especially with regard to the way market-order traders alter the size of the uniqueness region. This is summarised in the proposition 2 below.

[10]See Medrano (1996) and Vives (2008) for further discussion.

Proposition 2. For any $(\alpha_\theta, \alpha_\varepsilon, c) \in \mathbb{R}^3_+$, and as $\alpha_x \to +\infty$, (i) both M and $I\backslash M$ have strictly positive measure (i.e., the equilibrium is interior); (ii) $\bar{\gamma} \to \gamma(0)$ (i.e., market-order traders eventually overwhelm the market); (iii) $\alpha_z \underset{\alpha_x \to \infty}{\sim} (\mathrm{e}^{2\gamma(0)c} - 1)^{-1}\alpha_x$, so that α_z goes to infinity as the same rate α_x (while it does at the same rate as α_x^2 when $I\backslash M$ has exogenous, positive measure).

Proposition 2 emphasises several key properties of the equilibrium when α_x is large. Note that the property that the equilibrium is interior as $\alpha_x \to +\infty$ (point (ii)) is valid for any value of the cost c; in contrast, when α_x is small one can easily construct examples of corner solutions with a pure market-order (demand-schedule) market when c is sufficiently high (low). Points (ii) and (iii) are closely related. As discussed in Section 3, market-order traders tend to slow down information aggregation. It is precisely because they crowd out demand-schedule traders as $\alpha_x \to +\infty$ (point (ii)) that the precision of the endogenous public signal grows at the same rate as α_x, instead of α_x^2 when market shares are exogenous (point (iii)). To see how this may expand the uniqueness region, note that under exogenous shares from (4.9) we have $\alpha_z \underset{\alpha_x \to \infty}{\sim} (\nu/\gamma_{I\backslash M})\alpha_x^2$. Hence for $\nu > 0$ and α_x large enough, since $\left[(e^{2\gamma(0)c} - 1)^{-1}\alpha_x\right]/(\nu/\gamma_{I\backslash M})\alpha_x^2 \underset{\alpha_x \to \infty}{\to} 0$ it is necessarily the case that the precision of the price signal is greater under endogenous orders than under exogenous orders. Hence, whenever the uniqueness condition (4.3) is satisfied under exogenous shares, it is also so under endogenous shares, but the converse is not true. Figure 3 illustrates the relationship between α_z and α_x when α_x is large (i.e., $\sigma_x = \alpha_x^{-1/2}$ is small) and the shares of market-order and demand-schedules traders are endogenous.

Finally, note from (4.14) that heterogeneity in the cost c is formally equivalent to heterogeneity in risk aversion. To encompass both cases, rank traders in nondecreasing orders of $c(i)\gamma(i)$, assume that the function $g(i) = c(i)\gamma(i)$ is continuous, strictly increasing, that its reciprocal is continuous, and that $0 < g(0) < g(1) < +\infty$, and solve for the marginal trader exactly in the same way as in the case where $c(i) = c\ \forall i \in I$.

4.5 Concluding remarks

In this paper, we have analysed a two-stage global game wherein a market-based asset price determined at the trading stage of the game provides an endogenous public signal affecting traders' decisions in the coordination stage of the game. As we have shown, in this context the multiplicity region can be small even when private information is very precise –and especially so when traders optimise over their type of order (in addition to their amount of trade). The reason for this is that the presence of market-order traders limits information aggregation and hence the precision of the endogenous public signal that may serve as a coordination device when deciding whether or not to attack the regime. In this sense, a lower degree of informational efficiency (in the trading stage) may ultimately be stabilising (in the coordination stage). While this conclusion was derived under a specific barrier to full informational efficiency –market-order traders' willingness to avoid price risk–, we conjecture that it would also apply in a variety of contexts where information aggregation is impeded. This shall namely include markets with low liquidity, like FX markets. In this case, it might be the case that currencies traded on less liquid assets are safer bets for speculation, as they shield investors against self-fulfilling risks. Discussion about the implications of these findings for FX speculation is left for future research.

4.6 Appendix

A. Proof of Lemma 1

We restrict our attention to equilibrium price functions $p(\theta,\varepsilon)$ that are linear in (θ,ε), which implies that p is normally distributed. A trader i with risk aversion coefficient γ_i and information set G_i has a demand for assets $k_i(G_i) = \gamma_i^{-1}\mathbb{E}[\theta - p|G_i]/\mathbb{V}[\theta - p|G_i]$. We may thus write the demands by limit- and market-order traders as follows:

$$\forall i \in I\backslash M,\ k^i_{I\backslash M}(x_i,p) = \gamma_i^{-1} f_{I\backslash M}(x_i,p), \text{ with } f_{I\backslash M}(x_i,p) = \frac{\mathbb{E}[\theta|x_i,p] - p}{\mathbb{V}[\theta|x_i,p]},$$
$$\forall i \in M,\ k^i_M(x_i) = \gamma_i^{-1} f_M(x_i), \text{ with } f_M(x_i) = \frac{\mathbb{E}[\theta - p|x_i]}{\mathbb{V}[\theta - p|x_i]},$$

i.e., within each group asset demands are identical up to a risk tolerance correction γ_i^{-1}. Now conjecture that $f_{I\backslash M}(.)$ and $f_M(.)$ have the form $f_{I\backslash M}(x_i,p) = a(x_i - \overline{\theta}) + \zeta(p)$ and $f_M(x_i) = c(x_i - \overline{\theta})$, where a and b are normalised trading intensities (for a trader with $\gamma_i = 1$) and $\zeta(.)$ is linear. Using the convention that the average signal equals θ a.s., and recalling that γ_i is independent from ξ_i, the limit order book is given by

$$L(p) = \int_{I\backslash M} k^i_{I\backslash M}(x_i,p)\mathrm{d}i + \int_M k^i_M(x_i)\mathrm{d}i + \varepsilon = B\left[\theta + B^{-1}\varepsilon\right] - B\overline{\theta} + \zeta(p)\int_{I\backslash M}\gamma_i^{-1}\mathrm{d}i,$$

where $B = a\nu/\gamma_{I\backslash M} + c(1-\nu)/\gamma_M$. The market making sector observes $L(.)$, a linear function of p, and sets $p = \mathbb{E}[\theta|L(.)] = \mathbb{E}[\theta|z]$, where $z = \theta + B^{-1}\varepsilon$ is the public signal. From standard normal theory we infer that p is indeed linear, normal and given by equation (4.8).

We now need to identify a and c. From the joint distribution of (p, x_i, θ) we get:

$$\forall i \in I\backslash M, \begin{cases} \mathbb{E}[\theta|p,x_i] = \frac{B^2\alpha_\varepsilon z + \alpha_\theta\overline{\theta} + \alpha_x x_i}{B^2\alpha_\varepsilon + \alpha_\theta + \alpha_x} = \frac{(B^2\alpha_\varepsilon + \alpha_\theta)p + \alpha_x x_i}{B^2\alpha_\varepsilon + \alpha_\theta + \alpha_x}, \\ \mathbb{V}[\theta|p,x_i] = (B^2\alpha_\varepsilon + \alpha_\theta + \alpha_x)^{-1}. \end{cases} \tag{A1}$$

$$\forall i \in M, \begin{cases} \mathbb{E}[\theta - p|x_i] = \frac{(1-\lambda B)\alpha_x}{\alpha_x + \alpha_\theta}(x_i - \overline{\theta}), \\ \mathbb{V}[\theta - p|x_i] = (1-\lambda B)^2\ \mathbb{V}[\theta|x_i] + \frac{\lambda^2}{\alpha_\varepsilon} = \frac{(1-\lambda B)^2}{\alpha_x + \alpha_\theta} + \frac{\lambda^2}{\alpha_\varepsilon}. \end{cases} \tag{A2}$$

Hence, we obtain

$$k^i_{I\backslash M}(x_i, p; \gamma_i) = \frac{\mathbb{E}[\theta|p, x_i] - p}{\gamma_i \mathbb{V}[\theta|p, x_i]} = \frac{\alpha_x}{\gamma_i}(x_i - p), \;\; k^i_M(x_i; \gamma_i) = \frac{\mathbb{E}[\theta - p|x_i]}{\gamma_i \mathbb{V}[\theta - p|x_i]} = \frac{\beta}{\gamma_i}\left(x_i - \overline{\theta}\right),$$

where $\beta = (\alpha_x^{-1} + \alpha_\theta^{-1} - \left(\alpha_\theta + B^2\alpha_\varepsilon\right)^{-1})^{-1}$. In the special case where $\gamma_i = \gamma \;\forall i \in [0,1]$, we have $k^i_{I\backslash M}(x_i, p) = \gamma^{-1}\alpha_x(x_i - p)$, $k_M(x_i) = \gamma^{-1}\beta(x_i - \overline{\theta})$ and $p = (1 - \lambda B)\overline{\theta} + \lambda B z$, where B solves $B = \nu\gamma^{-1}\alpha_x + (1-\nu)\gamma^{-1}\beta$ and $\lambda = \frac{B\alpha_\varepsilon}{B^2\alpha_\varepsilon + \alpha_\theta}$.

Let us now turn to the parameter B. To establish that B is unique, positive and finite, define the function $f : \mathbb{R} \to \mathbb{R}$ as follows:

$$f : B \to B - \frac{\nu\alpha_x}{\gamma_{I\backslash M}} - \frac{1-\nu}{\gamma_M\left(\alpha_x^{-1} + \alpha_\theta^{-1} - (\alpha_\theta + B^2\alpha_\varepsilon)^{-1}\right)},$$

so that a root of $f(B)$ is a solution to (4.9). f is continuous and strictly increasing over $[0, +\infty)$ and such that $f(0) = -\alpha_x\left(\frac{\nu}{\gamma_{I\backslash M}} + \frac{(1-\nu)}{\gamma_M}\right) < 0$ and $\lim_{B \to +\infty} f(B) = +\infty$. Hence f is a bijection that admits a unique root $B_0 > 0$ over $[0, +\infty)$. Moreover, as $B \to \alpha_x^{-1} + \alpha_\theta^{-1} - \left(\alpha_\theta + B^2\alpha_\varepsilon\right)^{-1} > 0$ on $\mathbb{R}$, $f(.)$ is strictly negative on $\mathbb{R}_-$. Hence B_0 is the unique root of f in $\mathbb{R}$. In the numerical implementation of the model we use the exact solution for B, which is found using Cardano's method and gives:

$$B = \sqrt[3]{\frac{1}{2}(-(2\frac{a_2^3}{27} - \frac{a_1 a_2}{3} + a_0) + \sqrt{\frac{4a_1^3 + 4a_0 a_2^3 - (a_1 a_2)^2}{27} - \frac{2}{3}a_0 a_1 a_2 + a_0^2})}$$

$$+ \sqrt[3]{\frac{1}{2}(-(2\frac{a_2^3}{27} - \frac{a_1 a_2}{3} + a_0) - \sqrt{\frac{4a_1^3 + 4a_0 a_2^3 - (a_1 a_2)^2}{27} - \frac{2}{3}a_0 a_1 a_2 + a_0^2})} - \frac{a_2}{3},$$

where

$$a_0 = -\frac{\alpha_\theta}{(\alpha_x^{-1} + \alpha_\theta^{-1})\alpha_\varepsilon}\left(\frac{\nu}{\gamma_{I\backslash M}} + \frac{(1-\nu)}{\gamma_M}\right), \;\; a_1 = \frac{\alpha_x^{-1}\alpha_\theta}{(\alpha_x^{-1} + \alpha_\theta^{-1})\alpha_\varepsilon},$$

$$\text{and} \;\; a_2 = -\frac{\left[\frac{\nu}{\gamma_{I\backslash M}} + \frac{(1-\nu)}{\gamma_M} + \frac{\nu}{\gamma_{I\backslash M}}\alpha_x\alpha_\theta^{-1}\right]}{(\alpha_x^{-1} + \alpha_\theta^{-1})}.$$

B. Proof of Proposition 1

We know from Lemma 1 that $B \in \mathbb{R}_+^*$ uniquely solves (4.9). When $\nu = 0$, B solves $1/(\gamma_M x) = \alpha_x^{-1} + \alpha_\theta^{-1} - \left(\alpha_\theta + \alpha_\varepsilon x^2\right)^{-1}$. In this case $\lim_{\alpha_x \to +\infty} B$ is finite, hence the uniqueness condition (4.10) necessarily holds as $\alpha_x \to +\infty$. In contrast, when $\nu = 1$ we have $\lim_{\alpha_x \to +\infty} \frac{B^2}{\sqrt{\alpha_x}} = +\infty$, hence the uniqueness condition (4.10) is necessarily violated as $\alpha_x \to +\infty$.

C. Proof of Proposition 2

(i) We show that $k \equiv \gamma^{-1}(\bar{\gamma}) \in]0;1[$ for α_x sufficiently high, and that k is unique. Let us first define the function

$$\tilde{f} : \alpha_x, k \to \mathrm{e}^{2\gamma(k)c} - \left(1 - \frac{\alpha_\theta}{\alpha_x + \alpha_\theta} \frac{\alpha_x}{B(k,\alpha_x)^2\alpha_\varepsilon + \alpha_\theta}\right)\left(1 + \frac{\alpha_x}{B(k,\alpha_x)^2\alpha_\varepsilon + \alpha_\theta}\right), \tag{C.1}$$

where $B(k,\alpha_x)$ is the unique solution to

$$B(k,\alpha_x) = \alpha_x \int_0^k \gamma(i)^{-1}\mathrm{d}i + \left(\frac{1}{\alpha_x} + \frac{1}{\alpha_\theta} - \frac{1}{\alpha_\theta + B(k,\alpha_x)^2\alpha_\varepsilon}\right)^{-1} \int_k^1 \gamma(i)^{-1}\mathrm{d}i. \tag{C.2}$$

We have $\tilde{f}(\alpha_x, 1) \underset{\alpha_x \to \infty}{\to} \mathrm{e}^{2\gamma(1)c} - 1 > 0$ while $\tilde{f}(\alpha_x, 0) \underset{\alpha_x \to \infty}{\to} -\infty < 0$. Hence, by the intermediate value theorem there exists $\underline{\alpha} \in \mathbb{R}_+^*$,such that for all $\alpha_x \geq \underline{\alpha}$, $0 \in]\tilde{f}(\alpha_x, 0), \tilde{f}(\alpha_x, 1)[$. By continuity, $\forall \alpha_x \geq \underline{\alpha}$, $\exists k(\alpha_x) \in]0,1[$ such that $\tilde{f}(\alpha_x, k(\alpha_x)) = 0$. In this range of parameter, there exists an interior equilibrium allocation, and the corner solutions are ruled out (otherwise the polar traders would be better off switching positions).

To establish uniqueness, define $\tilde{\alpha} \equiv \alpha_\theta / (\alpha_x + \alpha_\theta)$ and $X \equiv \alpha_x / \left(B(k(\alpha_x), \alpha_x)^2\alpha_\varepsilon + \alpha_\theta\right)$, and rewrite $\tilde{f}(\alpha_x, k(\alpha_x)) = 0$ as $\mathcal{P}(X) = \tilde{\alpha}X^2 - (1-\tilde{\alpha})X + \mathrm{e}^{2\gamma(k(\alpha_x))c} - 1 = 0$. This polynomial has the following two real roots:

$$s^-, s^+ = \frac{1}{2\tilde{\alpha}}\left[(1-\tilde{\alpha}) \mp \sqrt{(1-\tilde{\alpha})^2 - 4\tilde{\alpha}\left(\mathrm{e}^{2\gamma(k(\alpha_x))c} - 1\right)}\right].$$

We prove by contradiction that $X = s^-$ is the only possible root of $\mathcal{P}(X) = 0$ when α_x becomes large enough. Formally,

$$\exists \underline{\alpha}^1 \geq \underline{\alpha}, \forall \alpha_x \in \mathbb{R}_+^*, \forall k \in]0;1[, \quad \left.\begin{array}{c} \alpha_x \geq \underline{\alpha}^1 \\ \tilde{f}(\alpha_x, k) = 0 \end{array}\right\} \Rightarrow \frac{\alpha_x}{B(k,\alpha_x)^2\alpha_\varepsilon + \alpha_\theta} = s^-. \tag{C.3}$$

To see this, suppose that $\forall \underline{\alpha}^1 \geq \underline{\alpha}, \exists \alpha_x \in \mathbb{R}_+^*, \exists k \in]0;1[$ such that

$$\left(\alpha_x \geq \underline{\alpha}^1\right) \wedge \left(\tilde{f}(\alpha_x, k(\alpha_x)) = 0\right) \wedge \left(\frac{\alpha_x}{B(k(\alpha_x),\alpha_x)^2\alpha_\varepsilon + \alpha_\theta} = s^+\right).$$

In particular, for $n \in \mathbb{N}$ large enough (say larger than $n_0 = \lceil \underline{\alpha} \rceil$),

$$\exists \alpha_x, \exists k, (\alpha_x \geq n) \wedge \left(\tilde{f}(\alpha_x, k) = 0\right) \wedge \left(\alpha_x / \left(B(k,\alpha_x)^2\alpha_\varepsilon + \alpha_\theta\right) = s^+\right)$$

For every $n \geq n_0$ we pick an α_x, and an associated $k(\alpha_x)$, satisfying $(\alpha_x \geq n) \wedge \left(\tilde{f}(\alpha_x, k(\alpha_x)) = 0\right) \wedge \left(\alpha_x / \left(B(k(\alpha_x),\alpha_x)^2\alpha_\varepsilon + \alpha_\theta\right)\right)$ and denote it α_n (resp. $k(\alpha_n)$), thereby constructing the series $(\alpha_n)_{n \geq n_0}$ (resp. $(k(\alpha_n))_{n \geq n_0}$). As $\alpha_n \underset{n\to\infty}{\to} \infty$, and since $k(\alpha_n)$ must belong to $[0;1]$ we have

$$\frac{4\tilde{\alpha}}{(1-\tilde{\alpha})^2}\left(e^{2\gamma(k(\alpha_n))c} - 1\right) = \frac{4\alpha_\theta}{\alpha_n + \alpha_\theta}\left(\frac{\alpha_n + \alpha_\theta}{\alpha_n}\right)^2\left(e^{2\gamma(k(\alpha_n))c} - 1\right) \underset{n\to\infty}{\to} 0, \tag{C.4}$$

and hence

$$\frac{X\tilde{\alpha}}{1-\tilde{\alpha}} = \frac{1}{2}\left[1 + \sqrt{1 - \frac{4\tilde{\alpha}}{(1-\tilde{\alpha})^2}\left(e^{2\gamma(k(\alpha_n))c} - 1\right)}\right] \underset{n\to\infty}{\to} 1.$$

This in turn implies that

$$\frac{\alpha_n}{B(k(\alpha_n),\alpha_n)^2\alpha_\varepsilon + \alpha_\theta} \underset{n\to\infty}{\sim} \frac{\alpha_n}{\alpha_\theta},$$

that is, $B(k(\alpha_n),\alpha_n)^2\alpha_\varepsilon \underset{n\to\infty}{\to} 0$. Since for each $n \geq n_0$, $B(k(\alpha_n),\alpha_n)^2 \geq B(0,\alpha_n)^2$ while $\left(B(0,\alpha_n)^2\right)_{n\in\mathbb{N}}$ admits a finite, non-zero limit as $n \to \infty$, we have a contradiction that proves (C.3). To summarise, for every given set of parameters, the function

$$k \to \frac{\alpha_x}{B(k,\alpha_x)^2\alpha_\varepsilon + \alpha_\theta} - \frac{1}{2\alpha}\left[(1-\tilde{\alpha}) - \sqrt{(1-\tilde{\alpha})^2 - 4\alpha\left(\mathrm{e}^{2\gamma(k)c} - 1\right)}\right] \tag{C.4}$$

is strictly decreasing and thus has a unique root. For every $\alpha_x \geq \underline{\alpha}^1$, we also have that $\alpha_x \geq \underline{\alpha}$, hence we know that there exists $k \in]0;1[$ such that $\tilde{f}(\alpha_x, k) = 0$. Moreover, as $\alpha_x \geq \underline{\alpha}^1$, we know that k is the unique root of (C.4). Hence, $k \in]0;1[$ exists and is unique. We will denote it $k(\alpha_x)$.

(ii) From above, $\forall \alpha_x \geq \underline{\alpha}^1$ we have $\alpha_x / \left(B(k(\alpha_x), \alpha_x)^2\alpha_\varepsilon + \alpha_\theta\right) = s^-$. Since $k(\alpha_x) \in]0;1[$ we have

$$\frac{\alpha_x}{B(k(\alpha_x),\alpha_x)^2\alpha_\varepsilon + \alpha_\theta} \geq \frac{(1-\tilde{\alpha}) - \sqrt{(1-\tilde{\alpha})^2 - 4\tilde{\alpha}\left(\mathrm{e}^{2\gamma(0)c} - 1\right)}}{2\tilde{\alpha}},$$

or, rearranging,

$$B(k(\alpha_x),\alpha_x)^2 \leq \frac{2\tilde{\alpha}}{(1-\tilde{\alpha}) - \sqrt{(1-\tilde{\alpha})^2 - 4\tilde{\alpha}\left(\mathrm{e}^{2\gamma(0)c} - 1\right)}}\frac{\alpha_x}{\alpha_\varepsilon}.$$

Moreover, from (4.16) we know that

$$\left(\frac{k(\alpha_x)}{\gamma(1)}\alpha_x\right)^2 \leq \alpha_x^2\left(\int_0^{k(\alpha_x)} \gamma(i)^{-1}\mathrm{d}i\right)^2 \leq B(k(\alpha_x),\alpha_x)^2$$

Hence, $\forall \alpha_x \geq \underline{\alpha}^1$,

$$0 \leq k(\alpha_x) \leq \frac{\gamma(1)}{\alpha_x}\sqrt{\frac{\alpha_x}{\alpha_\varepsilon}\frac{2\tilde{\alpha}}{(1-\tilde{\alpha}) - \sqrt{(1-\tilde{\alpha})^2 - 4\tilde{\alpha}\left(\mathrm{e}^{2\gamma(0)c} - 1\right)}}}$$

We know from (C.4) that $\frac{4\tilde{\alpha}}{(1-\tilde{\alpha})^2}\left(\mathrm{e}^{2\gamma(0)c}-1\right) \underset{\alpha_x\to\infty}{\to} 0$. It follows that

$$\frac{(1-\tilde{\alpha})-\sqrt{(1-\tilde{\alpha})^2-4\tilde{\alpha}\left(\mathrm{e}^{2\gamma(0)c}-1\right)}}{2\tilde{\alpha}} = \frac{(1-\tilde{\alpha})-(1-\tilde{\alpha})\sqrt{1-\frac{4\tilde{\alpha}}{(1-\tilde{\alpha})^2}\left(\mathrm{e}^{2\gamma(0)c}-1\right)}}{2\alpha}$$

$$= \frac{(1-\tilde{\alpha})}{4\tilde{\alpha}}\frac{4\tilde{\alpha}}{(1-\tilde{\alpha})^2}\left(\mathrm{e}^{2\gamma(0)c}-1\right) + \underset{\alpha_x\to\infty}{o}(1) = \frac{\mathrm{e}^{2\gamma(0)c}-1}{1-\tilde{\alpha}} + \underset{\alpha_x\to\infty}{o}(1) \underset{\alpha_x\to\infty}{\to} \mathrm{e}^{2\gamma(0)c}-1,$$

We infer that

$$\frac{\gamma(1)}{\alpha_x}\sqrt{\frac{\alpha_x}{\alpha_\varepsilon}\frac{2\tilde{\alpha}}{(1-\tilde{\alpha})-\sqrt{(1-\tilde{\alpha})^2-4\tilde{\alpha}\left(\mathrm{e}^{2\gamma(0)c}-1\right)}}} \underset{\alpha_x\to\infty}{\sim} \frac{\gamma}{\alpha_x}\sqrt{\frac{\alpha_x/\alpha_\varepsilon}{\mathrm{e}^{2\gamma(0)c}-1}} \underset{\alpha_x\to\infty}{\to} 0,$$

Hence $k(\alpha_x) \underset{\alpha_x\to\infty}{\to} 0$, and (by the continuity of $\gamma(.)$) $\bar{\gamma} = \gamma(k(\alpha_x)) \underset{\alpha_x\to\infty}{\to} \gamma(0)$.

(iii) Using (C.3) above we get

$$\frac{\alpha_x}{B(k(\alpha_x),\alpha_x)^2\alpha_\varepsilon+\alpha_\theta} = s^- = \frac{\mathrm{e}^{2\gamma(k(\alpha_x))c}-1}{1-\tilde{\alpha}} + \underset{\alpha_x\to\infty}{o}\left(\frac{\mathrm{e}^{2\gamma(k(\alpha_x))c}-1}{1-\tilde{\alpha}}\right) \underset{\alpha_x\to\infty}{\to} \mathrm{e}^{2\gamma(0)c}-1.$$

We infer that

$$\frac{B(k(\alpha_x),\alpha_x)^2\alpha_\varepsilon}{\alpha_x} = \frac{B(k(\alpha_x),\alpha_x)^2\alpha_\varepsilon+\alpha_\theta}{\alpha_x} - \frac{\alpha_\theta}{\alpha_x} \underset{\alpha_x\to\infty}{\to} \frac{1}{\mathrm{e}^{2\gamma(0)c}-1}$$

We conclude that $\alpha_z = B(k(\alpha_x),\alpha_x)^2\alpha_\varepsilon \underset{\alpha_x\to\infty}{\sim} (\mathrm{e}^{2\gamma(0)c}-1)^{-1}\alpha_x$.

Bibliography

[1] Angeletos, G.-M., Hellwig, C. and Pavan, A. (2006), Signalling in a global game: Coordination and policy traps, *Journal of Political Economy*, 114, 452–484

[2] Angeletos, G.-M. and Werning, I. (2006), Crises and prices: information aggregation, multiplicity, and volatility, *American Economic Review*, 96(5), 1720–1736.

[3] Atkeson, A. (2001). Discussion of Morris and Shin's "Rethinking multiple equilibria in macroeconomic modelling", in *NBER Macroeconomics Annual 2000*, Cambridge, MA: MIT Press.

[4] Brown, D.P. and Zhang, Z.M. (1997), Market orders and market efficiency, *Journal of Finance*, 52(1), 277-308.

[5] Brunnermeier, M. Nagel, S. and Pedersen, L.H. (2008) Carry Trades and Currency Crashes. Ed. K. Rogoff, M. Woodford, and D. Acemoglu. *NBER Macroeconomics Annual*, 23, pp. 313-347.

[6] Burnside, C. (2008), Carry Trades and Currency Crashes: A Comment. Ed. K. Rogoff, M. Woodford, and D. Acemoglu. *NBER Macroeconomics Annual*, 23, pp. 349-359.

[7] Carlsson, H. and van Damme, E. (1993), Global games and equilibrium selection, *Econometrica*, 61, 989–1018.

[8] Challe, E. and Chrétien, E. (2015), Market composition and price informativeness in a large market with endogenous order types, *Journal of Economic Theory*, 158, 679–696.

[9] Grossman, S.J. and Stliglitz, J. (1976), Information and competitive price system, *American Economic Review*, 66(2), pp. 246-253.

[10] Geromichalos, A. and Jung, K.M. (2016), An Over-The-Counter Approach to the FOREX Market, Mimeo

[11] Grossman, S.J. and Stliglitz, J. (1976), Information and competitive price system, *American Economic Review*, 66(2), pp. 246-253.

[12] Hellwig, C. (2002), Public information, private information, and the multiplicity of equilibria in coordiation games, *Journal of Economic Theory*, 107, 191–222.

[13] Hellwig, C., Mukherji, A. and Tsyvinski, A. (2006), Self-fulfilling currency crises: The role of interest rates, *American Economic Review*, 96(5), 1769–1787.

[14] Karnaukh, N., Ranaldo, A., and Söderlind, P. (2015), Understanding FX Liquidity, *Review of Financial Studies*, 28 (11), pp. 3073-3108.

[15] Kyle, A.S. (1989), Informed speculation under imperfect competition, *Review of Economic Studies*, 56(3), 317–355.

[16] Lyons, R.K. (2001), *The Microstructure Approach to Exchange Rates*, Cambridge, Mass., MIT Press

[17] Mancini, L., Ranaldo, A. and Wrampelmeyer, J. (2013), Liquidity in the Foreign Exchange Market: Measurement, Commonality, and Risk Premiums, *Journal of Finance*, 68(5), 1805-1841.

[18] Medrano, L.A. (1996), Market versus limit orders in an imperfectly competitive security market, UPF Working paper 165, february.

[19] Morris, S. and Shin, H.S. (2004), Coordination risk and the price of debt, *European Economic Review*, 48, 133–153.

[20] Morris, S. and Shin, H.S. (1998), Unique equilibrium in a model of self-fulfilling currency attacks, *American Economic Review*, 88(3), 587–597.

[21] Vives, X. (2008), *Information and Learning in Markets: The impact of market microstructure*, Princeton University Press.

[22] Vives, X. (1995), The speed of information revelation in a financial market mechanism, *Journal of Economic Theory*, 67(1), 178–204.

[23] Wald, J.K. and Horrigan, H.T. (2005), Optimal limit order choice, *Journal of Business*, 78(2), 597–619.

Chapter

5 Traditional and shadow banks during the crisis

We present a model of the interactions between traditional and shadow banks that explains their coexistence. In the 2007 financial crisis, some of shadow banks' assets and liabilities have moved to traditional banks, and assets were sold at fire sale prices. Our model is able to accommodate these stylized facts. The difference between traditional and shadow banks is twofold. First, traditional banks have access to a guarantee fund that enables them to issue claims to households in a crisis. Second, traditional banks have to comply with costly regulation. We show that in a crisis, shadow banks liquidate assets to repay their creditors, while traditional banks purchase these assets at fire-sale prices. This exchange of assets in a crisis generates a complementarity between traditional and shadow banks, where each type of intermediary benefits from the presence of the other. We find two competing effects from a small decrease in traditional banks' support in a crisis, which we dub a substitution effect and an income effect. The latter effect dominates the former, so that lower anticipated support to traditional banks in a crisis increases entry in the traditional banking sector ex-ante. We also propose a normative approach and show that, when both traditional and shadow banks coexist, there is in general too many shadow banks. Bankers, while allocating themselves towards the shadow banking system, fail to internalize the fact that this choice reduces the support to all shadow banks in terms of crises, hence limiting the investment made by all shadow banks in normal times.

JEL-code : *E32, E44, E61, G01, G21, G23, G38.*

Keywords*: Traditional banking, Shadow banking, Safe money-like claims, Financial crisis*

5.1 Introduction

Recent decades have seen the emergence of financial institutions that perform bank-like activities outside of the regulated (traditional) banking system. This so-called shadow banking system has now reached a size comparable to that of the traditional banking system, representing about one-fourth of total financial intermediation worldwide.[1] The collapse of shadow banking in 2007 to 2008 has arguably played a role in threatening traditional banks' stability and bringing about the financial crisis. The crisis started with a run on shadow banks that endangered the stability of the entire financial system, raising important questions. Why are there two types of banks? How do different types of banks interact in a crisis? Are traditional and shadow banks substitutes or complements? This paper offers the first model of financial intermediation where both a regulated and an unregulated financial sector coexist and interact, while replicating the following facts from the crisis: (i) liabilities transfer from shadow to traditional banks, (ii) assets transfer from shadow to traditional banks, and (iii) fire sales of assets.

Existing theories of traditional and shadow banking emphasize the substitutability between the two. Given that shadow banks are not subject to the regulations that pertain to traditional banks, these regulations might spur financial intermediation into shadow banking to exploit regulatory arbitrage. This view emphasizes the regulatory costs of traditional banks, failing to explain why traditional and shadow banks coexist and omitting the fact that the two bank types behaved differently in the crisis. As shown by He, Khang, and Krishnamurthy (31), a striking feature of the US financial crisis is that both assets and liabilities moved from shadow to traditional banks. Some assets such as mortgage-backed securities were sold at fire-sale prices. We document these three facts. First, almost \$600 billion of deposits and borrowings went into the largest traditional banks in 2008q3. This happened in less than a month, concomitantly to a wide run on the shadow banking system. Second, there was a mirror image in terms of asset flows, with

[1] This estimate is in terms of credit intermediation IMF (see 32). For empirical descriptions of shadow banking, see Pozsar, Adrian, Ashcraft, and Boesky (43) for the United States, ESRB (19) for the European Union, IMF (32) and FSB (20) for global estimates. Globally, shadow banks' assets were worth \$80 trillion in 2014, up from \$26 trillion more than a decade earlier (FSB (20)).

approximately \$800 billion assets out of shadow banks and \$550 billion into traditional banks from 2007q4 to 2009q1. Third, some assets were sold at fire sale prices; notably mortgage-backed government-agency securities. These facts suggest some form of complementarity between bank types. We propose a theory of the coexistence of traditional and shadow banks that accommodates the above stylized facts. The theory is based on their interaction in a crisis.

Our model describes the different technologies used by traditional and shadow banks to issue riskless claims against risky collateral. On the one hand, traditional banks have access to a guarantee fund to issue riskless claims in a crisis. This access also enables them to issue riskless claims outside a crisis, because these claims can be rolled-over in a crisis. Access to the guarantee fund comes at the cost of higher regulation for traditional banks. On the other hand, shadow banks rely upon traditional banks' ability to issue riskless claims in a crisis, to absorb their assets and provide them with enough liquidity to reimburse their creditors.

If traditional and shadow banks cannot trade assets in a crisis, traditional banks use their access to the guarantee fund only to roll-over their debt. If traditional banks cannot purchase assets from shadow banks in a crisis, these latter are unable to issue debt before a crisis. Therefore absent a secondary market for assets traditional banks are levered but shadow banks are not, and the two bank types are substitutes. Instead, if we allow traditional and shadow banks to trade assets in a crisis, both bank types gain from asset trade and traditional banks purchase assets from shadow banks. This interaction in the asset market enables traditional banks to intermediate government insurance to shadow banks, generating a complementarity between the two forms of financial intermediation. This complementarity is the key message of this paper: the more shadow banks in the system, the lower the price traditional banks have to pay for shadow banks' assets in a crisis, and the better off traditional banks. Conversely, the more traditional banks, the higher the (indirect) support from the guarantee fund to the entire financial system in a crisis and the better off the shadow banks. In that sense, traditional and shadow banks form an ecosystem.

Our model yields two sets of results. We endogenize bankers' choice to enter the shadow versus traditional banking sector, and the first result is that coexistence of traditional and shadow banks can arise in equilibrium. It is intrinsically linked to fire sales of assets from shadow to traditional banks in a crisis. In a crisis, traditional banks' ability to issue riskless debt is limited,[2] therefore traditional banks require a discount on asset prices in a crisis to forgo investment opportunities before a crisis, which endogenously determines asset (fire sale) prices in a crisis. Because shadow banks must sell assets in a crisis to repay their creditors, asset prices determine the amount of riskless debt that shadow banks can issue before a crisis. Both banks' expected profits and thereby bankers' choice to enter either type of banking sector is linked to the other banks' decisions through the channel of asset fire sales in a crisis. From a normative perspective, when traditionnal and shadow banks coexist, this fire sale is at the root of a form of pecuniary externality, which leads to a too high allocation of bankers into the shadow banking sector: bankers, when choosing their allocation between the two types of intermediation technology fail to internalize that allocating themselves towards the shadow banking sector yields a reduced support of the traditional banking sector in times of crises, hence constraining all shadow banks in their ability to leverage in normal times.

Second, we find two competing effects from a reduction in traditional banks' ability to issue deposits in a crisis.[3] On the one hand, this reduces traditional banks' advantage. Traditional banks have access to a lower level of guarantee, which limits their leverage at all times, hence exerting a downward pressure on their profits. This makes traditional banking relatively less profitable than shadow banking, thereby inducing bankers to enter the shadow banking sector. We call this direct effect a substitution effect.[4] On the other hand, note that shadow banks indirectly benefit from traditional banks' access to

[2]We think of the guarantee fund as a governmental entity. Therefore this limit captures some form of limited fiscal capacity that prevents the guarantee fund from insuring too large an amount of traditional banks debt.

[3]The reverse reasoning holds true for an increase in traditional banks' ability to issue deposits in a crisis. The reason why we consider small reductions in traditional banks' support in a crisis is that large reductions in traditional banks' support in a crisis wipe out traditional banks from the market, which is an effect already emphasized in the literature.

[4]It is reminiscent of the argument at play in existing models of shadow banking as regulatory arbitrage (see e.g. Plantin (42), Ordonez (41), Harris, Opp, and Opp (29)).

the guarantee fund through the (secondary) asset market in a crisis. Lowering traditional banks' guarantee in a crisis reduces the support that they provide to shadow banks. This decreases the price at which assets can be sold on the secondary market, hence reducing shadow banks' profits and bankers' incentives to enter the shadow banking sector. We call this effect an income effect. Overall, we find that lower support to traditional banks in a crisis reduces asset prices to such an extent that more bankers choose to enter the traditional banking sector *ex-ante*, i.e. the complementarity effect outweighs the substitution one.

Related literature We argue that traditional banks still compete with but also complement shadow banks, because of their ability to issue deposits in a crisis. Diamond and Dybvig (17) is the seminal paper providing a rationale for traditional banks' deposit insurance that is based on the elimination of depositors' incentives to run their bank. Merton (40) and Rajan (44, 45) are early discussions questioning the future of traditional banks,[5] which suggest that many of the services provided by traditional banks can be sustained by other types of banks in the modern institutional environment.

In our model, asset fire sales are key to understanding the relationship between traditional and shadow banks. As in Shleifer and Vishny (47), the price of assets sold during the crisis is the price at which the best users of these assets (traditional banks in our model) can pay, given they are limited in their ability to issue deposits. Shleifer and Vishny (48) and Gromb and Vayanos (27) model fire sales during which mispricing occurs due to frictions on arbitrageurs' funding capacity. Acharya, Gromb, and Yorulmazer (3) study interbank lending and asset sales when some banks have market power vis-a-vis other banks. Diamond and Rajan (18) discuss liquidity risks on both sides of banks' balance sheet, and inefficient exposure to fire sales.

A second group of theories relates to traditional banks' regulation and their coexistence with shadow banks.[6] Hanson, Shleifer, Stein, and Vishny (28) are interested in the

[5] Other early discussions of the evolution of the financial landscape are Boyd and Gertler (10) and James and Houston (33).

[6] Our model is in line with theories of financial intermediation as issuers of riskfree claims. A seminal paper is Gorton and Pennacchi (26), and other papers include Stein (50), DeAngelo and Stulz (16)

implications of traditional versus shadow banking businesses in terms of the assets that are held by financial intermediaries. Plantin (42) studies optimal bank capital regulation in the presence of shadow banking, and finds that the optimal regulation needs not be in line with current regulatory reforms. Ordonez (41) proposes a model in which reputational concerns are an effective disciplining device in the shadow banking sector. When reputation concerns are weak, banks can only operate using traditional banking. Harris et al. (29) develop a model where capital requirements reduce banks' risk-taking incentives while lowering their funding capacity, and discuss the cyclicality of optimal bank capital regulation in light of the amount of capital in the economy. Gornicka (24) and Luck and Schempp (38) present models where a crisis in the shadow banking sector transmits to the traditional banking sector through guarantees to shadow banks.

Finally, a strand of the literature ties banks' investment choices with asset markets, in line with the mechanism at play in our model. This approach substantially improves the quantitative dynamics of risk premia in crisis episodes where intermediaries' equity capital is scarce. Major contributions in this literature include Adrian and Boyarchenko (5), Brunnermeier and Sannikov (11), He and Krishnamurthy (30) and Viswanathan and Rampini (52).

The paper proceeds as follows. Section 5.2 presents stylized facts about traditional and shadow banks during the crisis. Section 5.3 presents the model, and we analyze the possible coexistence between shadow and traditional banks in Section 5.4. In Section 5.5 we discuss the substitution and complementarity effects, and the implications of our model. Section 5.6 concludes.

5.2 Motivating evidence

In this section, we document three stylized facts using data from the Financial Accounts of the United States (henceforth FAUS), the Federal Reserve H8 Releases and the quarterly Call Reports. From the FAUS data we define traditional banks as the private depository

---

and Plantin (42). As in Gennaioli, Shleifer, and Vishny (23), Krishnamurthy (36), Caballero and Krishnamurthy (13) or Caballero and Farhi (12), we model households' demand for safety as stemming from households' risk aversion.

institutions (L.110). Those institutions are composed of U.S.-chartered depository institutions (L.111), foreign banking offices (L.112), banks in U.S.-affiliated areas (L.113) and credit unions (L.114). Although our stylized facts do not rely on a precise definition of shadow banking using the FAUS, quantitative results depend on which institutions we identify as part of the shadow banking sector. We adopt the view of shadow banking as a chain of market-based transactions among legal institutions which, taken together, perform maturity transformation activities comparable to that of traditional banks. We choose to include money market mutual funds (L.121), mutual funds (L.122), issuers of asset-backed securities (L.127) and security brokers and dealers (L.130), as part of the shadow banking sector.[7] Finally, we use Krishnamurthy and Vissing-Jorgensen (37)'s definition of short-term debt from the FAUS, 60% of which is composed of small time and savings deposits in the 2007-09 period. We provide details on data construction in Appendix 5.A.1.

5.2.1 Fact 1: Liabilities flow from shadow to traditional banks

In the early phase of the 2007 financial crisis, investors stopped rolling over shadow banks' short-term funding. Gorton and Metrick (25) and Copeland, Martin, and Walker (15) document investors' run on their major yet unstable source of funding: the sale and repurchase market (the "repo" market). Another important fact we emphasize is the deposit inflow on traditional banks' balance-sheets. Table 5.1 shows the evolution of short-term debt for traditional and shadow banks from 2006q4 to 2011q1. It is apparent from Table 5.1 that there was a concomitant run on shadow banks and an inflow of short-term debt into traditional banks starting 2008q3. This inflow of deposits in turbulent times is the risk management motive emphasized in Kashyap, Rajan, and Stein (34) to explain why traditional banks combine demand deposits with loan commitments or lines

[7] We build on earlier descriptive work for our definition of the shadow banking system, see e.g. Pozsar et al. (43), or Adrian and Shin (6). Shin (46) and Hanson et al. (28) provide examples of what shadow banks are in the real world. The simplest example, from Hanson et al. (28), is the following: a money market fund that invests in assets that are riskless and issues what households perceive as deposits. The seniority and collateralization of sale and repurchase agreements ("repo" transactions), together with equity capital invested along the intermediation chain, renders shadow-banks "deposit-like" claims perfectly safe, and valued as such by households.

Cumulative flows since 2006q4		
	Shadow banks ($Bill)	Traditional banks ($Bill)
2007q2	+437	+106
2007q3	+606	+230
2007q4	+514	+454
2008q1	+378	+591
2008q2	+623	+732
2008q3	+284	+751
2008q4	+505	+1104
2009q1	-277	+1733
2009q2	-670	+1659
2009q3	-966	+ 1428
2009q4	-1132	+1436
2010q1	-1353	+1409
2010q2	-1354	+1431
2010q3	-1412	+1317
2010q4	-1440	+1420
2011q1	-1471	+1596
2011q2	-1398	+2011

Table 5.1: Traditional and shadow banks: short-term debt inflows (negative values denote outflows) source: Financial Accounts of the United States. We define traditional, shadow banks, and short-term debt in Appendix 5.A.1.

of credit. In a crisis, borrowers draw down on their credit lines while investors seek a safe haven for their wealth, turning to traditional banks because these latter provide insurance due to the government guarantee on their deposits.

Gatev and Strahan (22) emphasize that it is traditional banks' access to federal deposit insurance that causes economy's savings to move into traditional bank deposits during times of aggregate stress,[8] providing banks with the unique ability to hedge against systematic liquidity shocks. Nevertheless, Acharya and Mora (2) show and it is apparent from Figure 5.1 that it was not until the U.S. government's intervention just before the Lehman failure on September 15, 2008 that deposit flew into traditional banks. He et al.

[8]This explains why there is no evidence that funds flowed into the banking system when spreads widened during the 1920s, prior to the expansion of the federal safety net with the creation of federal deposit insurance.

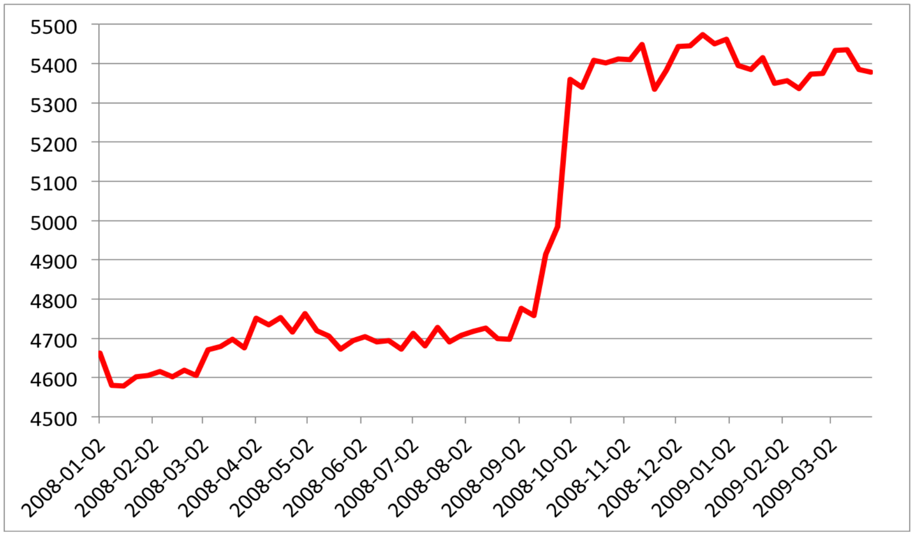


Figure 5.1: Large traditional banks: deposits and borrowings (stocks in $ bn)

source: Fed H8 Releases

(31) find that core deposits eventually increased by close to $800 billion by early 2009. Weekly times series in Figure 5.1 show a sudden $600 billion deposits and borrowings inflow into the largest US traditional banks in just a few days, from September 10th to October 1st, 2008.

The flow of deposits into traditional banks illustrates the fact that not all entities of the U.S. financial sector deleveraged in the crisis. Ang, Gorovyy, and van Inwegen (7) show that hedge fund leverage decreases prior to and during the financial crisis from mid-2007 onwards, He et al. (31) show that leverage of banks and investment banks continues to increase. This helps to put deleveraging into perspective. At the worst periods of the financial crisis in late 2008, hedge fund leverage is at its lowest while the leverage of banks is at its highest. Although traditional banks issued new equity during the crisis, in Appendix 5.A.2 we show the evolution of traditional banks' market and book equity over the crisis.[9] We find that traditional banks did not issue enough new equity during the

[9]See Baron (8) for evidence of banks' countercyclical equity issuance.

crisis to compensate for their market losses. As a result, the deposit flow into traditional banks increased their market leverage in the crisis.

5.2.2 Fact 2: Asset flow from shadow to traditional banks

Figure 5.2 illustrates that assets were massively reshuffled among financial intermediaries during the financial crisis. In particular, assets flew out of shadow banks and into traditional banks. Although our data does not allow us to identify whether these changes were due to changes in the value of assets or changes in ownership, empirical work by He et al. (31) and Bigio, Begenau, and Majerovitz (9) provide estimates of the amount of assets that were transferred from shadow to traditional banks during the crisis. From 2007q4 to 2009q1, He et al. (31) find that shadow banks decreased their holdings of securitized assets by approximately \$800 billion while traditional banks increased theirs by approximately \$550 billion. Looking at the wider period from 2007q1 to 2013q1 and considering total asset holdings, Bigio et al. (9) document a net asset outlfow of \$1702 billion out of shadow banks and an asset inflow of \$1595 billion into traditional banks.[10]

The main argument that explains why the shadow banking system developed is that traditional banks create off balance-sheets entities, because holding loans on balance sheets is not profitable for them see e.g. Gorton and Metrick (25), Acharya, Schnabl, and Suarez (4). For instance special conduits are comparable to regular banks in many ways, and they often are managed by traditional banks. Pozsar et al. (43) dubs shadow banks managed by traditional banks "internal shadow banking". However, the flip side of this off-balance sheet leverage in the shadow banking sector is that liquidity guarantees are provided by traditional banks. Acharya et al. (4) show that investors in conduits covered by guarantees were repaid in full. This implies a monetary transfer from the traditional to the shadow banking system in the crisis, and the mirror asset transfer from shadow to traditional banks that we see on Figure 5.2. Although we do not explicitly model the contractual relationship between traditional and shadow banks in terms of liquidity guarantees,[11] our

[10] Another important aspect of this asset transfer is the sizable purchase of assets from the Federal Reserve, which balance sheets increased by approximately \$1954 billion, as calculated in Bigio et al. (9).

[11] Luck and Schempp (38) give an account of how contractual linkages between regulated banking and shadow banking affect financial stability.

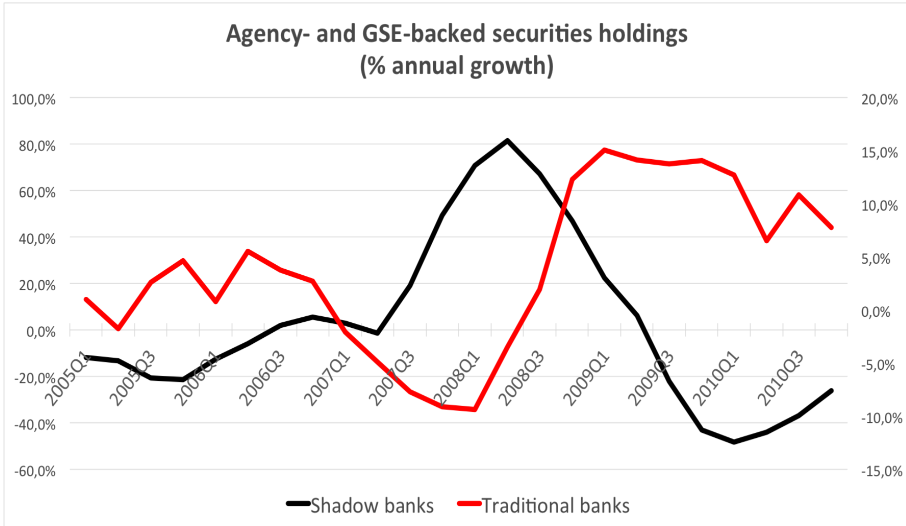


Figure 5.2: Traditional and shadow banks: asset inflows (negative values denote outflows)
source: Financial Account of the United States

model sheds light on the interaction between the two bank types and the cost/benefit analysis driving traditional banks' choices on whether or not to develop shadow banking.

One main testable prediction of our theory is that traditional banks are able to purchase assets from shadow banks in a crisis, insofar as they benefit from a guarantee on their deposits. This guarantee indeed enables them to attract deposits precisely when shadow banks have to reimburse their creditors. Publicly available data on purchases/sales of assets by traditional and shadow banks during the crisis is not available. Therefore we try to estimate purchases/sales of mortgage-backed securities (henceforth MBS) applying He et al. (31)'s methodology on traditional banks' regulatory data from the Call Reports. We observe the total value of MBS holdings by each traditional bank before the crisis (denote it $P\_2007q4 * MBS\_2007q4_i$ where P_2007q4 is the fair price of MBS securities in 2007q4 and MBS_2007q4 is the quantity of MBS held by bank i in 2007q4) and after the crisis ($P\_2009q1 * MBS\_2009q1_i$). Besides, denoting f the repayment/maturity rate of MBS net of the new issuance rate during the period from 2007q4 to 2009q1, the International Financial Reporting Standards (IFRS) give us the following accounting identity:

$$P\_2009q1 * MBS\_2009q1_i - P\_2007q4 * MBS\_2007q4_i * (1 - f) = MBSPurchases_i - MBSLosses_i$$

As in He et al. (31), we test three different scenarii based on (i) the total losses that traditional banks incurred on MBS assets during the 2008 crisis, and (ii) Bloomberg WDCI estimates for the net repayment rate f. Under scenario 1 the repayment rate used to construct the MBS_Purchases variable is 7% and total losses imputed to the financial sector are $500 billion.[12] Under scenario 2, the repayment rate is 12% and total losses are $176 billion. Under the "naive" scenario, we do not correct for the net repayment rate nor total losses.

We analyze the data formally by running the following OLS regression on changes in various items of traditional banks' balance sheets from 2007q4 to 2009q1:

$$\begin{aligned} MBSPurchases_i = \;& \beta_0 + \beta_1.Liquidity_i + \beta_2.Solvency_i + \beta_3.Insured.\_deposits_i \\ & + \beta_4.Non\_insured\_deposits_i + \beta_5.Credit\_granted\_i \\ & + \beta_6.Evolutions\_i + \beta_7.controls\_i + \epsilon_i \end{aligned}$$

where $MBSPurchases_i$ is our estimated purchases/sales of mortgage-backed securities by traditional bank i normalized by total assets (banks are aggregated to the top holder level in the Call Reports).

Details about the sample construction and the results are in Appendix 5.A.4. We find support for the central mechanisms of our theory: Traditional banks purchased mortgage-backed securities in the crisis, and the more so the more deposits they issued in the crisis. Those assets are precisely the ones shadow banks have sold (see Figure 5.2). Besides, we find that those banks that purchased assets in the crisis did so at the expense of credit. This is in line with Shleifer and Vishny (49) and Stein (51) who discuss how market

[12] Note that the only available estimate on MBS losses in the crisis is an aggregate over the traditional banking sector from the IMF's Global Financial Stability Report of October 2008 and Bloomberg WDCI (which explains why we test two scenarii thereafter). Denote those estimates for the entire traditional banking sector losses on MBS assets $MBSLosses$. Although we try to estimate MBS purchases/losses by taking into account potential losses on those assets when using the change in MBS holdings from 2007q4 to 2009q1 adjusted for the net repayment/maturity rate, we cannot account for differences in losses across traditional banks. We therefore assume that losses incurred by traditional banks are proportional to the amount of MBS they hold, so that $MBSLosses_i = \frac{MBS\_2007q4_i}{\sum_k MBS\_2007q4_k} * MBSLosses$ and $\sum_k MBSLosses_k = MBSLosses$.

conditions shape the allocation of scarce bank capital across lending and asset purchases. Using German data, Abbassi, Iyer, Peydró, and Tous (1) find comparable results.

5.2.3 Fact 3: Asset fire sales

Many examples in the literature suggest that asset prices have deviated significantly from "fundamental values" and were sold at fire-sale prices during the crisis. Using data on insurance companies, Merrill, Nadauld, Stulz, and Sherlund (39) show that risk-sensitive capital requirements, together with mark- to-market accounting, can cause financial intermediaries to engage in fire sales of RBMS securities. Krishnamurthy (35) discusses pricing relationships reflecting similar distortions on agency MBS, and notably the increasing option-adjusted spread of Ginnie Mae MBS versus the US Treasury with the same maturity. Gagnon, Raskin, Remache, and Sack (21) also document substantial spreads on MBS rates - well above historical norms. Such evidence of high spreads on a security which has no credit risk points to the scarcity of arbitrage capital in the marketplace and the large effects that this shortage can have on asset prices. Using micro-data on insurers' and mutual funds' bond holdings, Chernenko, Hanson, and Sunderam (14) finds that in order to meet their liquidity needs during the crisis, investors traded in more liquid securities such as government-guaranteed MBS. This strategy is consistent with theories of fire sales where investors follow optimal liquidation strategies: although spreads on GSE MBS were very high in the fall of 2008, those assets remained the most liquid ones in securitization markets at that time.

Our illustration of asset fire sales comes from Gorton and Metrick (25). The authors provide a snapshot of fire sales of assets that occurred due to the financial crisis that we reproduce on Figure 5.3. We see a negative spread between higher- and lower-rate bonds with the same maturity. Aaa-rated corporate bonds would normally trade at higher prices (i.e. lower spreads) than any lower-grade bonds with the same maturity (say, Aa-rated ones), and this negative spread is an evidence of such an important amount of Aaa-rated corporate bonds sales that the spread must rise to attract buyers.

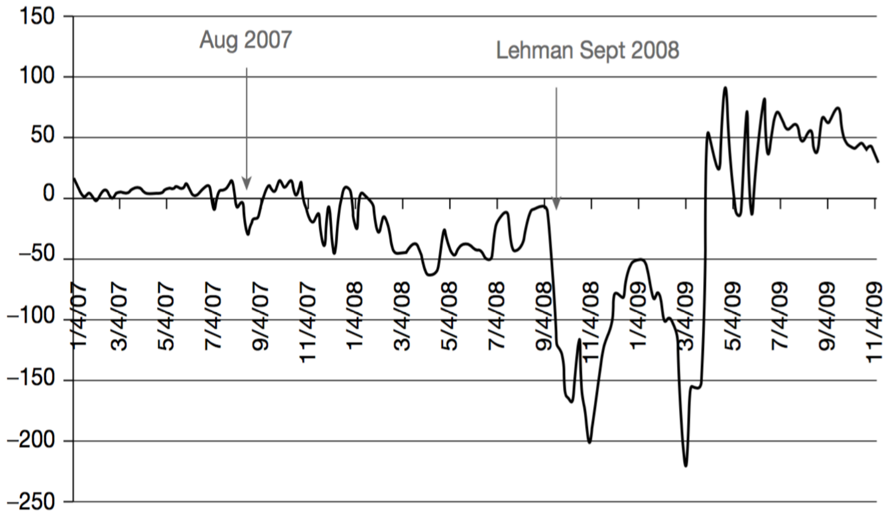


Figure 5.3: Interest rate spread: 5year AA-AAA Industrials
source: Gorton and Metrick (2009)

5.3 Model setting

The model features a closed economy, with three dates ($t = 0, 1, 2$), one production technology, two sets of agents (households and bankers) and two types of goods (consumption goods and capital goods).[13] Date 1 includes two states $\{G, B\}$ and date 2 three states $\{GG, BG, BB\}$.

5.3.1 Households

A unit mass of households is endowed with a large quantity of consumption goods at each date. They can consume at each date and do not discount future consumption. They have linear preferences over consumption at all dates.

At each date $t \in \{0,1\}$, and in each state ω, household's utility function writes as follows:

$$U_{t,\omega} = C_{t,\omega} + \mathbb{E}_{t,\omega}\left[U_{t+1}\right] \tag{5.1}$$

[13] The model is broadly inspired by Stein (50). However, we consider two types of banks: traditional and shadow banks. It is different from the framework of Hanson et al. (28) in that in our model traditional banks can issue debt in a crisis, and bankers endogenously choose which type of banking sector they enter.

with

$$U_{2,\omega} = C_{2,\omega} \tag{5.2}$$

Households are not able to invest directly in physical projects, and can only invest in financial claims issued by banks, which undertake the investment.[14]

5.3.2 Bankers

There is a unit mass of identical bankers, who start at $t = 0$ with an endowment n of consumption goods. As households, they are risk neutral and indifferent between consuming at $t = 0, 1, 2$. Each of them chooses the probability at which they are willing to set up a traditional bank (T-bank), $1 - \chi^S$, and a shadow bank (S-bank), χ^S. Once allocated to one sector or the other, they invest a quantity $n^i \in [0, n]$ ($i = \{S, T\}$) into the firm and consume whatever is left from this endowment. The funds invested in the bank constitute the own funds of the bank.

Banks' investment technology

Both T- and S-banks have access to a unique investment technology, whose payoffs are summarized on Figure 5.4.

Investing one unit of capital good in the investment technology at $t = 0$ yields a risky payoff $z \in \{R, r, 0\}$ in terms of consumption goods at $t = 2$, in each respective state of $\Omega_2 \equiv \{GG, BG, BB\}$. At this date, investment pays off and all capital goods is destroyed. At $t = 1$, information about the occurrence of date 2 states is revealed: when state G (good news) materializes (which occurs with probability p), it is known with certainty that state $\{GG\}$ will take place at $t = 2$ so that investment pays off R, and all uncertainty is resolved. However, when state B (bad news) materializes (which occurs with probability $1 - p$), it is learnt that there is a non-zero probability of 0 output at $t = 2$. The probabilities that

[14] As in Stein (50) we abstract from any agency problem between the intermediary and the firm manager and assume that the bank has all the bargaining power in the banking-firm relationship, thereby enabling it to extract all the profit from the investment and leaving the firm with no profit in expectation.

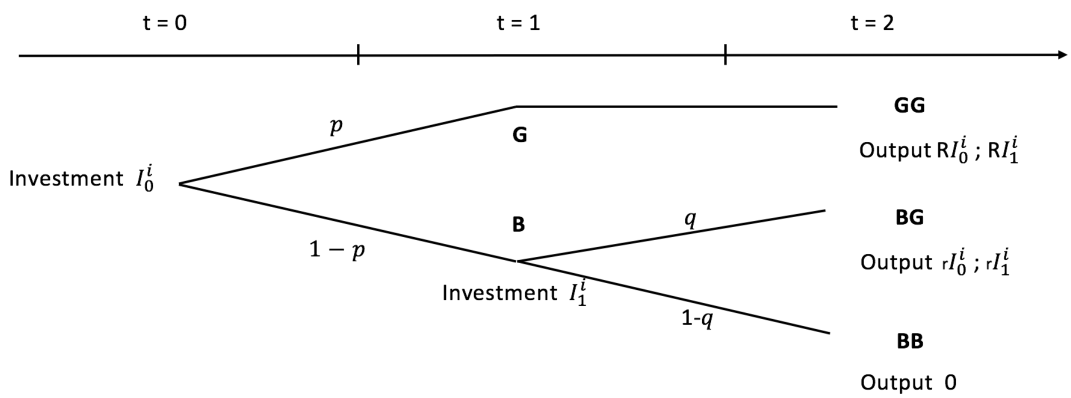


Figure 5.4: Investments payoffs (states in bold font)

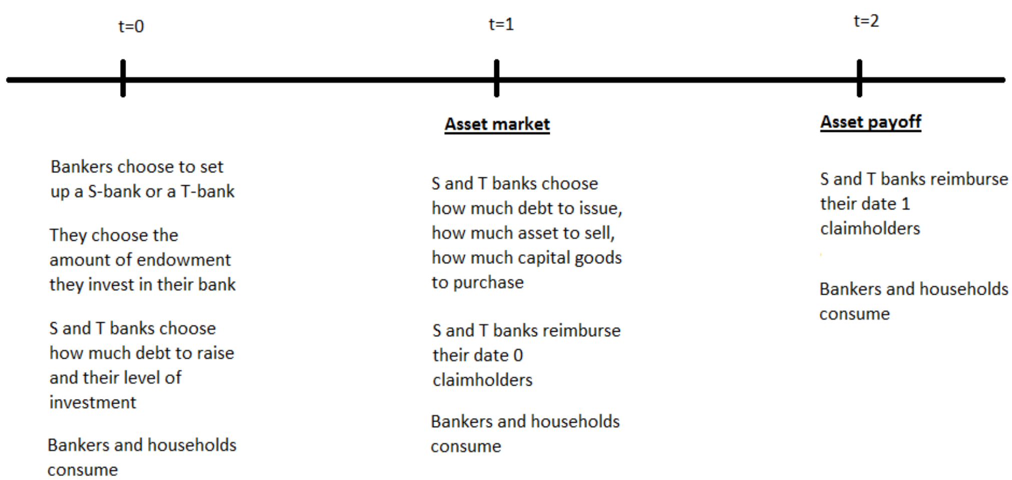


Figure 5.5: Timeline of the model

each state $\{BG, BB\}$ materializes are $(q, (1-q))$. In this case, there is an aggregate risk in the economy, that cannot be diversified away through other forms of investments.

Bank's choices

The timeline of the model is detailed in Figure 5.5.

Time 0 At $t = 0$, both T- and S-banks have the ability to transform consumption goods into capital goods (or physical assets) one-for-one, in order to invest I_0^i ($i = \{S, T\}$) units of capital goods in the long-term productive investment technology. To fund their investments, they can add to their endowment n^i, D_0^i units of funds raised from the households. Raising funds can be done in the exclusive form of riskless short-term debt, issued on a competitive market.

This assumption captures what we see as one fundamental role of banks: their ability to act as safety and liquidity providers to the households at all times. The key assumption here is that banks have to finance themselves with riskless financial contracts. Our results do not rely on the specific form of these contracts. [15]

Time 1 At $t = 1$, T- and S-banks can't either transform consumption goods into capital goods, or transform capital goods into consumption goods. They can however trade capital goods for consumption goods in a competitive secondary market, where all banks participate. In each state $\omega_1 \in \{B, G\}$, they purchase I_{1,ω_1}^i ($i = \{S, T\}$) units of capital goods at a market price $p_{1,B}qr$ at $t = 1$ in state B and $p_{1,G}R$ at $t = 1$ in state G. These capital goods are then reinvested in the same technology. [16]

Banks can also sell on the market a share $(1 - \alpha_{1,\omega_1}^i)$ of their stock of physical assets. We assume liquidation costs for banks such that a share $\varepsilon \in (0, 1]$ of assets sold is destroyed. This cost assumption is meant to capture the forsaken returns from liquidating illiquid projects. [17]

Finally, banks have the ability to raise additional funds from households, D_{1,ω_1}^i, by issuing risk-less short-term debt, either to roll previously issued debt over, or to finance purchases. The debt issued at $t = 0$ must be paid back, such that $r_0 D_0^i$ units of consumption

[15] Alternative model specifications closer to the safe asset literature would yield similar results. One could for instance endow the households with an infinite risk-aversion utility function as in Gennaioli et al. (23) or Epstein-Zin preferences with infinite relative risk aversion and infinite intertemporal elasticity of substitution as in Caballero and Farhi (12).

[16] We can alternatively think of such trade as a transfer of ownership rights on the payoffs generated by the capital goods

[17] One can also interpret ε as the cost of breaking up a lending relationship, or the loss associated to a loosened monitoring ability induced by a change of ownership. The adjustment cost $(1 - \varepsilon)$ is a form of technological illiquidity, whose importance is emphasized in Brunnermeier and Sannikov (11).

goods must be provided to the debt-holders, where r_0 is the interest rate on the debt issued at $t = 0$, pinned down on the market.

Time 2 At $t = 2$, the investment pays off in terms of consumption goods, and all capital goods are destroyed. T and S banks reimburse their date 1 debt-holders by providing them with $r_{1,\omega_1} D^i_{1,\omega_1}$ (with $\omega_1 \in \{B, G\}$) units of consumption goods, where r_{1,ω_1} is the interest rate on the debt issued at $t = 1$ in state $\omega_1 \in \{B, G\}$. It is important to note that in each date and state, banks are subject to limited liability constraints.

Differences between shadow and traditional banks

We distinguish traditional from shadow banking by making the following assumptions.

Assumption 1 *[Differences] Traditional and shadow banks differ in two ways:*

1. *T-banks have access to a guarantee fund at $t = 1$ in state B. This enables them to issue risk-less claims that promise to pay up to an amount $k > 0$.*

2. *T-banks have to cope with higher operating costs : At $t = 2$, T-banks only get a fraction $\delta \in [0, 1]$ of the payoff generated by their investments.*

T-banks benefit from an advantage over S-banks in $t = 1$ in state B they have access to a government guarantee which enables them to issue risk-less short-term debt from $t = 1$ in state B to $t = 2$, even if the productive technology is facing a risk of zero output. In our setting, it takes the form of a fairly-priced guarantee fund owned by the state: to get one unit of consumption good at $t = 2$ in state BB, bankers have to provide $\frac{1-q}{q}$ units of consumption goods to the fund at $t = 2$ in state BG such that the government is making no profit in expectation from setting this fund up, and provides no subsidy to the T-banks through this fund. We assume that the maximum guarantee each T-bank can benefit from is set to k, which is a structural parameter of our economy. One interpretation for this is that the government's ability to enforce payments made by T-banks to the fund at $t = 1$ in state G is limited, for instance due to limited fiscal capacity. Another interpretation for this parameter could be a reduced form for informational friction which prevents T-banks from taking too much debt at $t = 1$ in state B[18]. Note that, T-banks being subject to limited liability constraint at all times and states, k is not the only determinant of T-banks debt level, as will be emphasized below.

In return to this advantage, T-banks face regulatory costs, which take the form of a reduced payoff on their investment at $t = 2$ in each state. This cost captures a wide variety of costs associated to higher regulations imposed to the traditional banking sector : taxes to finance the regulating entities, the general functioning costs of the guarantee fund, costs to generate regulatory information, etc. [19]

5.3.3 Equilibrium definition

In this setting, we formally define an equilibrium as follows.

Definition 1 *[Equilibrium] An equilibrium is defined by a set* $\{\chi^S, n^S, n^T, v^S, v^T, p_{1,G}, p_{1,B}, r_0, r_{1,G}, r_{1,B}\}$ *with*

$$v^i = \{I_0^i, D_0^i, I_{1,G}^i, I_{1,B}^i, D_{1,G}^i, D_{1,B}^i, \alpha_{1,G}^i, \alpha_{1,B}^i\}$$

for $i \in \{S, T\}$ *such that*

1. v^T *maximizes traditional bank's date 0 value function* $V_0^T\left(p_{1,G}, p_{1,B}, r_0, r_{1,G}, r_{1,B}, n^T\right)$.
2. v^S *maximizes shadow bank's date 0 value function* $V_0^S\left(p_{1,G}, p_{1,B}, r_0, r_{1,G}, r_{1,B}, n^S\right)$.
3. n^i *maximizes the expected payoff of a banker's allocated in type i-bank (*$i \in \{S, T\}$*)* $V_0^{i,B}\left(p_{1,G}, p_{1,B}, r_0, r_{1,G}, r_{1,B}\right)$.
4. χ^S *is a solution to the allocation problem:*

$$\max_{\chi^S \in [0;1]} \chi^S V_0^{S,B}\left(\ p_{1,G}, p_{1,B}, r_0, r_{1,G}, r_{1,B}\right) + (1 - \chi^S) V_0^{T,B}\left(p_{1,G}, p_{1,B}, r_0, r_{1,G}, r_{1,B}\right).$$

[18] For such an interpretation, we will see that it is not restrictive not to impose the same constraint on S-bank: the existence of a risk of zero-payoff at $t = 2$ in state BB prevents S-banks to issue any type of risk-less debt at $t = 1$ in state B making such a constraint superfluous

[19] From a positive perspective, it can also be interpreted more broadly as a series of costs associated to the T-banks specificities in terms of business model: they have higher operating costs, employ more workers, provide more services to their customers etc.

5. *Markets for short-term debt clear at time 0, and 1, in states G and B for respective interest rates* $\{r_0, r_{1,G}, r_{1,B}\}$.

6. *Capital goods market clear at time 1 in states G and B at respective prices* $\{p_{1,G}qr, p_{1,B}R\}$.

We now turn to the analysis of this equilibrium, detail the banks and bankers' problems and value functions, as well as the market clearing conditions, and develop the solutions.

5.4 Model analysis

5.4.1 Assumptions

Assumptions on expected returns In Assumption 2 we make two restrictions about the investment technology.

Assumption 2 *[Returns] At* $t = 0$*, investing is efficient for both T- and S-banks:*

$$\delta\left(pR + (1-p)qr\right) > 1 \tag{5.3}$$

At $t = 1$ *in state* B*, expected returns are lower than one:*

$$qr < 1 \tag{5.4}$$

Condition (5.3) ensures that as of $t = 0$, investing is efficient. This implies that each banker invests her full endowment in the bank it sets up, be it a S-bank ($n^S = n$) or a T-bank ($n^T = n$), and each bank invests all her own funds in their time 0 investment technology. Condition (5.4) reflects the fact that in a crisis, asset returns are lower than in good state.

Assumptions on traditional bank's parameters We make two additional assumptions on the size of the guarantee traditional banks can benefit from, k, as well as its associated cost, δ, which helps us focus on the main cases of interest.

Assumption 3 *[Banks' parameters] The cost associated to traditional banking activity is low enough to prevent asset sales at $t = 1$ in state G:*

$$\delta > 1 - \varepsilon \tag{5.5}$$

The guarantee T-banks can benefit from on his time 1, B claims is the binding constraint for time 0, T-banks debt issuance:

$$k < k^* = \frac{\delta q r}{1 - \delta q r} n \tag{5.6}$$

Condition (5.5) enables us to rule out asset transfers at $t = 1$ in state G between the two types of intermediaries: it will always be optimal for any type of bank to choose to continue their time 0 investment, at $t = 1$ in state G rather than providing assets on the market for capital goods. This assumption could be relaxed at little cost. It is however convenient in keeping the analysis focused on the interaction between the two types of banks in the bad information state, time 1. Relaxing this assumption generates asset transfers from T-banks to S-banks in the good information state: indeed, if $\delta < 1 - \varepsilon$, it is more valuable for a T-bank to sell capital goods on the market to S-banks who value it more [20], and incur the illiquidity cost ε, than continuing its investment and incur the regulatory cost $1 - \delta$ on its time 2 payoff.

The second assumption (5.6) will imply that the maximum amount of risk-less debt issued at $t = 0$ that can be rolled-over at $t = 1$ in state B by a T-bank will be constrained by the size of the guarantee fund T-banks have access to at $t = 1$ in state B (hence k), and not the limited liability constraint of time 2, BG the bank has to comply with when financing this guarantee. This ensures that T-banks will be able to issue an amount k of risk-less debt both at $t = 0$ and at $t = 1$ in all states.

[20] Each unit of capital good generates a return R for a S-bank in date 2, GG instead of δR for T-banks, and none of them discount future payoffs

Technical assumption Finally, we set the following assumption, in order to ensure coexistence of the two types of institutions.

Assumption 4 *[Assumption for coexistence] We assume that the cost of regulation is high enough such that both S and T-banks can exist in equilibrium*

$$\delta < \frac{p}{\varepsilon\left(p\left(R-1\right)+\left(1-p\right)qr\right)+p} \tag{5.7}$$

We discuss later the fact that condition (5.7) is a necessary condition to find a range of asset prices at $t=1$ in state B such that both T and S-banks are willing to trade assets, given condition (5.6).[21]

5.4.2 Equilibrium implications of the assumptions

We now turn to the resolution of the equilibria of the game. In order to reduce the dimensionality of the exercise and ease further exposition we start by detailing some equilibrium conditions implied by our assumptions. We consider equilibrium conditions associated to the debt market, the asset market at $t=1$ in state G, and discuss their implications.

Debt market clearing conditions

Households are endowed with a large amount of consumption goods at each date, and the ability to generate safe short-term claims in the system being limited,[22] the real rates on short term debt issued by any type of bank is pinned down in equilibrium by the household's utility function. With a linear utility function and no time discounting, this rate is set to 1 in equilibrium.

[21] It is important to note that the five parametric restrictions generated by these assumptions are not mutually exclusive.

[22] It is indeed constrained by the guarantee that can be obtained from the government, which implies that the maximum amount of short term risk-less claims that can be issued at $t=0$ or $t=1$ cannot exceed $k(1-\chi^S)$

In equilibrium:

$$r_0 = r_{1,B} = r_{1,G} = 1$$

It is important to note that, the households being endowed with a large amount of consumption goods in each date and state, the limiting factor in banks' ability to expand will always be their ability to generate safe short-term claims. We prevent in this way mispricing of the debt instruments to play any role in the mispricing of assets that could occur when assets flow from one type of bank to the other.

Time 1, G asset market clearing condition

at $t = 1$ in state G, T and S-banks must reimburse their date 0 debt-holders $r_0 D_0^i = D_0^i$ ($i = S, T$). They can choose to issue short term risk-less debt, sell part of their capital goods to other banks on the asset market, and purchase capital goods sold by others. In this time and state, all uncertainty is resolved, such that both S and T-banks are able to generate short term safe claims $D_{1,G}^i$ ($i = S, T$) to roll their funding over as long as $D_{1,G}^T \leq \delta R I_0^T$ and $D_{1,G}^S \leq R I_0^S$ (due to limited liability constraint at $t = 2$ in state GG).

This strategy is a dominating one for S-banks as well as T-banks (due to assumption (5.5)), as it spares them the destruction of capital goods ε associated to sale of capital goods: if T-banks were to sell their assets on the market, they could get $(1 - \varepsilon) R I_0^T$ as S-banks would be willing to purchase the assets at fair price R, by issuing risk-less short term debt to finance this purchase. This being lower than $\delta R I_0^T$, the value of keeping the assets in place, T-banks will always choose to hold on their time 0 investments, as would S-banks do.

Finally, in equilibrium, no capital goods are supplied on the market, and no trade takes place. Without loss of generality, we simplify S and T-banks programs exposition by presenting and solving them as if no asset market existed at $t = 1$ in state G, and their only option were to roll their funding over, as these two problems are equivalent in equilibrium. We can now turn to the exposition of these modified banks' programs.

5.4.3 Shadow banks' program

We solve a S-bank's program given it has no access to the asset market at $t = 1$ in state G.

Date 1, state G. In such a program, at $t = 1$ in state G, S-banks can only roll their funding over. The value function of a S-bank at $t = 1$ in state G, with a date 0 investment level I_0^S, a debt level D_0^S writes, in case of no-default, as follows:

$$V_{1,G}^{S,ND}\left(I_0^S, D_0^S\right) = RI_0^S - D_0^S$$

and the S-bank doesn't default if and only if $RI_0^S - D_0^S \geq 0$.

At $t = 1$ in state G, if the S-bank has a higher debt level inherited from $t = 0$, it would have to default on it and would not be able to reimburse its creditors.

As S and T-banks can only issue risk-less short term debt in each date and state, such an event would not occur in equilibrium. We take the convention to set, in these cases, the value function to $-\infty$. S-banks value function at $t = 1$ in state G writes:

$$V_{1,G}^{S}\left(I_0^S, D_0^S\right) = \begin{cases} RI_0^S - D_0^S \text{ if } D_0^S \leq RI_0^S \\ -\infty \text{ otherwise} \end{cases}$$

Date 1, state B. At $t = 1$, in state B, S-banks choose how much (if any) funding to raise, how much capital goods to sell on the market, and how much capital goods to purchase on the market, for making additional investments in in their technology. They also have to reimburse their date 0 debt-holders.

At $t = 1$ in state B there is a non-zero probability for the investment technology to return a zero payoff at $t = 2$. Therefore, in the absence of any guarantee on the claims they issue, S-banks are not able to issue any risk-less claim at $t = 1$ in state B. The only way for S-banks to reimburse households is to sell capital goods on the market and channel the proceeds of this sale to households so as to meet their obligations. Shadow banks can also choose to interrupt a higher fraction of their date 0 investment, to sell more capital goods on the market, in order either to consume or purchase capital goods from other banks at $t = 1$ in state B.

Remark that S-banks can only invest at $t = 1$ in state B by interrupting previous investments made at $t = 0$. It can never be optimal for a S-bank to interrupt their time 0 investments, which induces an early liquidation cost ε, in order to sell capital goods on the market at a price $p_{1,B}qr$ by unit of capital goods, purchase new capital goods at the same price and reinvest it in the same investment technology. Such a strategy is indeed strictly dominated by the one that consists in keeping the investments without liquidation.

We therefore write the shadow bank's value function at $t = 1$ in state B in the case of no default, as

$$V_{1,B}^{S,ND}\left(I_0^S, D_0^S, p_{1,B}\right) = \max_{\alpha_{1,B}^S \in [0;1]} \alpha_{1,B}^S qr I_0^S + (1-\alpha_{1,B}^S)(1-\varepsilon)p_{1,B}qrI_0^S - D_0^S$$
$$\text{s.t. } (1-\alpha_{1,B}^S)\,(1-\varepsilon)\,p_{1,B}qrI_0^S \geq D_0^S$$

where I_0^S is the investment level of the S-bank at $t = 0$, $D_0^S$the investment level of the S-bank at $t = 0$, $p_{1,B}qr$ the market price of one unit of capital good, $\alpha_{1,B}^S$ is the share of investment made at $t = 0$ that the S-bank is willing to pursue at $t = 1$ in state B, and the S-bank is subject to a limited liability constraint, as well as positivity and no short-sale constraints.

Keeping the same convention, we set the value function in case of default to $-\infty$.

Lemma 1 *[S-banks time 1] At $t = 1$, in state B, S-banks don't default if and only if $D_0^S \leq (1-\varepsilon)\,p_{1,B}qrI_0^S$.*

If $p_{1,B} > 0$, their value function writes

$$V_{1,B}^S\left(I_0^S, D_0^S, p_{1,B}\right) = \begin{cases} \left((1-\varepsilon)p_{1,B}qrI_0^S - D_0^S\right)\max\left(\frac{1}{(1-\varepsilon)p_{1,B}}; 1\right) & \textit{if } D_0^S \leq (1-\varepsilon)\,p_{1,B}qrI_0^S \\ -\infty & \textit{otherwise} \end{cases}$$

If $p_{1,B} = 0$, their value function writes

$$V_{1,B}^S\left(I_0^S, D_0^S, p_{1,B}\right) = \begin{cases} qrI_0^S & \textit{if } D_0^S = 0 \\ -\infty & \textit{otherwise} \end{cases}$$

Date 0 At $t = 0$, Shadow banks will always choose levels of debt D_0^S and investment I_0^S consistent with an absence of default on their debt at $t = 1$,as they are only able to raise funding in such a way. They also have to face a funding constraint: their time 0 investment is funded with the banker's endowment and the debt raised from households at $t = 0$. [23]

For $p_{1,B} > 0$[24], the date 0 value function of a S-bank, with own funds n^S writes:

$$
\begin{aligned}
V_0^S\left(p_{1,B}, n^S\right) = \max_{D_0^S, I_0^S \geq 0} & \left[(1-p)\left((1-\varepsilon)p_{1,B}qrI_0^S - D_0^S\right)\max\left(\frac{1}{(1-\varepsilon)p_{1,B}}; 1\right)\right. \\
& \left. + p\left(RI_0^S - D_0^S\right) + (D_0^S + n^S - I_0^S)\right] \\
\text{s.t. } & D_0^S + n^S \geq I_0^S \\
& D_0^S \leq (1-\varepsilon)\, p_{1,B}qrI_0^S \\
& D_0^S \leq RI_0^S
\end{aligned}
$$

Denoting $\overline{p}_1^S \equiv \frac{1}{(1-\varepsilon)(qr + \frac{p(R-1)}{1-p})} < \frac{1}{1-\varepsilon}$, we obtain the following proposition.

Proposition 1 *[S-bank program] At $t = 0$, S-banks take the following decisions.*

1. *If $0 \leq p_{1,B} < \overline{p}_1^S$, $D_0^S = 0$, $I_0^S = n^S$. Shadow banks do not issue short-term claims at $t = 0$, i.e. they do not lever themselves. In this case, $V_0^S\left(p_{1,B}, n^S\right) = (1-p)qrn^S + pRn^S$*
2. *If $p_{1,B} = \overline{p}_1^S$, any $D_0^S \in \left[0; \frac{(1-\varepsilon)\overline{p}_1^S qr}{1-(1-\varepsilon)\overline{p}_1^S qr} n^S\right]$ is an equilibrium solution, and $I_0^S = n^S + D_0^S$. Shadow banks sell a fraction $\frac{D_0^S}{(1-\varepsilon)p_{1,B}qr\left(D_0^S + n^S\right)}$ of their capital goods at $t = 1, B$, so as to repay their date-0 creditors. In this case, $V_0^S\left(p_{1,B}, n^S\right) = (1-p)qrn^S + pRn^S$*

[23] Thanks to condition (5.3), we know that bankers will always choose to invest all their endowment into the bank they set up.

[24] If $p_{1,B} = 0$ this reduces to

$$
\begin{aligned}
V_0^S\left(0, n^S\right) = \max_{I_0^S} & (1-p)\left(qrI_0^S\right) + p\left(RI_0^S\right) \\
\text{s.t. } & n^S \geq I_0^S \\
& I_0^S \geq 0
\end{aligned}
$$

3. *If* $\overline{p}_1^S < p_{1,B} < \frac{1}{(1-\varepsilon)qr}$*,* $D_0^S = \frac{(1-\varepsilon)p_{1,B}qr}{1-(1-\varepsilon)p_{1,B}qr}n^S$*,* $I_0^S = n^S + D_0^S$*. Shadow banks sell all their capital goods at* $t = 1$ *in state* B*, so as to repay their date-0 creditors. In this case,* $V_0^S\left(p_{1,B}, n^S\right) = p\left(\frac{R-(1-\varepsilon)p_{1,B}qr}{1-(1-\varepsilon)p_{1,B}qr}\right)n^S$

4. *If* $p_{1,B} \geq \frac{1}{(1-\varepsilon)qr}$*,* $D_0^S = +\infty$*,* $I_0^S = +\infty$ *and* $V_0^S\left(p_{1,B}, n^S\right) = +\infty$

Proof. See appendix. □

Although S-banks do not have access to the guarantee fund, they can issue riskless debt at $t = 0$ insofar as they are backed by the liquidation value of the fraction $(1 - \alpha_{1,B}^S)$ of their existing investment they sell at $t = 1$ in state B. It is S-banks' ability to pull the plug in the crisis that enables them to issue safe short-term debt at $t = 0$. When liquidating at $t = 1$ in state B, proceeds from selling capital goods are $(1 - \alpha_{1,B}^S)(1 - \varepsilon)p_{1,B}qrI_0^S$ where $p_{1,B}qr$ is the price of a unit of capital good in the secondary market at $t = 1$ in state B. The proceeds of this sale depend on T-banks' ability to purchase capital goods in the crisis, which itself relies on the guarantee fund these latter can access.

Indirectly, S-banks therefore rely on the guarantee fund via T-banks, thereby granting T-banks the ability to play the role of government support intermediary. They are able to purchase guarantee from the government at fair price and sell it to S-banks at a potential premium.

5.4.4 Traditional banks' program

We now turn to T-banks' program, which we expose and solve again by backward induction.

As for S-banks, and for the clarity of exposition, we expose and solve, without loss of generality, the modified problem, of a T-bank with no access to capital goods market at $t = 1$ in state G, and whose only option is to roll its funding over.

Date 1, state G. As for S-banks, at $t = 1$ in state G, T-banks roll their funding over. With the same convention on value function in case of default, T-banks value function at $t = 1$ in state G writes:

$$V_{1,G}^{T}\left(I_0^T, D_0^T\right) = \begin{cases} \delta R I_0^T - D_0^T \text{ if } D_0^T \leq \delta R I_0^T \\ -\infty \text{ otherwise} \end{cases}$$

Date 1, state B In contrast to S-banks, T-banks are able to issue safe claims at $t = 1$ in state B because they can access a guarantee fund that makes these claims safe despite a non-zero probability of a zero output at $t = 2$. The possibility of a zero output at $t = 2$ therefore does not deter T-banks from issuing claims to households at $t = 1$ in state B, but T-banks are subject to two constraints in this respect. They have (i) to fairly pay for the guarantee on their short-term debt, while complying with their time 2 limited liability conditions and (ii) to comply with their debt constraint k.

Constraint (i) puts an upper bound on the amount of risk-less debt that can be reimbursed at $t = 2$ in states BB and BG (or similarly issued at $t = 1$ in state B): this amount cannot exceed the expected payoff of the productive investment at $t = 2$.

Indeed, the guarantee fund does not provide any subsidy to the T-bank: the T-bank must reimburse its date 1, B debt $D_{1,B}^T$ at $t = 2$, either by reimbursing the debt $D_{1,B}^T$ directly (at $t = 2$ in state BG), or by fairly financing its time 2, BB guarantee at $t = 2$ in state BG, hence providing the guarantee fund with $\frac{1-q}{q} D_{1,B}^T$ at $t = 2$ in state BG, such that the expected net payment made to the fund, from a time 1, B perspective is:

$$q \frac{1-q}{q} D_{1,B}^T + (1-q)(-D_{1,B}^T) = 0$$

For a T-bank, with a date 0 investment level I_0^T, who chooses to keep a share $\alpha_{1,B}^T$ on its balance sheet, and purchases $I_{1,B}^T$ units of capital goods at $t = 1$ in state B, the limited liability constraint at $t = 2$ in state BG rewrites (with $D_{1,B}^T$ the amount of riskless short term debt issued at $t = 1$ in state B) :

$$\frac{1-q}{q} D_{1,B}^T + D_{1,B}^T \leq \delta r \left(I_{1,B}^T + \alpha_{1,B}^T I_0^T\right)$$

Or equivalently

$$D_{1,B}^{T} \leq \delta qr\left(I_{1,B}^{T} + \alpha_{1,B}^{T} I_{0}^{T}\right)$$

Constraint (ii) writes as follows

$$D_{1,B}^{T} \leq k$$

In addition to these constraints, the bank is facing a limited liability constraint at $t = 1$ in state B and must reimburse its date 0 debt-holders and finance its date 1, T purchase either by raising new debt, or by selling part of its capital goods on the market. It also faces the same no short-sale constraint, and positivity constraints as the S-banks

In the no-default case, the value function of a T-bank at $t = 1$ in state B, writes

$$\begin{aligned} V_{1,B}^{T,ND}\left(I_{0}^{T}, D_{0}^{T}, p_{1,B}\right) = \max_{\left(\alpha_{1,B}^{T}, D_{1,B}^{T}, I_{1,B}^{T}\right) \in [0;1]\times[0;k]\times\mathbb{R}_{+}} & \left(\delta - p_{1,B}\right) qrI_{1,B}^{T} + \alpha_{1,B}^{T}\delta qrI_{0}^{T} + (1-\alpha_{1,B}^{T})p_{1,B}qrI_{0}^{T}\left(1-\varepsilon\right) - D_{0}^{T} \\ \text{s.t. } & (1-\alpha_{1,B}^{T})p_{1,B}qrI_{0}^{T}\left(1-\varepsilon\right) + D_{1,B} \geq D_{0}^{T} + p_{1,B}qrI_{1,B}^{T} \\ & D_{1,B} \leq q\delta\left(r\alpha_{1,B}^{T}I_{0}^{T} + rI_{1,B}^{T}\right) \end{aligned}$$

We keep the convention of setting $V_{1,B}^{T}\left(I_{0}^{T}, D_{0}^{T}, p_{1,B}\right) = -\infty$ if the bank defaults on its time 0 debt-holders.

As shown in the appendix, the T-bank does not default on its debt if and only if $D_{0}^{T} \leq \overline{D}_{0}^{T}\left(I_{0}^{T}, p_{1,B}\right)$, where $\overline{D}_{0}^{T}\left(I_{0}^{T}, p_{1,B}\right)$ is given in the appendix. The value function of a T-bank at $t = 1$ in state B is given in Proposition 2.

Proposition 2 *[Time 1, T-banks] For any $I_{0}^{T} \geq 0$, $D_{0}^{T} \geq 0$ and $p_{1,B} > 0$,*

$$V_{1,B}^{T}\left(I_{0}^{T}, D_{0}^{T}, p_{1,B}\right) = \begin{cases} V_{1,B}^{T,ND}\left(I_{0}^{T}, D_{0}^{T}, p_{1,B}\right) & \text{if } D_{0}^{T} \leq \overline{D}_{0}^{T}\left(I_{0}^{T}, p_{1,B}\right) \\ -\infty & \text{otherwise} \end{cases}$$

with

$$\begin{aligned} V_{1,B}^{T,ND}\left(I_{0}^{T}, D_{0}^{T}, p_{1,B}\right) = & \left(\frac{\left(\delta - p_{1,B}\right)_{+}}{p_{1,B}}\left(k - D_{0}^{T}\right)_{+} + \left(k - D_{0}^{T}\right)_{-}\frac{\left(\delta - p_{1,B}\left(1-\varepsilon\right)\right)_{+}}{p_{1,B}\left(1-\varepsilon\right)}\right) \\ & + \left(p_{1,B}\left(1-\varepsilon\right) - \delta\right)_{+} qrI_{0}^{T} + \delta qrI_{0}^{T} - D_{0}^{T} \end{aligned}$$

Moreover, if $p_{1,B} = 0$, $V_{1,B}^{T,ND}\left(I_0^T, D_0^T, p_{1,B}\right) = +\infty$

Proof. See appendix ☐

The intuition goes as follows. If the debt raised at $t = 0$ is higher than the maximum guarantee traditional banks can benefit on the time 1, B claims they issue $D_0^T > k$, the T-bank is not able to roll over all its funding. According to the price level, it will either choose to roll over as much of its previous debt as possible, hence k, (when the price is low), and sell just a high enough share of its assets, to cover for the remaining debt that must be reimbursed, or might choose not to roll over any debt, and sell capital goods on the market if the price gets high enough.

However, if the debt raised at $t = 0$ can be fully reimbursed by raising short term debt $t = 1$ in state B, the T-bank shall choose to roll all its funding over for low level of prices. Moreover, if prices are low enough, T-banks will choose to bind their k-constraint (set $D_{1,B}^T \leq k$) by raising additional debt such as to purchase under-priced capital goods. For higher level of prices, T-banks would choose not to invest, and only raise additional debt in order to reimburse its date 0 claim-holders, as well as potentially increase its time 1 consumption. Finally, if the price is high enough, T-banks might also be willing to sell all its capital goods on the market.

If the date 0 debt level is higher than the payoff that can be obtained with these strategies, the T-bank would not be able to reimburse all its date 0 debt-holders, while still complying with the different constraints it has to face. It would then only be able to reimburse less than the debt face value, which would lead to a default. In this case, the convention we chose is to put the value function to $-\infty$.

From a date 0 perspective, as T-banks can only finance themselves through risk-less short term debt, such a high level of debt cannot be chosen in equilibrium.

Date $t = 0$**.** At $t = 0$, T-banks choose how much funds to raise, and how much consumption goods to transform into capital goods. As for S-banks, T-banks have to choose a debt level such that no-default occurs. Moreover, the same funding constraint applies.

T-banks value function at $t = 0$ writes:

$$V_0^T\left(p_{1,B}, n^T\right) = \max_{\substack{(D_0^T, I_0^T) \in \mathbb{R}_+^2 \\ D_0^T \leq \overline{D}_{0,B}\left(I_0^T, p_{1,B}\right) \\ D_0^T \leq \delta R I_0^T \\ D_0^T + n^T \geq I_0^T}} p\left(\delta R I_0^T - D_0^T\right) + (1-p) V_{1,B}^{T,ND}(D_0^T, I_0^T, p_{1,B}) + (D_0^T + n^T - I_0^T)$$

Denoting $p_{1,L}^T \equiv \frac{\delta}{\delta qr + \frac{p(\delta R - 1)}{1-p}} < \delta$, and $p_{1,H}^T = \frac{p_{1,L}^T}{1-\varepsilon} < \frac{\delta}{1-\varepsilon}$, T-banks' program yields the following proposition.

Proposition 3 *[T-banks' program] Under condition (5.6), T-bank's program solves as follows*

1. *If* $0 < p_{1,B} < p_{1,L}^T$, $D_0^T = 0, I_0^T = n$, $V_0^T = p\left(\delta R n\right) + (1-p)\left(\delta qrn - k + \frac{\delta}{p_{1,B}} k\right)$
2. *If* $p_{1,B} = p_{1,L}^T$, *any* $D_0^T \in [0; k], I_0^T = n + D_0^T$ *is an equilibrium solution and* $V_0^T = p\left(\delta R(n+k) - k\right) + (1-p)\left(\delta qr(n+k) - k\right)$
3. *If* $p_{1,L}^T < p_{1,B} < p_{1,H}^T$, $D_0^T = k, I_0^T = k + n$ *and* $V_0^T = p\left(\delta R(n+k) - k\right) + (1-p)\left(\delta qr(n+k) - k\right)$
4. *If* $p_{1,B} = p_{1,H}^T$, *any* $D_0^T \in [k; \frac{k\left(1 - \frac{p_{1,B}(1-\varepsilon)}{\delta}\right) + (1-\varepsilon) p_{1,B} qrn}{1 - p_{1,B}(1-\varepsilon) qr}]$ *is an equilibrium solution,* $I_0^T = n + D_0^T$ *and* $V_0^T = p\left(\delta R(n+k) - k\right) + (1-p)\left(\delta qr(n+k) - k\right)$
5. *If* $p_{1,H}^T < p_{1,B} \leq \frac{\delta}{1-\varepsilon}$, $D_0^T = \frac{k\left(1 - \frac{p_{1,B}(1-\varepsilon)}{\delta}\right) + (1-\varepsilon) p_{1,B} qrn}{1 - p_{1,B}(1-\varepsilon) qr}$, $I_0^T = n + D_0^T$ *and* $V_0^T = p\left(\delta R\left(D_0 + n\right) - D_0\right)$
6. *If* $\frac{\delta}{1-\varepsilon} \leq p_{1,B} < \frac{1}{(1-\varepsilon)qr}$, $D_0^T = \frac{p_{1,B}(1-\varepsilon)qrn}{\left(1 - p_{1,B}(1-\varepsilon)qr\right)}$, $I_0^T = D_0^T + n$, *and* $V_0^T = p\left(\delta R\left(D_0 + n\right) - D_0\right)$
7. *If* $p_{1,B} \geq \frac{1}{(1-\varepsilon)qr}$, $D_0^T = +\infty$, $I_0^T = +\infty$ *and* $V_0^T = +\infty$

Moreover, if $p_{1,B} = 0$, $V_0^T = +\infty$

Depending on the price of capital goods on the secondary market at $t = 1$ in state B, T-banks choose how much short-term debt to issue at $t = 0$ to invest in positive NPV projects, versus how much buffer to keep to purchase capital goods from S-banks at $t = 1$ in state B. Although the guarantee fund enables T-banks to issue short-term debt at $t = 1$ in state B, they have to trade-off between those two investment opportunities because their issuance of short-term debt is limited by the size of the support.

For low level of prices, the return T-banks get on purchasing capital goods on the market overcomes the one of investing in the positive NPV projects they are faced with at $t = 0$. They prefer not issuing short term debt at $t = 0$, to keep slack in order to purchase goods at $t = 1$ in state B. When prices get higher, they are better off investing even more at $t = 0$, and selling part of his capital goods in the market when a crisis hits. As shall be seen shortly, this would not occur in equilibrium.

One implication which is worth emphasizing is that the size of the support the T-bank can benefit from the government at $t = 1$ in state B has an impact on its time 0 maximum debt level. Indeed, thanks to the guarantee fund, T-banks get the ability to raise risk-free debt at $t = 0$ up to an amount k under condition (5.6), without the need to delever at $t = 1$ in state B.[25] The T-bank is then able to issue up to k short term risk-less deposits whatever price prevails on the market at $t = 1$ in state B, generating a positive spillover from guarantee at $t = 1$ in state B to $t = 0$.

5.4.5 Banker's endowment allocation

Finally, given $V_0^i\left(p_{1,B}, n^i\right)$, bankers who choose to set up a i-bank ($i \in \{S;T\}$), allocate their initial endowment n between a part n^i they invest as bank's own funds, and the remaining part $(n - n^i)$ they consume.

When investing $n^i \in [0;n]$ units of their endowment into a i-bank, they obtain the residual payoffs of the bank. This provides them with an expected utility $V_0^i\left(p_{1,B}, n^i\right)$ from a date 0 perspective. Their problem writes as follows

$$V_0^{i,B}\left(p_{1,B}\right) = \max_{n^i \in [0;n]} \left(n - n^i\right) + V_0^i\left(p_{1,B}, n^i\right)$$

Then

[25] The bank has always the choice to roll its time 0 debt over in this case

1. If $p_{1,B} < \frac{1}{(1-\varepsilon)qr}$, they choose $n^i = n$. Indeed, each unit of funds a banker allocates to the i-bank can at least be transformed into capital goods and invested into the investment technology the bank has access to. The expected payoff generated by such a strategy is at least $\delta(pR + (1-p)qr)$, which provides more utility to the banker than immediate consumption of the funds (assumption 5.3)

2. If $p_{1,B} \geq \frac{1}{(1-\varepsilon)qr}$ the banker is indifferent between the different levels of endowment allocation.

In any case,

$$V_0^{i,B}(p_{1,B}) = V_0^i(p_{1,B}, n)$$

5.4.6 Capital goods market clearing

In this section, we derive the market-clearing conditions for the capital goods market at $t = 1$ in state B, taking the shares χ^S ($(1-\chi^S)$) of S-banks (T-banks) as given. Let us define an equilibrium on the secondary market for capital goods at $t = 1$ in state B.

Definition 2 *[Market equilibrium definition] A market equilibrium at $t = 1$ in state B is defined by*

1. *A quantity $S(p_{1,B})$ of capital goods supplied*

2. *A quantity $D(p_{1,B})$ of capital goods demanded*

3. *A price $p_{1,B}$ such that $D(p_{1,B}) = S(p_{1,B})$*

We have the following proposition

Proposition 4 *[Supply and demand] In date 1, B, with $\chi^S \in [0;1]$ denoting the share of S-banks and $(1-\chi^S)$ that of T-banks, the aggregate demand for capital goods writes:*

$$D(p_{1,B}) = \begin{cases} +\infty & \text{if } p_{1,B} = 0 \\ \frac{k}{p_{1,B}qr}(1-\chi^S) & \text{if } 0 < p_{1,B} < \overline{p}_{1,L}^T \\ \in \left[0; \frac{k}{p_{1,B}qr}(1-\chi^S)\right] & \text{if } p_{1,B} = \overline{p}_{1,L}^T \\ 0 & \text{if } p_{1,B} > \overline{p}_{1,L}^T \end{cases}$$

and the aggregate supply of capital goods writes:

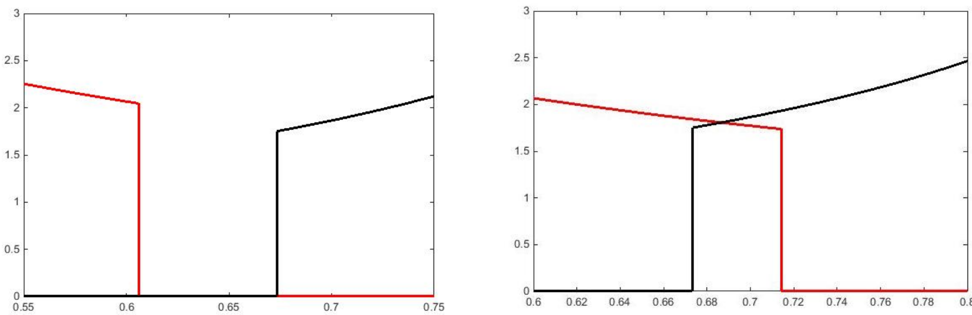
$$S(p_{1,B}) = \begin{cases} 0 & \text{if } 0 \leq p_{1,B} < \overline{p}_1^S \\ \in \left[0; \frac{n(1-\varepsilon)}{1-(1-\varepsilon)p_{1,B}qr}\chi^S\right] & \text{if } p_{1,B} = \overline{p}_1^S \\ \frac{n(1-\varepsilon)}{1-(1-\varepsilon)p_{1,B}qr}\chi^S & \text{if } \overline{p}_1^S < p_{1,B} < \overline{p}_{1,H}^T \\ \in \left[\frac{n(1-\varepsilon)}{1-(1-\varepsilon)p_{1,B}qr}\chi^S; \frac{n(1-\varepsilon)}{1-(1-\varepsilon)p_{1,B}qr}\chi^S + \frac{(1-\varepsilon)}{qr}\frac{\left(-\frac{k}{\delta}\right)+qr(n+k)}{1-p_{1,B}(1-\varepsilon)qr}\left(1-\chi^S\right)\right] & \text{if } p_{1,B} = \overline{p}_{1,H}^T \end{cases}$$

If $p_{1,B} > \overline{p}_{1,H}^T$, the aggregate supply of capital goods is strictly larger than the aggregate demand such that the market cannot clear.

Proof. See appendix □

The supply and demand for capital goods are illustrated in Figure 5.6 and follow from our previous analysis: when the price is low, there is a high demand for capital goods by T-banks. They prefer keeping slack at $t = 1$, to take the chance of buying underpriced assets at $t = 1$ in state B. When prices increase, they are better off issuing debt and investing at $t = 0$.

When the price is low shadow banks are not willing to lever themselves at $t = 0$, because the price they would get by selling assets on the capital goods market is too low, which makes the cost of time 0 debt too high for them, the intermediation price of government guarantee being too high. When prices increase, this reduces the cost of making their short-term debt safe, hence the cost of levering. This tradeoff is akin to the one emphasized in Stein (50).



(a) Case $\delta = 1$
No asset trade

(b) Case $\delta < 1$ ($\delta = 0.8$)
Asset trade

Figure 5.6: Numerical illustration of the capital goods market at $t = 1$ in state B (T-banks' demand in red, S-banks' supply in black)

One notices that $S(p_{1,B}) > 0$ when $p_{1,B} > \overline{p}^T_{1,H}$. Indeed, for such prices, both S and T-banks are willing to sell capital goods, which induces a positive supply whatever the allocation of intermediaries between the two types of banks. This restricts the set of prices that can prevail in equilibrium to $[0; \overline{p}^T_{1,H}]$.

Our technical assumption (5.7) ensures that $\overline{p}^S_1 < \overline{p}^T_{1,L} < \overline{p}^T_{1,H}$. According to the share of S-banks, different equilibria on the capital goods market can prevail.

Proposition 5 *[Market equilibrium] In equilibrium, we have*

1. *Either* $\chi^S = 0$, $D = S = 0$, *and* $\overline{p}^T_{1,L} \leq p_{1,B} \leq \overline{p}^T_{1,H}$. *No assets are traded.*

2. *Or* $\chi^S \in (0;1)$, $D = S$, *and*

   a) *Either* $\dfrac{\frac{k}{\overline{p}^S_1 qr}}{\frac{n(1-\varepsilon)}{1-(1-\varepsilon)\overline{p}^S_1 qr} + \frac{k}{\overline{p}^S_1 qr}} \leq \chi^S$ *and* $D = S = \frac{k}{\overline{p}^S_1 qr}(1 - \chi^S)$, *and* $p_{1,B} = \overline{p}^S_1$

   b) *Or* $\chi^S \in \left[\dfrac{\frac{k}{\overline{p}^T_{1,L} qr}}{\frac{k}{\overline{p}^T_{1,L} qr} + \frac{n(1-\varepsilon)}{1-(1-\varepsilon)\overline{p}^T_{1,L} qr}}; \dfrac{\frac{k}{\overline{p}^S_1 qr}}{\frac{n(1-\varepsilon)}{1-(1-\varepsilon)\overline{p}^S_1 qr} + \frac{k}{\overline{p}^S_1 qr}}\right]$ *and* $p_{1,B} = \dfrac{1}{\frac{\chi^S n}{k(1-\chi^S)} + 1} \dfrac{1}{qr(1-\varepsilon)} \in [\overline{p}^S_{1,}; \overline{p}^T_{1,L}]$,
   $D = S = \frac{k}{p_{1,B} qr}(1 - \chi^S) = \frac{n(1-\varepsilon)}{1-(1-\varepsilon)p_{1,B} qr}\chi^S$

   c) *Or* $\dfrac{\frac{k}{\overline{p}^T_{1,L} qr}}{\frac{k}{\overline{p}^T_{1,L} qr} + \frac{n(1-\varepsilon)}{1-(1-\varepsilon)\overline{p}^T_{1,L} qr}} \geq \chi^S$ *and* $D = S = \frac{n(1-\varepsilon)}{1-(1-\varepsilon)\overline{p}^T_{1,L} qr}\chi^S, p_{1,B} = \overline{p}^T_{1,L}$

3. *Or* $\chi^S = 1$, $D = S = 0$, *and* $0 \leq p_{1,B} \leq \overline{p}^S_1$. *No assets are traded*

5.4.7 The allocation program

We now endogenize bankers' choice to enter the T-or a S-banking sector by observing that bankers will choose whatever banking business is more profitable. They compare *ex ante* value functions as of $t = 0$, and choose an allocation χ^S such as to solve

$$\max_{\chi^S \in [0;1]} \chi^S V_0^{S,B}(p_{1,B}) + (1 - \chi^S) V_0^{T,B}(p_{1,B})$$

where $p_{1,B}$ is a market price for capital goods at $t = 1$ in state B.

Let's define:

$$\Delta : p_{1,B} \to V_0^{T,B}(p_{1,B}) - V_0^{S,B}(p_{1,B})$$

Recalling that equilibrium market prices for capital goods at $t = 1$ in state B are to be found in $[0; \overline{p}^T_{1,H}]$, it is sufficient to study Δ on this interval. On $[0; \overline{p}^T_{1,H}]$, Δ is a continuous, strictly decreasing function.

As, $\Delta(0) = +\infty$, it can cancel at most once on this interval. We end up with the following proposition

Proposition 6 *[Allocation program] Defining* $S = \Delta^{-1}(0) \cap [0; \overline{p}^T_{1,H}]$, *the allocation program solves as follows*

1. *If* $S = \emptyset$, $\chi^S = 0$

2. *Otherwise, denoting* $p_1^* = \Delta^{-1}(0) \cap [0; \overline{p}^T_{1,H}]$, *we have*

$$\chi^S = \begin{cases} 0 & \text{if } p_{1,B} \in [0; p_1^*) \\ \in [0;1] & \text{if } p_{1,B} = p_1^* \\ 1 & \text{if } p_{1,B} \in (p_1^*; \overline{p}^T_{1,H}] \end{cases}$$

5.4.8 The game equilibria

Having detailed the different parts of our equilibrium, we can now focus on the equilibrium determination of our game. We have the following result, which follows from the above propositions.

Proposition 7 *[Equilibria description] In all equilibria, bankers invest all their initial endowment in the bank they set up. The bank invests all the funds obtained from the bankers in the productive investment technology.At $t = 1$ in state G no assets are traded on the capital goods market, and banks roll over their time 0 short-term risk-free debt ($p_{1,G} \in [1; \frac{\delta}{1-\varepsilon}]$). In each date and state, risk-less short term debt (if any) is sold at par to the households, and the households purchase all the short term debt sold to them by the bankers.*

And:

1. *Either $\Delta(\overline{p}_1^S) < 0$. In this case, there is a unique $p^*_{1,B} \in (0; \overline{p}_1^S)$ such that $\Delta\left(p^*_{1,B}\right) = 0$. Then $\chi^S = 1$ is the unique equilibrium allocation, and any $p_{1,B} \in [p^*_{1,B}; \overline{p}_1^S]$ is an equilibrium market price at $t = 1$ in state B. No assets are traded in these equilibria, and only S banks exist. They don't issue any form of riskless debt.*

2. *Or $\Delta(\overline{p}_1^S) = 0$. In this case, any χ^S such that $\frac{\frac{k}{\overline{p}_1^S qr}}{\frac{n(1-\varepsilon)}{1-(1-\varepsilon)\overline{p}_1^S qr} + \frac{k}{\overline{p}_1^S qr}} \leq \chi^S$ is an equilibrium allocation, and $p_{1,B} = \overline{p}_1^S$. Either only S-banks exist, and they don't issue any type of riskless debt, and invest all their endowment in the productive investment, or shadow and traditional banks coexist, and interact on the asset market: T-banks raise funds at $t = 0$ to invest in the productive investment technology, and at $t = 1$ in state B, to purchase capital goods on the market.*

3. *Or $\Delta(\overline{p}_1^S) > 0$ and $\Delta(\overline{p}_1^T) < 0$. Then, there is a unique $p^*_{1,B} \in (\overline{p}_1^S; \overline{p}_1^T)$ such that $\Delta\left(p^*_{1,B}\right) = 0$. In this case $\chi^S = \frac{1}{1+\frac{p^*_{1,B} qr}{1-(1-\varepsilon)p^*_{1,B} qr}\frac{n(1-\varepsilon)}{k}}$ is the unique equilibrium allocation, and $p_{1,B} = p^*_{1,B}$ is the unique equilibrium market price at $t = 1$ in state B. Shadow and traditional banks coexist, and interact on the asset market at $t = 1$ in state B: T-banks don't issue short-term debt at $t = 0$, but issue short term debt at $t = 1$ in*

state B in order to purchase underpriced capital goods from S-banks. S-banks lever at $t = 0$, and sell capital goods on the market at $t = 1$ in state B.

4. *Or $\Delta(\overline{p}_1^T) = 0$. In this case, any χ^S such that $\frac{\frac{k}{\overline{p}_1^T qr}}{\frac{n(1-\varepsilon)}{1-(1-\varepsilon)\overline{p}_1^T qr} + \frac{k}{\overline{p}_1^T qr}} \geq \chi^S$ is an equilibrium allocation, and $p_{1,B} = \overline{p}_1^T$. Either only T-banks exist, and they issue k units of riskless debt at $t = 0$, or both S and T-banks exist, and T-banks issue less claims at $t = 0$, to keep slack to purchase capital goods at $t = 1$ in state B from S-banks. Shadow and traditional banks can coexist in this case, in which case assets are traded between the two types of intermediaries.*

5. *Or $\Delta(\overline{p}_1^T) > 0$. In this case $\chi^S = 0$ and any $p_{1,B} \in \left[\overline{p}_{1,L}^T ; \overline{p}_{1,H}^T\right]$ such that $\Delta(p_{1,B}) \geq 0$ is an equilibrium market price. No assets are traded in these equilibria. Only T banks exist, and they issue k units of riskless debt at $t = 0$, which they roll over at $t = 1$ in state B and in state G.*

The conclusion we draw from Proposition 7 is that both types of intermediaries can coexist and interact on capital goods market. When they do, S-banks lever themselves thanks to T-banks' ability to purchase capital goods sold by shadow banks in a crisis (i.e. at $t = 1$ in state B). In such a situation, fire-sales always occur.

Indeed, T-banks' have a limited ability to issue short-term debt, and thereby the total quantity of investment they can make at $t = 0$ and $t = 1$ in state B is limited. In order to purchase capital goods from S-banks in the crisis, T-banks need to be compensated for foregoing investment opportunities at $t = 0$. T-banks face a trade-off between issuing short-term debt at $t = 0$ and keeping some buffer in order to issue short-term debt at $t = 1$ in state B so as to purchase S-banks' capital goods. Interestingly this makes the fire sale prices always lower than the price at which the traditional banks value the asset (i.e. δ): the fire sale is not entirely driven by the need for T-banks to be compensated for higher functioning costs. However, the fire-sale price cannot be too low either: S-banks have to be be better off paying such a cost at $t = 1$ in state B to benefit from the intermediated government guarantee, in order to benefit from increased leverage in good times, rather than not levering up at all.

These trade-offs are key to understand the occurrence of fire-sales in a crisis and the interaction between traditional and shadow banks.

5.5 Discussion

5.5.1 The effects of assets transfer

Complementarity between T-banks and S-banks In the model, S-banks can issue riskless claims by relying upon T-banks' asset purchases in a crisis. T-banks are better off when there are more shadow banks because they have to pay a lower price for shadow banks' assets in a crisis. In that sense, our model exhibits a form of complementarity between T-banks and S-banks that goes through the asset market at $t = 1$ in state B (henceforth "in a crisis").

One implication of this complementarity is T-banks channel the support from the guarantee fund to the rest of the financial system. Indeed, T-banks have the possibility to intermediate this support by providing back-up to S-banks in times of troubles. In our setting, the more T-banks, the higher the (indirect) support from the guarantee fund to the financial system in a crisis. The higher this support, the better off the S-banks. Conversely, the more S-banks, the higher the amount of capital goods that needs to be absorbed by T-banks when S-banks need to delever. This induces a decrease in prices, as T-banks are limited in their ability to issue riskless debt in a crisis. This also increases T-bank's profit from asset purchases. T-banks are therefore better off when there are more S-banks.

The complementarity induced by the secondary asset market has consequences in terms of bankers' allocation between the two types of banks, and the aggregate level of investment in the economy (hence welfare).

Sectoral allocation effect and welfare impact To clarify this point, let us conduct the following thought experiment. Take our model and assume that there is no secondary asset market in a crisis. In such a case, there cannot be asset trade in a crisis, which means that neither T- nor S-banks can use the proceeds of asset sales to reimburse their creditors. A notable difference between T- and S-banks is that T-banks have access to the guarantee fund making their debt safe in a crisis, while shadow banks do not. As a result, it is impossible for shadow banks to issue debt at $t = 1$ in state B and - backwardly - neither at $t = 0$, because there is a nonzero probability that S-banks' output goes to zero at future date $t = 2$. In contrast, because they can access the guarantee fund, T-banks can issue risk-less debt in a crisis. From the viewpoint of $t = 0$, T-banks are able to issue risk-less debt insofar as they are able to refinance this debt at future dates, whatever state materializes at $t = 1$.

When there is no secondary asset market in a crisis, bankers' choice to enter the T- or S-banking sector becomes the following. One the one hand, T-banks enjoy deposit insurance in a crisis. Even though they do not benefit from fire-sales of S-banks' assets, this enables them to issue debt at $t = 0$. Again this comes at the cost of being regulated, which is captured by the parameter $\delta < 1$ in the model. On the other hand, S-banks do not have to comply with regulatory costs, and they do not enjoy deposit insurance in a crisis. Even without a secondary asset market, the absence of deposit insurance is detrimental to S-banks because they cannot issue debt at $t = 0$, thus no debt at all. Bankers therefore choose either to enter a levered though regulated T-banking sector, or an unregulated though unlevered S-banking sector.

Keeping the same notations as in the model, the time-0 value function of a banker entering the S-banking sector in an environment where there is no asset market in a crisis writes:

$$V_0^{SB,NM} = [pR + (1-p)qr]n \tag{5.8}$$

where we denote "NM" for "no market". Meanwhile, time-0 value function of a banker entering the T-banking sector when there is no asset market in a crisis writes:

$$V_0^{TB,NM} = p\delta[R(n+k) - k] + (1-p)(\delta qr(n+k) - k) \tag{5.9}$$

This has two implications. First, it should be noted that theses value functions are a lower bound for the value function of bankers entering the T- and S-banking sectors in the model with an asset market. This implies that opening up the asset market can only generate gains from trade, hence can only positively impact the welfare of the economy. One example of such a welfare improvement associated to the existence of this market is detailed in Figure 5.7.

Moreover, in a general way, the existence of such an asset market impacts the allocation of bankers in one type of bank or the other. In Figure 5.7, we provide an example where T-banking is dominated when markets are closed ($V_0^{TB,NM} < V_0^{SB,NM}$). Once we open up the asset market, bankers start allocating themselves in the T-bank sector. We interpret this as a potential rationale for why T-banks continue existing, even though their business model is dominated absent the opportunity to earn a profit from fire sales, the latter stemming from their ability to benefit from crises by issuing deposits when other intermediaries cannot.[26] The shaded area represents the welfare gains from trade on the asset market at $t = 1$ in state B.

[26]The allocation of bankers between the two forms of intermediation technology is illustrated in Figure 5.8, both with and without asset market, for different values of k.

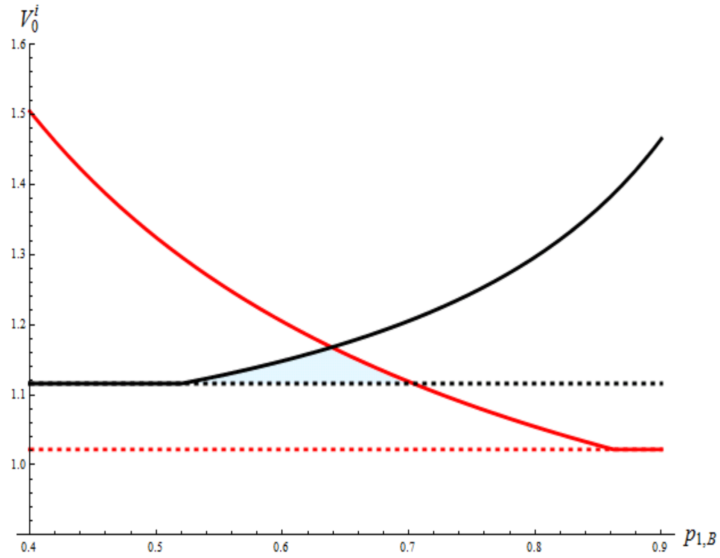


Figure 5.7: Value functions against asset price in a crisis (V_0^T in red, V_0^S in black, NM dashed)

5.5.2 Impact of changes in the level of guarantees (k) on the relative size of the traditional and shadow banking sectors

We now turn to the analysis of comparative statics of the model. In so doing, we focus on equilibria of type 3, as described in point 3 of proposition 7.

In an equilibrium of type 3, recall that bankers allocate themselves between T-banks and S-banks such that the value function at $t = 0$ of entering the T- or S-banking sector is the same (the solution is interior). We have $V_0^T = V_0^S$ with

$$V_0^T = \left[p\,(\delta R) + (1-p)\left(\delta q r + \frac{\delta - p_{1,B}^*}{p_{1,B}^*} \frac{k}{n} \right) \right] n$$

$$V_0^S = \left[\frac{p(R - (1-\varepsilon) p_{1,B}^* q r)}{1 - (1-\varepsilon) p_{1,B}^* q r} \right] n$$

where $p_{1,B}^*$ is the market equilibrium price for capital goods at $t = 1$ in state B.

Denoting $\chi_*^S \in [0,1]$ the equilibrium share of S-banks, market clearing condition in a crisis yields:

$$p_{1,B}^* = \frac{\frac{k}{n}}{(1-\varepsilon)\, qr\left(\frac{\chi_*^S}{(1-\chi_*^S)} + \frac{k}{n}\right)}$$

Recall that for parameter values such that this equilibrium exists, it is uniquely defined. Combining these three conditions we obtain

$$V_0^S = n\frac{p(R-1)}{\frac{\chi_*^S}{(1-\chi_*^S)}}\left(\frac{\chi_*^S}{\left(1-\chi_*^S\right)} + \frac{k}{n}\right) + np \tag{5.10}$$

$$V_0^T = np\delta R + (1-p)\left(\delta qrn + \left(\delta\,(1-\varepsilon)\,qr - 1\right)k + \delta\,(1-\varepsilon)\,qr\frac{\chi_*^S}{\left(1-\chi_*^S\right)}n\right) \tag{5.11}$$

which enables us to link the equilibrium sectoral allocation to the value function of T-banks and S-banks.

We use expressions (5.10) and (5.11) to discuss how the relative size of the two banking sectors is modified when allowing for a slight change in the size of the government guarantee.

Change in the amount of deposit insurance k For clarity we discuss the effects of lowering k. We first focus on a change in k low enough to ensure that the considered modified equilibrium stays of type 3. Two competing effects are at stake.

1. On the one hand, keeping all parameters constant, a lower k reduces the size of the advantage of being a T-bank: T-banks have access to a lower government guarantee, which limits their leverage at all times, hence exerting a downward pressure on T-bank's profits. This makes T-banking relatively less profitable than S-banking, thereby inducing bankers to enter the S-banking sector. We call this direct effect a substitution effect.[27]

[27] It is reminiscent of the argument at play in existing models of shadow banking as regulatory arbitrage (see e.g. Plantin (42), Ordonez (41), Harris et al. (29)).

2. On the other hand, S-banks indirectly benefit from T-banks' guarantee fund through the (secondary) asset market in a crisis: this is at the root of the complementarity effect we pointed out in the previous subsection. Lowering k reduces the support that can be provided to S-banks through T-banks, reducing the price at which capital goods can be sold on the secondary market, hence reducing S-banks' profits and the incentive of entering the S-banking sector. We call this effect an income effect.

In our setting, the asset price adjustment associated to a lower asset demand by T-banks in a crisis outweighs the direct effect of lowering k on T-banks' profits: the net effect of lowering k on T-banks' expected profits is positive taking bankers' allocation as given.[28] We find that lower support to T-banks in a crisis reduces asset prices to such an extent that more bankers choose to enter the T-banking sector *ex-ante*, i.e. the income effect outweighs the substitution one. Bankers therefore react by allocating themselves with a higher probability in the traditional banking sector.

However, this does not hold true for high changes in k: holding δ fixed, when k gets low enough, T-banking becomes too inefficient for T-banks to continue existing. Indeed, for low levels of k, the T-banks would require too high a compensation for purchasing the asset at $t = 1$ in state B that S-banks would be better off not issuing short term debt. Moreover, in this range, entering the unlevered shadow banking sector is more valuable than entering the traditional banking sector where banks issue k units of risk-less claims at $t = 0$ to invest in the productive technology. Finally, bankers allocate themselves only in S-banking, and no short-term debt is issued.

Figure 5.8 illustrates the allocation of bankers into the shadow banking system when the parameters are such that S-banking is dominating absent a market for assets in a crisis (i.e., at $t = 1$ in state B). We take $n = 1, p = 0.9, r = 1, q = 0.99, R = 1.13, \varepsilon = 0.11, \delta = 0.9$, for $0 < k < k^*$. The dashed red line illustrates the fact that all bankers allocate to the S-banking sector if T-banks cannot purchase capital goods from S-banks in the secondary market. The plain black line illustrates the two effects at play for different values of

[28] This is evidenced by the fact that, in equation (5.10), the coefficient of k, $(1-p)(\delta(1-\varepsilon)qr-1)\frac{1}{n}$ is negative.

$0 < k < k^*$.[29] When k is low, T-banks' advantage of issuing debt in a crisis do not compensate for their costs, and all bankers allocate to the S-banking sector ($\chi^S = 1$). Once T-banks' advantage enables them to purchase assets from S-banks at fire-sale prices, not all bankers allocate to the S-banking sector and T- and S-banks coexist. By the same reasoning as previously, increases in the amount of deposits guaranteed in a crisis pushes more bankers to allocate to the S-banking sector.

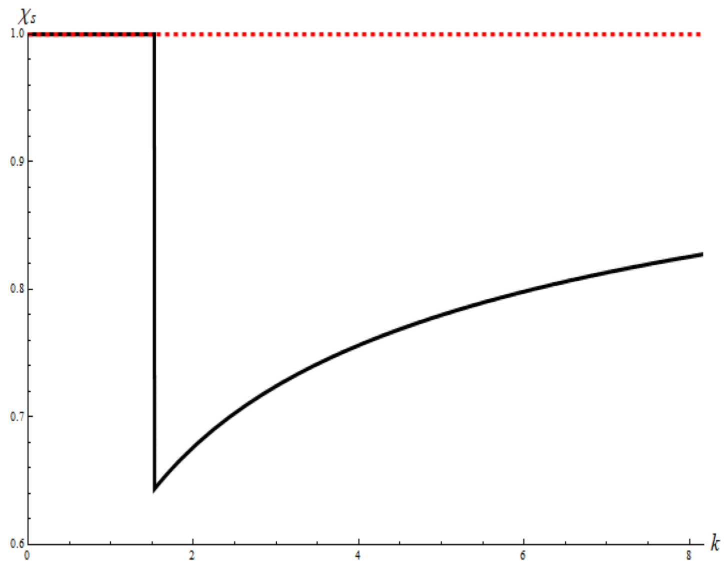


Figure 5.8: Share of shadow banks according to the level of guarantee (plain with asset market, dashed without asset market)

Analogy with capital requirements A positive analogy that can be developed is to consider a model where the structural limit on the size of T-banks' support in a crisis (k) comes at the cost of complying with capital requirements, that are imposed to T-banks for reasons outside of the model.

We assume that at $t = 1$ in state B, T-banks face a capital requirement of the form:

$$D^T_{1,B} \leq (1 - c)\left[\alpha I^T_0 + p_{1,B} q r I^T_1\right] \tag{5.12}$$

[29] Notice that, for these values of the parameters, coexistence of the different forms of intermediation is ensured only when asset markets are open.

where $D^T_{1,B}$ denotes the amount of funds borrowed from households and $p_{1,B}$ the price of capital goods.[30] Such a model would yield similar results, where our k counterpart would be found in $\frac{(1-c)n}{c}$. As is the case in our model, imposing such a capital requirement on desposits issued in a crisis spills over to date 0, as T-banks needs to roll all their debt over at $t = 1$ in state B, when no other intermediary is able to purchase assets. The same argument as in our model can then be adapted to this case, where increased capital requirement in a crisis (thereby at all dates and states) would reduce T-bank's ability to purchase assets in a crisis, making T-banks better off and reducing bankers' incentives to set up a S-bank *ex-ante*. Conversely, lower capital requirement in a crisis would reduce the impact of fire sales: In that way, policies such as countercyclical capital buffers could have the potentially unwanted consequences of increasing the support to the shadow banking system at the time when shadow banks need it the most, hence driving more intermediaries into setting up unregulated entities.

Post-crisis banking reforms There have been several policy initiatives to impose restrictions on banks' trading activities since the crisis. Prohibiting regulatory arbitrage is the paradigm in Section 619 of the Dodd-Frank Act in the U.S. (known as the "Volcker Rule"), in the Financial Services Act of 2013 in the U.K. (based on the Report of the Vickers Commission), as well as the 2012 Report of the European Commission's High-level Expert Group on Bank Structural Reform in the E.U. (known as the Liikanen Report). Those regulation proposals include a prohibition of proprietary trading by commercial banks ("Volcker Rule"), a separation between different risky activities (Liikanen Report), and ring-fencing of depository institutions and systemic activities (Report of the Vickers Commission, enacted in 2013 in the Financial Services Act).

The philosophy of these reforms is to prevent traditional banks from doing regulatory arbitrage by sustaining shadow banking activities, which Pozsar et al. (43) refer to as "in-

[30] At $t = 1$ in state B, we would set the following constraint, for the capital constraint to be stringent enough (rhs) while not incitating T-banks to delever at $t = 1$ in state B (lhs):

$$\delta(1-\varepsilon)qr \leq (1-c) \leq \delta qr \tag{5.13}$$

ternal shadow banking". Through the lens of our model, we indeed acknowledge the fact that contractual relationships on liquidity guarantees from traditional to shadow banks tend to favor shadow banking by increasing asset prices in a crisis. However, our key finding is that, even absent contractual relationships between traditional and shadow banks, the two types of banks coexist because they are not only substitutes but also complements. The complementarity between the two bank types comes from shadow banks' asset (fire) sales to traditional banks in a crisis. One implication of this complementarity is that traditional banks channel the support from the guarantee fund to shadow banks, even absent a contractual relationship between the two. Although it might then seem like a good idea to prevent traditional banks from trading assets in a crisis, this paper argues that traditional banks' profits from shadow banks' fire-sold assets in a crisis outweigh the (regulatory) costs that traditional banks have to comply with. This explains why traditional banks still exist even though shadow banks perform comparable activities while not having to comply with regulations. When designing banking reforms one needs to think hard about the implications of reforms such as a shutdown of the (interbank) asset market, in light of the reasons why regulated and unregulated banks coexist in the first place. This paper provides a framework to do so.

5.5.3 Normative approach

In addition to the positive aspects developped above, this framework provides a setting for a normative analysis. To do so, we will look at the aggregate surplus in the economy, hence the aggregate profit of the traditional and shadow banking systems. A way to think about it is to consider that bankers don't consume the profits they make, but that they provide the households with it in the end of period 2. This point of view is close to the normative measure put forward by Stein (50).

Among the questions that can be adressed, we focus here on the analysis of the efficiency of the banker's allocation between the two types of banking institutions, with respect to the measure above. We will be assuming that having fixed the allocations of bankers between the two forms of intermediation technology, they are free to make the investment decisions that occur in the decentralized market.

In this model, the advantage of shadow banks is that they don't have to bear the cost of regulation δ (and the cost of early investment disruption ε), while the advantage of traditional banks is that they can get support from the government in times of crises k. These two distinctions might lead to some form of inefficiencies, associated to pecuniary externalities linked to a misallocation of agents.

Let's now consider the allocation problem faced by the central planner. The problem of the central planner writes as the allocation program of the agents. However, he recognizes the impact of its allocation choice on the equilibrium price $p_{1,B}$ which clears the secondary market in time 1, B. The planner chooses an allocation $\chi^S \in [0;1]$ such as to solve the central planner's program:

$$\max_{\chi^S \in [0;1]} \chi^S V_0^{S,B}(p_{1,B}) + (1-\chi^S) V_0^{T,B}(p_{1,B})$$

where $p_{1,B}$ is a market price for capital goods at $t=1$ in state B. χ_S being fixed, shadow and traditional banks make the same choices as in the decentralized equilibrium and $p_{1,B}$ is expressed as a functionf of χ_S in the same way as in Proposition 5.

The objective function of the central planner can then be rewritten as a piecewise linear function of χ_S:

$$W(\chi_S) = \begin{cases} \chi_S V_0^{S,B}(\overline{p}_{1,L}^T) + (1-\chi_S) V_0^{T,B}(\overline{p}_{1,L}^T) & if \chi_S \in [0; \underline{\chi_S}] \\ \alpha_0 + \alpha_1 \chi_S & if \chi_S \in [\underline{\chi_S}; \overline{\chi_S}] \\ \chi_S V_0^{S,B}(\overline{p}_1^S) + (1-\chi_S) V_0^{T,B}(\overline{p}_1^S) & if \chi_S \in [\overline{\chi_S}; 1] \end{cases}$$

with $\alpha_0 = p(R-1)k + np\delta R + (1-p)[\delta qrn + (\delta(1-\varepsilon)qr - 1)k]$, $\alpha_1 = n[pR - \delta[pR + (1-p)\varepsilon qr]] + k[1 - [pR + (1-p)\delta(1-\varepsilon)qr]]$, $\underline{\chi_S} = \frac{\frac{k}{\overline{p}_1^S qr}}{\frac{n(1-\varepsilon)}{1-(1-\varepsilon)\overline{p}_1^S qr} + \frac{k}{\overline{p}_1^S qr}}$, $\overline{\chi_S} = \frac{\frac{k}{\overline{p}_{1,L}^T qr}}{\frac{k}{\overline{p}_{1,L}^T qr} + \frac{n(1-\varepsilon)}{1-(1-\varepsilon)\overline{p}_{1,L}^T qr}}$.

For the following, we restrict our attention on parametric restrictions where the decentralized equilibrium is such that both traditional and shadow banks coexist and the market price is such that $p_{1,B}^D \in (p_1^S; p_{1,L}^T)$. This corresponds to an equilibrium of type 3 in the typology of Proposition 7. In this case, the banker's allocate themselves with a uniquely defined probability $\chi_S^D \in (\underline{\chi_S}; \overline{\chi_S})$ into the shadow banking system.

Such parameter restrictions first ensure that W(.) is strictly increasing on $[0; \underline{\chi_S}]$ and strictly decreasing on $[\overline{\chi_S}; 1]$. Moreover, from $V_0^{T,B}(\overline{p}_1^S) > V_0^{S,B}(\overline{p}_1^S)$, we can infer that $\alpha_1 < 0$ [31] In turn, W(.) is strictly decreasing on $[\underline{\chi_S}; \overline{\chi_S}]$ and the constrained optimum allocation is uniquely obtained for $\chi_S^C = \underline{\chi_S} < \chi_S^D$. The market allocation of bankers into the shadow banking sector is in this case always excessive. We illustrate this situation in figure 5.9, where the parameter values used in figure 5.7 have been chosen.

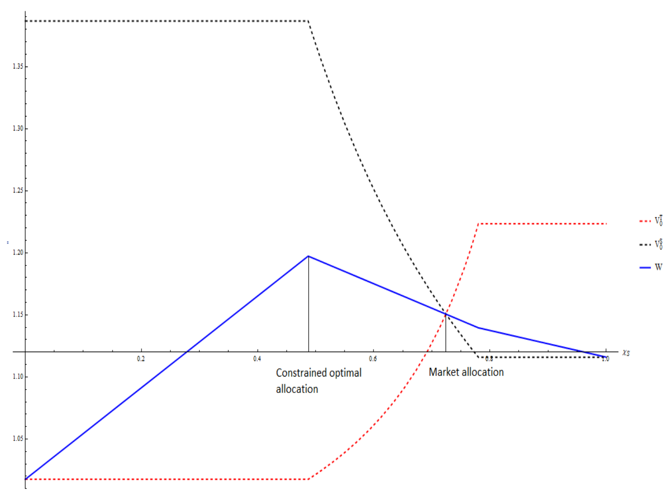


Figure 5.9: Centralized and decentralized equilibrium allocation

With these parameters' sets, bankers allocate themselves too much into the shadow banking system, or conversely too little into the traditional banking system. The key mechanism at play in this inefficiency is reminiscent of Stein (50), and stems from pecuniary externalities; bankers, by choosing to allocate themselves into the shadow banking system fail to internalize that this will lower the support brought to shadow banks in bad times. This in turns hinders the ability of traditional banks to provide a backstop in times of crises, hence the ability of all shadow banks in growing and investing in normal times;

[31]Indeed, $V_0^{T,B}(\overline{p}_1^S) > V_0^{S,B}(\overline{p}_1^S)$ rewrites $[pR + (1-p)qr][1-\delta]n < k[(1-\varepsilon)\delta[p(R-1) + (1-p)qr] - (1-p)]$ which implies $n[pR(1-\delta) - \delta(1-p)\varepsilon qr] < k[pR + (1-p)\delta(1-\varepsilon)qr - 1]$ or alternatively $\alpha_1 < 0$.

in a nutshell, when bankers allocate themselves into the shadow banking system, they fail to take into account that this choice reduces the ability of traditional banks to purchase assets in bad times, hence the ability of all shadow banks to issue riskless claims in time 0.

In this case, transfers between the two forms of intermediation technology can be beneficial in terms of total Welfare. The aim is here to incentivize more agents to allocate themselves towards the traditional banking system, by subsidizing traditional banking activities and financing this subsidy through taxes levied on shadow banks. Welfare can always be improved by imposing lump sum taxes on shadow banks, in order to subsidize traditional banking activities.

It should be noted that the market allocation needs not be always inefficient. For instance, if the decentralized allocation is one in such shadow banks are the only form of intermediaries existing, decentralized and centralized allocations will coincide, as exemplified in the following figure.

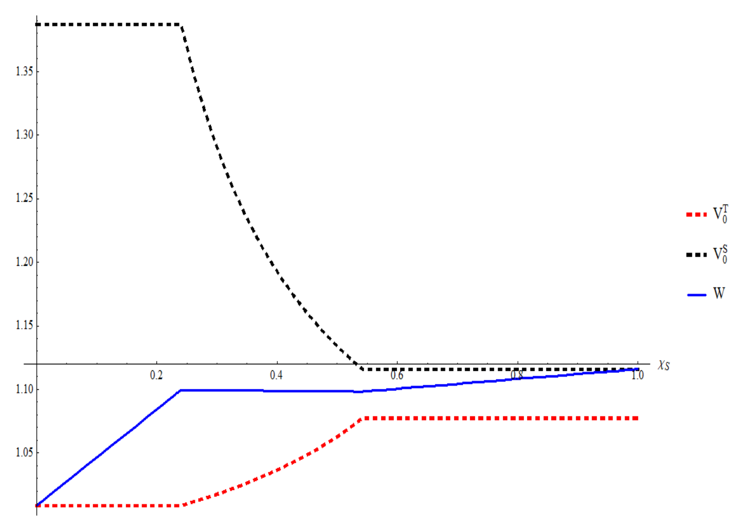


Figure 5.10: Constrained optimal allocation

5.6 Conclusion

We document and integrate stylized facts from the 2007 financial crisis into a simple model that rationalizes the coexistence of traditional and shadow banks. This paper offers the first model of financial intermediation where both a regulated and an unregulated sector coexist and interact, while replicating the following facts from the crisis: (i) liabilities transfer from shadow to traditional banks, (ii) assets transfer from shadow to traditional banks, and (iii) fire sales of assets.

The model describes the different technologies used by traditional and shadow banks in order to issue safe claims against risky collateral. On the one hand, traditional banks rely on a guarantee fund to be able to issue riskless claims in a crisis. Therefore this guarantee also enables them to issue riskless claims outside a crisis. This access to a guarantee fund comes at a cost of higher regulation. On the other hand, shadow banks rely upon traditional banks' ability to issue riskless claims in a crisis, to absorb their assets and provide them with enough liquidity to reimburse their debt holders. This interaction in the asset market enables traditional banks to intermediate government insurance to the rest of the financial system, generating a complementarity between the two forms of financial intermediation. This complementarity is the key message of this paper: the more shadow banks in the system, the lower the price traditional banks have to pay for shadow banks assets in a crisis, and the better off the traditional banks. We see this form of complementarity as a main driver of the coexistence of these two banking sectors in the financial system.

We find that when shadow and traditional banks coexist in an economy, a small reduction in traditional banks' ability to issue deposits in a crisis induces a shift of intermediation towards the traditional banking sector. Indeed, two opposite effects are at play. One is the direct effect of hindered traditional banks' ability to raise funds in a crisis, which is to reduce their ability to lever up and thereby their expected profits. This substitution effect induces bankers to shift to the shadow banking sector, as this latter is not directly impacted by such a reduction in traditional banks ability to issue risk-less debt. The other one is indirect and akin to an income effect: a reduction in traditional banks' ability to issue riskless claims induces a lower asset demand in the crisis, creating a downward pressure on the equilibrium price of assets transferred from the shadow to the traditional banking sector. This lowers shadow banks' expected profits and increases traditional banks' expected profit, thereby pushing more bankers into the regulated sector. We show that this latter effect is sizeable and outweigh the former. All in all, a reduction in traditional banks' ability to lever up in a crisis, if not too strong, leads to an increase in the relative size of the traditional banking sector.

From a normative perspective, we show that, when traditional and shadow banks coexist, the pecuniary externality stemming from the fire sales of shadow banks assets in times of crises usually induces inefficiencies: bankers tend to allocate too much into the shadow banking system, failing to internalize the fact that their allocation towards the shadow banking system reduces the overall support that can be brought to the shadow banking system in times of crises. Lump sum transfers between shadow and traditional banks in an attempt to increase the number of traditional banks have an overall positive impact on the welfare of the economy.

Technical Appendix

5.A The data

5.A.1 Stylized balance sheets of US financial intermediaries

Once we aggregate those financial intermediaries that we include in our definition of the shadow banking sector, we define short-term debt using the FAUS by using the same terminology as Krishnamurthy and Vissing-Jorgensen (2015). We build stylized balance-sheets by consolidation of the financial balance sheets of the legal institutions for which we have data in the Financial Accounts of the United States (FAUS).

The list of items included in shadow banks' short-term debt is: Security repurchase agreements (net), Depository institution loans n.e.c., Trade payables, Security credit (Customer credit balances), Security credit (U.S.-chartered institutions), Security credit (foreign banking offices in U.S.), Taxes payable, Commercial paper, Open market paper.

5.A.2 Book versus market value of equity

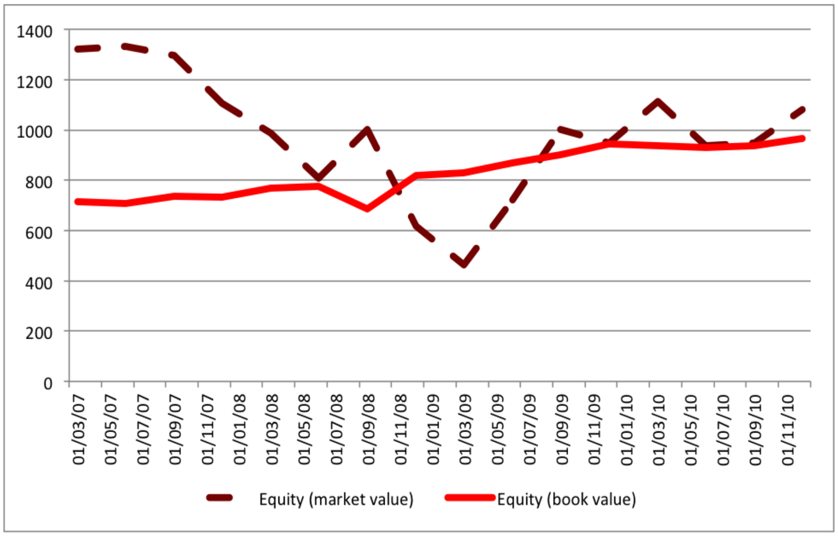


Figure 5.A.1: Traditional banks: Book versus market value of equity

He et aL.(2010) and Begenau et al. (2016) also find that traditional banks' book equity increased by around US $250 billion during the crisis. Figure 5.A.1 provides evidence of this increase in the stock of book equity of the US traditional banking sector through the crisis. This Figure is based on reported book value of equity, which is the leverage measure most used for regulatory purposes. However, there are reasons to believe that the true level of capital for the traditional banking sector was lower. We use data from CRSP to measure the market value of traditional banks' equity and we see that most of the increase in book value of equity disappears when one looks at market value of equity.

5.A.3 Deposit inflows into traditional banks

We take the definition of the largest US bank-holding companies on Figure 5.1 from the Federal Reserve's website (https://www.ffiec.gov/nicpubweb/nicweb/HCSGreaterThan10B.aspx/).

5.A.4 Regression: Traditional banks' MBS purchases in the crisis

Table 5.A.1: Traditional banks: Determinants of MBS purchases during the crisis

	(1) Scenario 1	(2) Scenario 2	(3)
Change_insured_deposits	0.231***	0.227***	0.223***
	(0.049)	(0.044)	(0.039)
Change_brokered_deposits	0.016	0.031	0.046
	(0.046)	(0.042)	(0.038)
Change_ir_large_deposits	−0.777***	−0.465**	−0.160
	(0.233)	(0.188)	(0.162)
Change_Credit	−0.112***	−0.102***	−0.092***
	(0.032)	(0.028)	(0.024)
Change_Capital_ratio	−0.695***	−0.583***	−0.474***
	(0.117)	(0.109)	(0.105)
Change_Net_Wholesale_fund	0.119***	0.116***	0.113***
	(0.037)	(0.034)	(0.032)
Change_Real_estate	−0.016	−0.034	−0.052
	(0.049)	(0.041)	(0.037)
Change_NPL_Ratio	−0.363***	−0.287***	−0.213***
	(0.078)	(0.066)	(0.059)
Change_Liquidity	−0.374***	−0.332***	−0.292***
	(0.044)	(0.037)	(0.032)
Liquidity_ratio_2007q4	−0.019	−0.001	0.017
	(0.027)	(0.023)	(0.020)
NPL_ratio_2007q4	−0.303***	−0.290***	−0.277***
	(0.114)	(0.093)	(0.080)
Mat_Gap_2007q4	−0.119***	−0.079***	−0.040***
	(0.014)	(0.012)	(0.011)
ir_large_deposits_2007q4	−0.979***	−0.613***	−0.255
	(0.226)	(0.181)	(0.156)
Net_Wholesale_fund_ratio_2007q4	0.025	0.016	0.008
	(0.021)	(0.017)	(0.015)
Unused_commitments_ratio_2007q4	−0.004	−0.024	−0.043*
	(0.033)	(0.028)	(0.024)
Real_Estate_2007q4	−0.055***	−0.045***	−0.034***
	(0.016)	(0.013)	(0.011)
Capital_ratio_2007q4	−0.058	−0.046	−0.035
	(0.059)	(0.051)	(0.048)
Tag_deposits	0.198***	0.153***	0.109***
	(0.048)	(0.042)	(0.039)
log_assets	0.007***	0.004***	0.000
	(0.002)	(0.001)	(0.001)
Constant	0.081***	0.074***	0.068***
	(0.031)	(0.025)	(0.021)
Observations	3954	3954	3954
adjusted R^2	0.1989	0.2229	0.2374

The data come from the quarterly Call Reports and He et aL.(2010)'s estimates. We use the procedure described in Acharya and Mora (2015) to construct our sample. Variables ending in 2007q4 represent variable levels as of 2007q4. Variables starting with "Change" are growth rates from 2007q4 to 2009q1, normalized by total assets as of 2007q4. The dependent variable MBS_Purchases represents purchases of mortgage-backed securities by traditional banks between 2007q4 and 2009q1, normalized by total assets as of 2007q4. As in He et aL.(2010) we test many scenarios in terms of MBS repayment rate and total losses on assets, to make sure that what our dependent variables capture are actual purchases of MBS by traditional banks. We report three of these scenarios, including the "naive" one. Under scenario 1 the repayment rate used to construct the MBS_Purchases variable is 7% and total losses imputed to the financial sector are \$500 billion. Under scenario 2, the repayment rate is 12% and total losses are \$176 billion. Under the "naive" scenario, we do not correct for the net repayment rate nor total losses. The White robust standard error estimator is used. Stars indicate p values of 10% (one star), 5% (two stars) and 1% (three stars). Table 5.A.2 below details the construction of variables mainly following Acharya and Mora (2015).

Variable	Variable Name	Call Report Items
Insured deposits	insured__deposits	rconf049 +rconf045
Brokered deposits	brokered__deposits	rcon2365
Interest rate on large deposits	ir__large__deposits	rconf049 + rconf045
Unused commitment		rcfd3814+rcfd3816+rcfd3817+rcfd3818+rcfd6550+rcfd3411
Credit	Credit	rcfd1400 + Unused__commitments
Unused commitments ratio	Unused__commitments__ratio	unused commitments/(unused commitments+rcfd1400)
Cash		rcfd0010
Federal Funds Sold		rcfd1350+rconb987
		(rconb987+rcfdb989 if after 2002/03/30)
MBS		rcfd1699+rcfd1705+rcfd1710+rcfd1715+rcfd1719+rcfd1734
		+rcfd1702+rcfd1707+rcfd1713+rcfd1717+rcfd1732+rcfd1736
Securities (MBS excluded)		rcfd1754+rcfd1773-(rcfd8500+rcfd8504+rcfdc026+rcfd8503+rcfd8507+rcfdc027)
Liquid assets		Securities (MBS excluded)+ Federal Funds Sold+Cash
Liquidity ratio	Liquidity__ratio	Liquid Assets/rcfd2170
Wholesale funding		rcon2604+rcfn2200+rcfd3200+rconb993+rcfdb995+rcfd3190
Wholesale funding ratio		Wholesale funding/rcfd2170
Net Wholesale fund ratio	Net__Wholesale__fund	Wholesale funding -(Securities (MBS excluded)+Federal Funds Sold+Cash)
Non performing loan		rcfd1407+rcfd1403
Non performing loan ratio	NPL__ratio	Non performing loan/rcfd1400
Capital ratio	Capital__ratio	(rcfd3210+rcfd3838)/rcfd2170
Real Estate Loan Share	Real__Estate	rcfd1410/rcfd1400
Residential Mortgages		(rcfdf070+rcfdf071)/rcfd2170
Financial Assets		rcfd0081+rcfd0071+rcfda570+rcfda571+rcona564+rcona565
		+rcfd1350+rcfda549+rcfda550+rcfda556+rcfda248
Short Term Liabilities		rcon2210+rcona579+rcona580+rcona584+rcona585+rcfd2800+rcfd2651+rcfdb571
Maturity Gap	Mat__Gap	(Financial Assets- Short Term Liabilities) / rcfd2170
Tag deposits	Tag__deposits	rcong167

Figure 5.A.2: Variables definitions

Note: All missing observations are considered equal to zero. Banks are aggregated to top holder level (RSSD9348). We follow the same procedure as Acharya and Mora (2015). We now turn to the technical appendix related to the model.

5.B Shadow bank's program

5.B.1 Proof of lemma 1

No default occurs at $t = 1$ in state B if and only if the S-bank is able to obtain enough funds when selling assets, to finance its debt level D_0^S .

For any $t = 0$ investment level $I_0^S > 0$, $t = 1$-state B capital good price $p_{1,B}qr > 0$, there is an upper level $\overline{D}_0^S \left(I_0^S, p_{1,B}\right)$ of debt that can be reimbursed at $t = 1$ in state B:

$$\begin{aligned}\overline{D}_0^S\left(I_0^S\right) &= \max_{\alpha_S^1 \in [0;1]} \left(\left(1-\alpha_1^S\right)(1-\varepsilon)\, p_{1,B} qr I_0^S\right) \\ &= (1-\varepsilon)\, p_{1,B} qr I_0^S\end{aligned}$$

where $\left(1-\alpha_1^S\right)$ is the share of S-bank's assets that is liquidated. at $t = 1$ in state B. If $D_0^S > (1-\varepsilon)\, p_{1,B} qr I_{0,S}$, the S-bank must default on its debt issued at $t = 0$. In case of default, we set $V_{1,B}^{S,D}\left(I_0^S, D_0^S, p_{1,B}\right) = -\infty$. This ensures that the S-bank is not willing to default on its debt at $t = 1$ in state B.

In case of no-default, the program writes:

$$\begin{aligned}V_{1,B}^{S,ND}\left(I_0^S, D_0^S, p_{1,B}\right) = \max_{\alpha_{1,B}^S \in [0;1]} \ & \alpha_{1,B}^S qr I_0^S + (1-\alpha_{1,B}^S)(1-\varepsilon) p_{1,B} qr I_0^S - D_0^S \\ \text{s.t. } & \left(1-\alpha_{1,B}^S\right)(1-\varepsilon)\, p_{1,B} qr I_0^S \geq D_0^S\end{aligned}$$

Denoting $\nu_{\alpha_{1,B}^S \geq 0}$ the Lagrange multiplier associated to the constraint $\alpha_{1,B}^S \geq 0$, $\nu_{\alpha_{1,B}^S \leq 1}$ the Lagrange multiplier associated to the constraint $\alpha_{1,B}^S \leq 1$ and $\mu_{1,B}^S$ the Lagrange multiplier associated to the funding constraint, the Lagrangian of the problem rewrites:

$$\begin{aligned}L = {} & \alpha_{1,B}^S qr I_0^S + (1-\alpha_{1,B}^S)(1-\varepsilon) p_{1,B} qr I_0^S - D_0^S \\ & + \nu_{\alpha_{1,B}^S \geq 0} \alpha_{1,B}^S + \nu_{\alpha_{1,B}^S \leq 1}\left(1-\alpha_{1,B}^S\right) \\ & + \mu_{1,B}^S \left(\left(1-\alpha_{1,B}^S\right)(1-\varepsilon)\, p_{1,B} qr I_0^S - D_0^S\right)\end{aligned}$$

and the first order condition writes as follows

$$\frac{dL}{d\alpha_{1,B}^S} = qr I_0^S \left(1-(1-\varepsilon) p_{1,B}\right) + \nu_{\alpha_{1,B}^S \geq 0} - \nu_{\alpha_{1,B}^S \leq 1} - \mu_{1,B}^S (1-\varepsilon)\, p_{1,B} qr I_0^S = 0$$

And solves as follows:

1. If $1-(1-\varepsilon)p_{1,B}<0$, $\nu_{\alpha^S_{1,B}\geq 0}>0$ and $\alpha^S_{1,B}=0$, $V^S_{1,B}\left(I^S_0,D^S_0,p_{1,B}\right)=(1-\varepsilon)p_{1,B}qrI^S_0-D^S_0$

2. If $1-(1-\varepsilon)p_{1,B}=0$, $\nu_{\alpha^S_{1,B}\geq 0}-\nu_{\alpha^S_{1,B}\leq 1}-\mu^S_{1,B}\left(1-\varepsilon\right)p_{1,B}qrI^S_0=0$ and, either $\mu^S_{1,B}>0$ and $\nu_{\alpha^S_{1,B}\geq 0}>0$ and $\alpha^S_{1,B}=0$, $\left(1-\varepsilon\right)p_{1,B}qrI^S_0=D^S_0$ or any $\alpha^S_{1,B}\in[0;1]$ such that $\left(1-\alpha^S_{1,B}\right)\left(1-\varepsilon\right)p_{1,B}qrI^S_0\geq D^S_0$ is an equilibrium solution, and $V^S_{1,B}\left(I^S_0,D^S_0,p_{1,B}\right)=qrI^S_0-D^S_0$

3. If $1-(1-\varepsilon)p_{1,B}>0$, $\nu_{\alpha^S_{1,B}\leq 1}+\mu^S_{1,B}\left(1-\varepsilon\right)p_{1,B}qrI^S_0>0$. Hence, either $\alpha^S_{1,B}=1$, which can hold if and only if $D^S_0=0$, or $\mu^S_{1,B}>0$. In this case, $\alpha^S_{1,B}=1-\frac{D^S_0}{(1-\varepsilon)p_{1,B}qrI^S_0}$ and $V^S_{1,B}\left(I^S_0,D^S_0,p_{1,B}\right)=\alpha^S_{1,B}qrI^S_0=\frac{(1-\varepsilon)p_{1,B}qrI^S_0-D^S_0}{(1-\varepsilon)p_{1,B}}$

To summarize, either $p_{1,B}>0$ and in any case,

$V^{S,ND}_{1,B}\left(I^S_0,D^S_0,p_{1,B}\right)=\left(\left(1-\varepsilon\right)p_{1,B}qrI^S_0-D^S_0\right)\max\left(\frac{1}{(1-\varepsilon)p_{1,B}};1\right)$,

which holds true if $D^S_0=I^S_0=0$. Besides, if $I^S_0>0$ and $p_{1,B}qr>0$, we have

$$\alpha^S_{1,B}=\begin{cases} 0 & \text{if } p_{1,B}<\frac{1}{1-\varepsilon} \\ \in\left[0;1-\frac{D^S_0}{(1-\varepsilon)p_{1,B}qrI^S_0}\right] & \text{if } p_{1,B}=\frac{1}{1-\varepsilon} \\ 1-\frac{D^S_0}{(1-\varepsilon)p_{1,B}qrI^S_0} & \text{if } p_{1,B}>\frac{1}{1-\varepsilon}\end{cases}.$$

Or $p_{1,B}=0$, and $\overline{D}^S_0\left(I^S_0,p_{1,B}\right)=0$. In this case, the value function in case of no-default, writes $V^{S,ND}_{1,B}\left(I^S_0,D^S_0,p_{1,B}\right)=qrI^S_0$.

This proves lemma 1.

5.B.2 Proof of proposition 1

Recall that by assumption, S-banks raise riskless short-term debt. Therefore whatever the amount D^S_0 of debt they issue at $t=0$, they have to repay their creditors at $t=1$ in both states G and B.

For $p_{1,B} > 0$, the value function of a S-bank at $t = 0$ writes:

$$
\begin{aligned}
V_0^S(p_{1,B}, n_S) = \max_{D_0^S, I_0^S \geq 0} & \left[(1-p)\left((1-\varepsilon)p_{1,B}qrI_0^S - D_0^S\right)\max\left(\frac{1}{(1-\varepsilon)p_{1,B}}; 1\right) \right. \\
& \left. + p\left(RI_0^S - D_0^S\right) + \left(D_0^S + n_S - I_0^S\right)\right] \\
\text{s.t. } & D_0^S, I_0^S \geq 0 \\
& D_0^S \leq (1-\varepsilon)\, p_{1,B}qrI_0^S \\
& D_0^S \leq RI_0^S
\end{aligned}
$$

For $p_{1,B} = 0$, $V_0^S(0, n) = (1-p)qrn + pRn$ and $I_0^S = n$, $D_0^S = 0$.

Case $p_{1,B} < \frac{1}{1-\varepsilon}$**.**

If $\frac{1}{(1-\varepsilon)} > p_{1,B}$, $(1-\varepsilon)\, p_{1,B}qrI_0^S < RI_0^S$. Therefore we get rid of the last constraint and the program rewrites:

$$
\begin{aligned}
V_0^S(p_{1,B}, n_S) = \max_{D_0^S, I_0^S \geq 0} & (1-p)\left(qrI_0^S - \frac{D_0^S}{(1-\varepsilon)p_{1,B}}\right) + p\left(RI_0^S - D_0^S\right) + \left(D_0^S + n - I_0^S\right) \\
\text{s.t. } & D_0^S + n_S \geq I_0^S \\
& D_0^S \leq (1-\varepsilon)\, p_{1,B}qrI_0^S
\end{aligned}
$$

Denoting $\nu_{D_0^S \geq 0}$ the Lagrange multiplier associated to the constraint $D_0^S \geq 0$, $\nu_{I_0^S \geq 0}$ the Lagrange multiplier associated to the constraint $I_0^S \geq 0$, μ_0^S the Lagrange multiplier associated to the funding constraint and $\lambda_{1,B}^S$ the Lagrange multiplier associated to the debt constraint, the Lagrangian of the problem rewrites:

$$
\begin{aligned}
L = & (1-p)\left(qrI_0^S - \frac{D_0^S}{(1-\varepsilon)p_{1,B}}\right) + p\left(RI_0^S - D_0^S\right) + \left(D_0^S + n - I_0^S\right) \\
& + \nu_{I_0^S \geq 0} I_0^S + \nu_{D_0^S \geq 0} D_0^S + \mu_0^S\left(D_0^S + n - I_0^S\right) + \lambda_{1,B}^S\left((1-\varepsilon)\, p_{1,B}qrI_0^S - D_0^S\right)
\end{aligned}
$$

and the first order condition on I_0^S yields:

$$\frac{dL}{dI_0^S} = (1-p)qr + pR - 1 + \nu_{I_0^S \geq 0} + \lambda_{1,B}^S (1-\varepsilon) p_{1,B} qr - \mu_0^S = 0$$

This implies $\mu_0^S > 0$ and $D_0^S + n = I_0^S$.

One can replace I_0^S and rewrite the problem as

$$\begin{aligned} V_0^S(p_{1,B}, n) = \max_{D_0^S \geq 0} & (1-p)\left(qrI_0^S - \frac{D_0^S}{(1-\varepsilon)p_{1,B}}\right) + p\left(RI_0^S - D_0^S\right) \\ \text{s.t. } & D_0^S + n = I_0^S \\ & D_0^S \leq (1-\varepsilon) p_{1,B} qr I_0^S \end{aligned}$$

Or

$$\begin{aligned} V_0^S(p_{1,B}, n) = \max_{D_0^S \geq 0} & \left[(1-p)\left(qr - \frac{1}{(1-\varepsilon)p_{1,B}}\right) + p(R-1)\right] D_0^S + (pR + (1-p)qr)\, n \\ \text{s.t. } & D_0^S + n = I_0^S \\ & D_0^S \leq (1-\varepsilon) p_{1,B} qr \left(D_0^S + n\right) \end{aligned}$$

We denote $\overline{p}_1^S \equiv \frac{1}{(1-\varepsilon)} \frac{1}{qr + \frac{p(R-1)}{1-p}}$. Using $qr + \frac{p(R-1)}{1-p} > 1$ we have $\overline{p}_1^S < \frac{1}{(1-\varepsilon)}$. In the first case, the first order condition solves as follows.

1. If $0 < p_{1,B} < \overline{p}_1^S$, $\left[(1-p)\left(qr - \frac{1}{(1-\varepsilon)p_{1,B}}\right) + p(R-1)\right] < 0$ and $D_0^S = 0$, $I_0^S = n$, $V_0^S(p_{1,B}, n) = (pR + (1-p)qr)\, n$.
2. If $p_{1,B} = \overline{p}_1^S$, $\left[(1-p)\left(qr - \frac{1}{(1-\varepsilon)p_{1,B}}\right) + p(R-1)\right] = 0$, and any $D_0^S \in \left[0; \frac{(1-\varepsilon)\overline{p}_1^S qr}{1-(1-\varepsilon)\overline{p}_1^S qr} n\right]$ is an equilibrium value of D_0^S, $I_0^S = n + D_0^S$ and $V_0^S(p_{1,B}) = (pR + (1-p)qr)\, n$.
3. If $\overline{p}_1^S < p_{1,B} < \frac{1}{1-\varepsilon}$, $D_0^S = \frac{(1-\varepsilon)p_{1,B}qr}{1-(1-\varepsilon)p_{1,B}qr} n$, $I_0^S = n + D_0^S = \frac{n}{1-(1-\varepsilon)p_{1,B}qr}$, and $V_0^S(p_{1,B}) = p\left(RI_0^S - D_0^S\right) = p\left(\frac{R-(1-\varepsilon)p_{1,B}qr}{1-(1-\varepsilon)p_{1,B}qr}\right) n$

Case $p_{1,B} \geq \frac{1}{1-\varepsilon}$.

If $p_{1,B} \geq \frac{1}{1-\varepsilon}$, the program rewrites

$$\begin{aligned}V_0^S(p_{1,B},n_S) = \max_{D_0^S,I_0^S\geq 0} & (1-p)\left((1-\varepsilon)p_{1,B}qrI_0^S - D_0^S\right) + p\left(RI_0^S - D_0^S\right) + \left(D_0^S + n - I_0^S\right)\\ \text{s.t. } & D_0^S + n_S \geq I_0^S\\ & D_0^S \leq (1-\varepsilon)\,p_{1,B}qrI_0^S\\ & D_0^S \leq RI_0^S\end{aligned}$$

Denoting $\nu_{D_0^S\geq 0}$ the Lagrange multiplier associated to the constraint $D_0^S \geq 0$, $\nu_{I_0^S\geq 0}$ the Lagrange multiplier associated to the constraint $I_0^S \geq 0$, μ_0^S the Lagrange multiplier associated to the funding constraint, $\lambda_{1,B}^S$ the Lagrange multiplier associated to the debt constraint $D_0^S \leq (1-\varepsilon)\,p_{1,B}qrI_0^S$, and $\lambda_{1,G}^S$ the Lagrange multiplier associated to the debt constraint $D_0^S \leq RI_0^S$, the Lagrangian of the problem writes:

$$\begin{aligned}L = & (1-p)\left((1-\varepsilon)p_{1,B}qrI_0^S - D_0^S\right) + p\left(RI_0^S - D_0^S\right) + \left(D_0^S + n - I_0^S\right)\\ & + \nu_{I_0^S\geq 0}I_0^S + \nu_{D_0^S\geq 0}D_0^S + \mu_0^S\left(D_0^S + n - I_0^S\right) + \lambda_{1,B}^S\left((1-\varepsilon)\,p_{1,B}qrI_0^S - D_0^S\right)\\ & + \lambda_{1,G}^S\left(RI_0^S - D_0^S\right)\end{aligned}$$

and first order condition on I_0^S yields:

$$\frac{dL}{dI_0^S} = (1-p)(1-\varepsilon)p_{1,B}qr + pR - 1 + \nu_{I_0^S\geq 0} + \lambda_{1,B}^S(1-\varepsilon)\,p_{1,B}qr + \lambda_{1,G}^S R - \mu_0^S = 0$$

This implies $\mu_0^S > 0$ and $D_0^S + n = I_0^S$. Therefore the debt constraint $D_0^S \leq RI_0^S$ is always satisfied, so that we can rewrite the problem as:

$$\begin{aligned}V_0^S(p_{1,B},n_S) = \max_{D_0^S\geq 0} & \left[(1-p)\left((1-\varepsilon)p_{1,B}qr - 1\right) + p(R-1)\right]D_0^S + (1-p)(1-\varepsilon)p_{1,B}qrn + pRn\\ \text{s.t. } & D_0^S + n_S = I_0^S\\ & D_0^S \leq (1-\varepsilon)\,p_{1,B}qrI_0^S\end{aligned}$$

which implies :

1. If $\frac{1}{1-\varepsilon} \leq p_{1,B} < \frac{1}{(1-\varepsilon)qr}$, $D_0^S = \frac{(1-\varepsilon)p_{1,B}qr}{1-(1-\varepsilon)p_{1,B}qr}n$, $I_0^S = \frac{n}{1-(1-\varepsilon)p_{1,B}qr}$, and $V_0^S(p_{1,B}) = p\left(RI_0^S - D_0^S\right) = p\left(\frac{R-(1-\varepsilon)p_{1,B}qr}{1-(1-\varepsilon)p_{1,B}qr}\right)n$

2. If $p_{1,B} \geq \frac{1}{(1-\varepsilon)qr}$, $D_0^S = +\infty$, $I_0^S = +\infty$,and $V_0^S(p_{1,B}) = +\infty$

Proposition 1 obtains.

5.C Traditional bank's program

5.C.1 T-banks' debt constraint

Lemma 2 *[Debt constraint] For a given level of investment I_0^T made at $t = 0$, the maximum amount of short term debt that can be reimbursed at $t = 1$ in state B is:*

$$\overline{D}_{0,B}^T\left(I_0^T, p_{1,B}\right) = \begin{cases} k + \frac{p_1}{\delta}\left(\delta qr I_0^T - k\right)_- + \frac{p_{1,B}(1-\varepsilon)}{\delta}\left(\delta qr I_0^T - k\right)_+ & \text{if } 0 \leq p_{1,B} \leq \delta \\ k + \left(\delta qr I_0^T - k\right)_- + \frac{p_{1,B}(1-\varepsilon)}{\delta}\left(\delta qr I_0^T - k\right)_+ & \text{if } \delta \leq p_{1,B} \leq \frac{\delta}{1-\varepsilon} \\ p_{1,B} qr I_0^T (1-\varepsilon) & \text{if } p_{1,B} \geq \frac{\delta}{1-\varepsilon} \end{cases}$$

Similarly, the maximum amount of short term debt that can be reimbursed at $t = 1$ in state G is:

$$\overline{D}_{0,G}(I_0) = \delta R I_0^T$$

Proof. At $t = 1$ in state B, T-banks can generate funds either by selling a share $1 - \alpha_{1,B}^T$ of their assets, or by newly raising funds $D_{1,B}^T$. They are subject to a (i) limited liability constraint at $t = 2$, and (ii) a limit k on the amount of debt that can be guaranteed at $t = 1$ in state B. For a given level of investment $I_0^T > 0$ made at $t = 0$ and a price of capital goods $p_{1,B}qr > 0$, the maximum amount of debt issued at $t = 0$ that can be reimbursed at

$t = 1$ in state B is:

$$
\begin{aligned}
\overline{D}^T_{0,B}\left(I^T_0, p_{1,B}\right) = \max_{\alpha^T_{1,B}, I^T_{1,B}, D^T_{1,B}} & (1-\alpha^T_{1,B})p_{1,B}qrI^T_0\left(1-\varepsilon\right) + D^T_{1,B} - p_{1,B}qrI^T_{1,B} \\
\text{s.t. } & D^T_{1,B}, I^T_{1,B}, \alpha^T_{1,B} \geq 0, \alpha^T_{1,B} \leq 1 \\
& D^T_{1,B} \leq q\delta\left(\alpha^T_{1,B} r I^T_0 + rI^T_{1,B}\right) \\
& D^T_{1,B} \leq k
\end{aligned}
$$

Denoting $\lambda^T_{1,B} \geq 0$ the Lagrange multiplier associated to the funding constraint $D^T_{1,B} \leq q\delta\left(\alpha^T_{1,B} r I^T_0 + rI^T_{1,B}\right)$, $\nu_{D^T_{1,B}\leq k}$ the Lagrange multiplier associated to the constraint $D^T_{1,B} \leq k$, $\nu_{D^T_{1,B}\geq 0}$ the Lagrange multiplier associated to the constraint $D^T_{1,B} \geq 0$, $\nu_{I^T_{1,B}\geq 0}$ the Lagrange multiplier associated to the constraint $I^T_{1,B} \geq 0$, $\nu_{\alpha^T_{1,B}\geq 0}$ the Lagrange multiplier associated to the debt constraint $\alpha^T_{1,B} \geq 0$, and $\nu_{\alpha^T_{1,B}\leq 1}$ the Lagrange multiplier associated to the constraint $1-\alpha^T_{1,B} \geq 0$, the Lagrangian of the problem writes:

$$
\begin{aligned}
L = & (1-\alpha^T_{1,B})p_{1,B}qrI^T_0\left(1-\varepsilon\right) + D^T_{1,B} - p_{1,B}qrI^T_{1,B} \\
& + \lambda^T_{1,B}\left[q\delta\left(\alpha^T_{1,B} r I^T_0 + rI^T_{1,B}\right) - D^T_{1,B}\right] \\
& + \nu_{D^T_{1,B}\leq k}\left(k - D^T_{1,B}\right) \\
& + \nu_{D^T_{1,B}\geq 0} D^T_{1,B} + \nu_{I^T_{1,B}\geq 0} I^T_{1,B} \\
& + \nu_{\alpha^T_{1,B}\geq 0}\alpha^T_{1,B} + \nu_{\alpha^T_{1,B}\leq 1}\left(1-\alpha^T_{1,B}\right)
\end{aligned}
$$

The first order conditions yield

$$
\begin{aligned}
\frac{dL}{d\alpha^T_{1,B}} &= -p_{1,B}qrI^T_0\left(1-\varepsilon\right) + \lambda^T_{1,B}q\delta rI^T_0 + \nu_{\alpha^T_{1,B}\geq 0} - \nu_{\alpha^T_{1,B}\leq 1} = 0 \\
\frac{dL}{dD^T_{1,B}} &= 1 - \lambda^T_{1,B} - \nu_{D^T_{1,B}\leq k} + \nu_{D^T_{1,B}\geq 0} = 0 \\
\frac{dL}{dI^T_{1,B}} &= -p_{1,B}qr + \lambda^T_{1,B}q\delta r + \nu_{I^T_{1,B}\geq 0} = 0
\end{aligned}
$$

Hence

$$\left(\lambda^T_{1,B}\delta - p_{1,B}\left(1-\varepsilon\right)\right) qrI^T_0 + \nu_{\alpha^T_{1,B}\geq 0} - \nu_{\alpha^T_{1,B}\leq 1} = 0 \tag{5.14}$$

$$1 + \nu_{D^T_{1,B}\geq 0} = \lambda^T_{1,B} + \nu_{D^T_{1,B}\leq k} \tag{5.15}$$

$$\lambda^T_{1,B} q\delta r + \nu_{I^T_{1,B}\geq 0} = p_{1,B}qr \tag{5.16}$$

Putting (5.14) and (5.16) together we obtain:

$$\nu_{\alpha^T_{1,B}\leq 1} + \nu_{I^T_{1,B}\geq 0} I^T_0 = \nu_{\alpha^T_{1,B}\geq 0} + p_{1,B}qr\varepsilon I^T_0$$

First, we solve the problem for $I^T_0 > 0$.

1. If $p_{1,B} = 0$, $\lambda^T_{1,B} = \nu_{I^T_{1,B}\geq 0} = 0$, $\nu_{D^T_{1,B}\leq k} > 0$, $\nu_{\alpha^T_{1,B}\leq 1} = \nu_{\alpha^T_{1,B}\geq 0} = 0$. We then have $\overline{D}^T_{0,B}\left(I^T_0, p_{1,B}\right) = k$.

2. If $0 < p_{1,B} < \delta$, using (5.16) we have $\lambda^T_{1,B} < 1$, and using (5.15) we have $\nu_{D^T_{1,B}\leq k} > 0$ and $D^T_{1,B} = k$. Then, either $\lambda^T_{1,B} = 0$ and $q\delta\left(\alpha^T_{1,B} rI^T_0 + rI^T_{1,B}\right) > D^T_{1,B}$, in which case $\nu_{I^T_{1,B}\geq 0} > 0$, $\nu_{\alpha^T_{1,B}\geq 0} > 0$ which implies $\alpha^T_{1,B} = 0$, $I^T_{1,B} = 0$. This is impossible because $D^T_{1,B} = k > 0$. We must then have $\lambda^T_{1,B} > 0$ and $q\delta\left(\alpha^T_{1,B} rI^T_0 + rI^T_{1,B}\right) = D^T_{1,B}$. In that case, (5.14) and (5.16) yield $\left(\lambda^T_{1,B}\delta - p_{1,B}\left(1-\varepsilon\right)\right) qrI^T_0 = p_{1,B}qr\varepsilon I^T_0 - \nu_{I^T_{1,B}\geq 0} I^T_0 = \nu_{\alpha^T_{1,B}\leq 1} - \nu_{\alpha^T_{1,B}\geq 0}$. Rearranging, we obtain $\nu_{\alpha^T_{1,B}\leq 1} + \nu_{I^T_{1,B}\geq 0} I^T_0 = \nu_{\alpha^T_{1,B}\geq 0} + p_{1,B}qr\varepsilon I^T_0$. Two subcases arise.

   a) Either $\nu_{\alpha^T_{1,B}\leq 1} > 0$, and $I^T_{1,B} = \frac{k - q\delta rI^T_0}{q\delta r}$, $k = q\delta\left(rI^T_0 + rI^T_{1,B}\right)$. We then have $\overline{D}^T_{0,B}\left(I^T_0, p_{1,B}\right) = D^T_{1,B} - p_{1,B}qrI^T_{1,B} = k + \frac{p_{1,B}}{\delta}\left(q\delta rI^T_0 - k\right)$. This solution is an optimum if and only if $\delta qrI^T_0 \leq k$.

   b) Or $\nu_{I^T_{1,B}\geq 0} > 0$, $D^T_{1,B} = q\delta\alpha^T_{1,B} rI^T_0 = k$. We then have $\overline{D}^T_{0,B}\left(I^T_0, p_{1,B}\right) = (q\delta rI - k)\frac{p_{1,B}(1-\varepsilon)}{\delta} + k$. This solution is an optimum if and only if $\delta qrI \geq k$.

   In a nutshell $\overline{D}^T_{0,B}\left(I^T_0, p_{1,B}\right) = k + \frac{p_{1,B}}{\delta}\left(\delta qrI^T_0 - k\right)_- + \frac{p_{1,B}(1-\varepsilon)}{\delta}\left(\delta qrI^T_0 - k\right)_+$.

3. If $p_{1,B} = \delta$, two subcases arise:

a) Either $\lambda^T_{1,B} = 1$, and $\nu_{I^T_{1,B}\geq 0} = \nu_{D^T_{1,B}\geq 0} = \nu_{D^T_{1,B}\leq k} = 0, \nu_{\alpha^T_{1,B}\leq 1} > 0$. We then have $\overline{D}^T_{0,B}\left(I^T_0, p_{1,B}\right) = q\delta r I^T_0$. This solution is an optimum if and only if $\delta q r I^T_0 \leq k$.

b) Or $0 < \lambda^T_{1,B} < 1$, $\nu_{I^T_{1,B}\geq 0} > 0, \nu_{D^T_{1,B}\leq k} > 0$ and $D^T_{1,B} = k$, $I^T_{1,B} = 0$, $q\delta\alpha^T_{1,B} r I^T_0 = D^T_{1,B} = k$. We then have $\overline{D}^T_{0,B}\left(I^T_0, p_{1,B}\right) = (1-\varepsilon)\left(\delta q r I^T_0 - k\right) + k$. This solution is an optimum if and only if $\delta q r I^T_0 \geq k$.

In a nutshell $\overline{D}^T_{0,B}\left(I^T_0, p_{1,B}\right) = k + \left(\delta q r I^T_0 - k\right)_- + (1-\varepsilon)\left(\delta q r I^T_0 - k\right)_+$.

4. If $\delta < p_{1,B} < \frac{\delta}{1-\varepsilon}$, by (5.16) we have either $\lambda^T_{1,B} > 1$ or $\nu_{I^T_{1,B}\geq 0} > 0$. If $\lambda^T_{1,B} > 1$, $\nu_{D^T_{1,B}\geq 0} > 0$ by (5.15). This implies $\alpha^T_{1,B} = I^T_{1,B} = 0$. However $\lambda^T_{1,B} > 1$ also implies $\left(\lambda^T_{1,B}\delta - p_{1,B}(1-\varepsilon)\right) q r I^T_0 > 0$, which imposes $\nu_{\alpha^T_{1,B}\leq 1} > 0$ by (5.14) and contradicts $\alpha^T_{1,B} = 0$. Hence, in equilibrium $\nu_{I^T_{1,B}\geq 0} > 0$ and $0 < \lambda^T_{1,B} \leq 1$.

a) If $\lambda^T_{1,B} = 1$, $\nu_{D^T_{1,B}\geq 0} = \nu_{D^T_{1,B}\leq k} = 0$, $\nu_{\alpha^T_{1,B}\leq 1} > 0$, $D^T_{1,B} = q\delta r I^T_0 \leq k$. We then have $\overline{D}^T_{0,B}\left(I^T_0, p_{1,B}\right) = q\delta r I^T_0$. This solution is an optimum if and only if $\delta q r I^T_0 \leq k$.

b) If $\lambda^T_{1,B} < 1, \nu_{D^T_{1,B}\leq k} > 0, D^T_{1,B} = \alpha^T_{1,B} q\delta r I^T_0 = k$. We then have $\overline{D}^T_{0,B}\left(I^T_0, p_{1,B}\right) = \frac{(1-\varepsilon)p_{1,B}}{\delta}\left(\delta q r I^T_0 - k\right) + k$. This solution is an optimum if and only if $\delta q r I^T_0 \geq k$.

In a nutshell $\overline{D}^T_{0,B}\left(I^T_0, p_{1,B}\right) = k + \left(\delta q r I^T_0 - k\right)_- + \frac{(1-\varepsilon)p_{1,B}}{\delta}\left(\delta q r I^T_0 - k\right)_+$.

5. If $p_{1,B} = \frac{\delta}{1-\varepsilon}$, by (5.16) we have either $\lambda^T_{1,B} > 1$ or $\nu_{I^T_{1,B}\geq 0} > 0$. If $\lambda^T_{1,B} > 1$, $\nu_{D^T_{1,B}\geq 0} > 0$ by (5.15). This implies $\alpha^T_{1,B} = I^T_{1,B} = 0$. However $\lambda^T_{1,B} > 1$ also implies $\left(\lambda^T_{1,B}\delta - p_{1,B}(1-\varepsilon)\right) q r I >$ 0, which imposes $\nu_{\alpha^T_{1,B}\leq 1} > 0$ by (5.14) and contradicts $\alpha^T_{1,B} = 0$. Hence, in equilibrium $\nu_{I^T_{1,B}\geq 0} > 0$ and $0 < \lambda^T_{1,B} \leq 1$.

a) If $\lambda^T_{1,B} = 1, \nu_{D_{1,B^T}\geq 0} = \nu_{D^T_{1,B}\leq k} = 0, \nu_{\alpha^T_{1,B}\leq 1} = 0, D^T_{1,B} = \alpha^T_{1,B} q\delta r I^T_0 \leq k$. We then have $\overline{D}_0 = q\delta r I^T_0$. This solution is an optimum if and only if $\delta q r I^T_0 \leq k$.

b) If $\lambda^T_{1,B} < 1, \nu_{D^T_{1,B}\leq k} > 0, D^T_{1,B} = \alpha^T_{1,B} q\delta r I = k$. We then have $\overline{D}^T_{0,B}\left(I^T_0, p_{1,B}\right) = \left(\delta q r I^T_0 - k\right) + k$. This solution is an optimum if and only if $\delta q r I^T_0 \geq k$.

In a nutshell $\overline{D}^T_{0,B}\left(I^T_0, p_{1,B}\right) = \delta q r I^T_0$.

6. If $p_{1,B} > \frac{\delta}{1-\varepsilon}$, by (5.14) either $\nu_{\alpha^T_{1,B}\leq 1} > 0$, which implies $\left(\lambda^T_{1,B}\delta - p_{1,B}\left(1-\varepsilon\right)\right) > 0$ and $\lambda^T_{1,B} > 1$, therefore $\nu_{D^T_{1,B}\geq 0} > 0$ by (5.15) which contradicts $\left(\lambda^T_{1,B}\delta - p_{1,B}\left(1-\varepsilon\right)\right) > 0$ therefore this not possible. We must then have $\nu_{\alpha^T_{1,B}\leq 1} = 0$, and $\nu_{I^T_{1,B}\geq 0} > 0$ by (5.16). If $\lambda^T_{1,B}\delta - p_{1,B}\left(1-\varepsilon\right) = 0$, then $\lambda^T_{1,B} > 1$, $\nu_{D^T_{1,B}\geq 0} > 0$ by (5.15) and $D^T_{1,B} = I^T_{1,B} = \alpha^T_{1,B} = 0$. Otherwise $\lambda^T_{1,B}\delta - p_{1,B}\left(1-\varepsilon\right) < 0$ and $\nu_{\alpha^T_{1,B}\geq 0} > 0$. We then have $\overline{D}^T_{0,B}\left(I^T_0, p_{1,B}\right) = p_{1,B}qrI^T_0\left(1-\varepsilon\right)$.

Second, we solve the problem for $I^T_0 = 0$. In this case the program rewrites

$$\begin{aligned}
\overline{D}^T_{0,B}\left(I^T_0, p_{1,B}\right) &= \max_{I^T_{1,B}, D^T_{1,B}} \; D^T_{1,B} - p_{1,B}qrI^T_{1,B} \\
D^T_{1,B} &\leq q\delta\left(rI^T_{1,B}\right) \\
D^T_{1,B} &\leq k \\
D^T_{1,B}, I^T_{1,B} &\geq 0
\end{aligned}$$

It is easily shown that the previous optima also hold true when $I^T_0 = 0$. Lemma 2 obtains.

□

5.C.2 T-Banks: program at $t = 1$ in state B

We now turn to the resolution of T-banks' program at $t = 1$ in state B in case of no default, and show the following proposition.

Proposition 8 *[T-banks, time 1, B] For any $I^T_0 > 0$, $p_{1,B} > 0$ and $D^T_{0,B} \in [0; \overline{D}_{0,B}\left(I^T_0, p_{1,B}\right)]$, the time 1 equilibria and value functions, in the bad information state solve as follows:*

1. *If $D^T_0 \leq k$ then*

   a) *If $0 < p_{1,B} < \delta$, the equilibrium is such that $\alpha^T_{1,B} = 1$, $D^T_{1,B} = k$, $I^T_{1,B} = \frac{k-D^T_0}{p_{1,B}qr}$, $V^{T,ND}_{1,B}\left(I^T_0, D^T_0, p_{1,B}\right) = \frac{(\delta - p_{1,B})}{p_{1,B}}\left(k - D^R_0\right) + \delta qrI^T_0 - D^T_0$*

   b) *If $p_{1,B} = \delta$, the equilibria are such that $0 \leq I^T_{1,B} \leq \frac{k-D^T_0}{\delta qr}$, $D^T_{1,B} \in [D^T_0 + \delta qrI^T_{1,B}; \delta qrI^T_0 + \delta qrI_1]$, $\alpha^T_{1,B} = 1$, $V^{T,ND}_{1,B}\left(I^T_0, D^T_0, p_{1,B}\right) = \delta qrI - D^T_0$.*

c) *If* $\delta < p_{1,B} < \frac{\delta}{1-\varepsilon}$,*the equilibria are such that* $\alpha^T_{1,B} = 1, I^T_1 = 0$, $D^T_{1,B} \in [D^T_0; \min(\delta qr I^T_0; k)]$ *and* $V^{T,ND}_{1,B}\left(I^T_0, D^T_0, p_{1,B}\right) = \delta qr I^T_0 - D^T_0$

d) *If* $p_{1,B} = \frac{\delta}{1-\varepsilon}$, *the equilibria are* $I^T_{1,B} = 0$, $\alpha^T_{1,B} \in [0;1]$, *and* $D^T_{1,B} \in [0; \min(k; \delta q \alpha^T_{1,B} r I^T_0)]$, *such that* $D^T_0 \leq D^T_{1,B} + (1-\alpha^T_{1,B}) p_{1,B} qr I^T_0 (1-\varepsilon)$ *and* $V^{T,ND}_{1,B}\left(I^T_0, D^T_0, p_{1,B}\right) = \delta qr I^T_0 - D^T_0$.

e) *If* $p_{1,B} > \frac{\delta}{1-\varepsilon}$,*the equilibrium is such that* $D^T_{1,B} = \alpha = I^T_{1,B} = 0$ *and* $V^{T,ND}_{1,B}\left(I^T_0, D^T_0, p_{1,B}\right) = (1-\varepsilon) p_{1,B} qr I^T_0 - D^T_0$.

2. *If* $D^T_0 \geq k$, *then*

a) *If* $0 < p^T_{1,B} < \frac{\delta}{1-\varepsilon}$, *the equilibrium is such that* $I^T_{1,B} = 0$, $D^T_{1,B} = k$, $\alpha = 1 - \frac{D^T_0 - k}{p_{1,B} qr I^T_0 (1-\varepsilon)}$, $V^{T,ND}_{1,B}\left(I^T_0, D^T_0, p_{1,B}\right) = \left(1 - \frac{D^T_0 - k}{p_{1,B} qr I^T_0 (1-\varepsilon)}\right) \delta qr I^T_0 - k$

b) *If* $p_{1,B} = \frac{\delta}{1-\varepsilon}$, *the equilibria are such that* $I^T_{1,B} = 0$, $\alpha^T_{1,B} \in [0;1]$, *and* $D^T_{1,B} \in [0; \min(k; \delta q \alpha^T_{1,B} r I^T_0)]$, *with* $D^T_0 \leq D^T_{1,B} + (1-\alpha^T_{1,B}) p_{1,B} qr I^T_0 (1-\varepsilon)$ *and* $V^{T,ND}_{1,B}\left(I^T_0, D^T_0, p_{1,B}\right) = \delta qr I^T_0 - D^T_0$.

c) *If* $p_{1,B} > \frac{\delta}{1-\varepsilon}$, *the equilibrium is such that* $D^T_{1,B} = \alpha^T_{1,B} = I^T_{1,B} = 0$ *and* $V^{T,ND}_{1,B}\left(I^T_0, D^T_0, p_{1,B}\right) = p_{1,B}(1-\varepsilon) qr I^T_0 - D^T_0$.

If $p_{1,B} = 0$, $V^{T,ND}_{1,B}\left(I^T_0, D^T_0, p_{1,B}\right) = +\infty$. *Moreover, the value functions extend to the case where* $I^T_0 = 0$

Proof. We'll first solve the time 1, B bank program taking as given D^T_0 and I^T_0, in the set which ensures no-default in time 1,B. In this case, the value function of the T-bank is:

$$\begin{aligned} V^T_{1,B}\left(I^T_0, D^T_0, p_{1,B}\right) = \max_{(\alpha^T_{1,B}, D^T_{1,B}, I^T_{1,B}) \in [0;1]\times[0;k]\times\mathbb{R}_+} & \left(\delta - p_{1,B}\right) qr I^T_{1,B} + \alpha^T_{1,B} \delta qr I^T_0 + (1-\alpha^T_{1,B}) p_{1,B} qr I^T_0 (1-\varepsilon) - D^T_0 \\ \text{s.t. } & (1-\alpha^T_{1,B}) p_{1,B} qr I^T_0 (1-\varepsilon) + D_{1,B} \geq D^T_0 + p_{1,B} qr I^T_{1,B} \\ & D_{1,B} \leq q\delta \left(r \alpha^T_{1,B} I^T_0 + r I^T_{1,B}\right) \end{aligned}$$

The Lagrangian writes as follows:

$$
\begin{aligned}
L_{1,B}^T = & (\delta - p_{1,B})\, qrI_{1,B}^T + \alpha_{1,B}^T \delta qr I_0^T + (1-\alpha_{1,B}^T) p_{1,B} qr I_0^T (1-\varepsilon) - D_0^T \\
& + \lambda_1 \left((1-\alpha_{1,B}^T) p_{1,B} qr I_0^T (1-\varepsilon) + D_{1,B}^T - D_0^T - p_{1,B} qr I_{1,B}^T \right) \\
& + \lambda_2 \left(\delta q \left(\alpha_{1,B}^T r I_0^T + r I_1 \right) - D_{1,B}^T \right) \\
& + \nu_{D_{1,B}^T \le k} \left(k - D_{1,B}^T \right) \\
& + \nu_{D_{1,B}^T \ge 0} D_{1,B}^T + \nu_{\alpha_{1,B}^T \ge 0} \alpha_{1,B}^T + \nu_{\alpha_{1,B}^T \le 1} \left(1 - \alpha_{1,B}^T \right) + \nu_{I_{1,B}^T \ge 0} I_{1,B}^T
\end{aligned}
$$

where ν_c is the Lagrange multiplier associated to the constraint c, λ_1 the Lagrange multiplier associated to

which yields

$$\frac{dL_{1,B}^T}{dD_{1,B}^T} = \lambda_1 - \lambda_2 - \nu_{D_{1,B}^T \le k} + \nu_{D_{1,B}^T \ge 0} = 0 \tag{5.17}$$

$$\frac{dL_{1,B}^T}{d\alpha_{1,B}^T} = qrI_0^T \left(\delta - p_{1,B} (1-\varepsilon) \right) - \lambda_1 p_{1,B} qr I_0^T (1-\varepsilon) + \lambda_2 \delta qr I_0^T + \nu_{\alpha_{1,B}^T \ge 0} - \nu_{\alpha_{1,B}^T \le 1} = 0 \tag{5.18}$$

$$\frac{dL_{1,B}^T}{dI_{1,B}^T} = (\delta - p_{1,B})\, qr - \lambda_1 p_{1,B} qr + \lambda_2 \delta qr + \nu_{I_{1,B}^T \ge 0} = 0 \tag{5.19}$$

We will focus on cases where $k > 0$, and $I_0^T > 0$ to start with.

Equation (5.17) together with equation (5.19) yield:

$$(\delta - p_{1,B})\, qr\, (1+\lambda_2) + \nu_{I_{1,B}^T \ge 0} + (\nu_{D_{1,B}^T \ge 0} - \nu_{D_{1,B}^T \le k}) p_{1,B} qr = 0 \tag{5.20}$$

and equation (5.17) together with equation (5.18):

$$qrI_0^T \left(\delta - p_{1,B} (1-\varepsilon) \right) (1+\lambda_2) + \nu_{\alpha_{1,B}^T \ge 0} - \nu_{\alpha_{1,B}^T \le 1} + \left(\nu_{D_{1,B}^T \ge 0} - \nu_{D_{1,B}^T \le k} \right) p_{1,B} qr I_0^T (1-\varepsilon) = 0 \tag{5.21}$$

Finally, equations (5.20) and (5.21) give:

$$\varepsilon qr I_0^T \delta \left(1+\lambda_2\right) + \nu_{\alpha_{1,B}^T \geq 0} = \nu_{\alpha_{1,B}^T \leq 1} + \nu_{I_{1,B}^T \geq 0} I_0^T \left(1-\varepsilon\right) \tag{5.22}$$

Different cases should be treated:

Case $p_{1,B} = 0$

In this case

$$\begin{aligned} \lambda_1 - \lambda_2 - \nu_{D_{1,B} \leq k} + \nu_{D_{1,B} \geq 0} &= 0 \\ qr I_0^T \delta + \lambda_2 \delta qr I_0^T + \nu_{\alpha_{1,B}^T \geq 0} - \nu_{\alpha_{1,B}^T \leq 1} &= 0 \\ \delta qr + \lambda_2 \delta qr + \nu_{I_{1,B}^T \geq 0} &= 0 \end{aligned}$$

Hence, $I_1 = +\infty, \alpha = 1$ and $V_{1,B}^T = +\infty$.

Case $0 < p_{1,B} < \delta$

In this case, $\lambda_1 > 0$ and $\nu_{D_{1,B} \leq k} > 0$ (from equations (5.19) and (5.20))

Hence $D_{1,B} = k$, $\nu_{D_{1,B} \geq 0} = 0$, and $(1-\alpha_{1,B}^T) p_{1,B} qr I_0^T \left(1-\varepsilon\right) + D_{1,B}^T = D_0^T + p_{1,B} qr I_{1,B}^T$

Equation (5.22) implies either $\nu_{\alpha_{1,B}^T \leq 1} > 0$ or $\nu_{I_{1,B}^T \geq 0} > 0$

1. Hence, either $\nu_{\alpha_{1,B}^T \leq 1} > 0$ and $\alpha_{1,B}^T = 1$, $D_{1,B} = k$, $I_{1,B}^T = \frac{k - D_0^T}{p_{1,B} qr}$. In this case $V_{1,B}^T = \frac{(\delta - p_{1,B})}{p_{1,B}} \left(k - D_0^T\right) + \delta qr I_0^T - D_0^T$. This would be an optimum if and only if $D_0^T \leq k$.

2. Or $\nu_{I_{1,B}^T \geq 0} > 0$, and $I_{1,B}^T = 0$, $(1-\alpha) p_{1,B} qr I_{0,B}^T \left(1-\varepsilon\right) + k = D_0^T$, hence $\alpha_{1,B}^T = 1 - \frac{D_0^T - k}{p_{1,B} qr I_0^T (1-\varepsilon)}$. In this case, $V_{1,B}^T = \left(1 - \frac{D_0^T - k}{p_{1,B}^T qr I_0^T (1-\varepsilon)}\right) \delta qr I_0^T - k$. This would be an optimum if and only if $D_0^T \geq k$.

Case $p_{1,B} = \delta$

Again, equation (5.22) yields

1. Either $\nu_{\alpha^T_{1,B}\leq 1} > 0$ and $\alpha^T_{1,B}$.

    a) Then, either $\lambda_1 = \lambda_2, \nu_{D^T_{1,B}\leq k} = \nu_{D^T_{1,B}\geq 0} = 0, \nu_{I^T_{1,B}\geq 0} = 0$. In this case any $0 \leq I^T_{1,B} \leq \frac{k-D^T_0}{\delta qr}$ is an equilibrium value, any $D^T_{1,B}$ such that $\delta qrI^T_0 + \delta qrI^T_{1,B} \geq D^T_{1,B} \geq D^T_0 + \delta qrI^T_{1,B}$, $V^T_{1,B} = \delta qrI^T_{0,B} - D^T_0$, as well, and these are equilibria if and only if $D^T_0 \leq k$.

    b) Or $\lambda_1 > \lambda_2, \nu_{I^T_{1,B}\geq 0} > 0$, $\nu_{D^T_{1,B}\leq k} > 0$ and $I^T_{1,B} = 0, D^T_{1,B} = k, \alpha^T_{1,B} = 1 - \frac{D^T_0 - k}{\delta qrI^T_0(1-\varepsilon)} = 1$. In this case, $V^T_{1,B} = \delta qrI^T_0 - k$. This is an equilibrium if and only if $D^T_0 = k$.

2. Or $\nu_{I^T_{1,B}\geq 0} > 0$ and $I^T_{1,B} = 0$.

    a) In this case, $\lambda_2 < \lambda_1$, $\lambda_1 > 0$, $\nu_{D^T_{1,B}\leq k} > 0$, $\alpha^T_{1,B} = 1 - \frac{D^T_0 - k}{\delta qrI^T_0(1-\varepsilon)}$, and $I^T_{1,B} = 0, D^T_{1,B} = k$. In this case, $V^T_{1,B} = \left(1 - \frac{D^T_0 - k}{\delta qrI^T_0(1-\varepsilon)}\right)\delta qrI^T_0 - k$. This is an equilibrium if and only if $k \leq D^T_0$.

Case $\delta < p_{1,B} < \frac{\delta}{1-\varepsilon}$

Then

$$\nu_{D^T_{1,B}\geq 0}p_{1,B}qr + \nu_{I^T_{1,B}\geq 0} = -\left(\delta - p_{1,B}\right)qr\left(1+\lambda_2\right) + \nu_{D^T_{1,B}\leq k}p_{1,B}qr$$

$$\lambda_1 + \nu_{D^T_{1,B}\geq 0} = \lambda_2 + \nu_{D^T_{1,B}\leq k}$$

$$\left(\nu_{D^T_{1,B}\leq k} - \nu_{D^T_{1,B}\geq 0}\right)p_{1,B}qrI^T_0\left(1-\varepsilon\right) + \nu_{\alpha^T_{1,B}\leq 1} = qrI^T_0\left(\delta - p_{1,B}\left(1-\varepsilon\right)\right)\left(1+\lambda_2\right) + \nu_{\alpha^T_{1,B}\geq 0}$$

With $qrI^T_0\left(\delta - p_{1,B}\left(1-\varepsilon\right)\right)\left(1+\lambda_2\right) > 0$ and $-\left(\delta - p_{1,B}\right)qr\left(1+\lambda_2\right) > 0$.

Hence, either $\nu_{I^T_{1,B}\geq 0} = 0, \nu_{D^T_{1,B}\geq 0} > \nu_{D^T_{1,B}\leq k} = 0, \nu_{\alpha^T_{1,B}\leq 1} > 0$ and $\lambda_2 > 0$. This cannot occur as $I^T_0 > 0$.

This implies $\nu_{I^T_{1,B}\geq 0} > 0$.

1. Then, either $\nu_{D^T_{1,B}\geq 0} > \nu_{D^T_{1,B}\leq k} = 0, \nu_{\alpha^T_{1,B}\leq 1} > 0$ and $\lambda_2 > 0$ which can't occur as $I^T_0 > 0$.

2. Finally $\nu_{D^T_{1,B}\geq 0} = 0$.

a) Then, either $\nu_{D^T_{1,B}\leq k} = 0, \nu_{\alpha^T_{1,B}\leq 1} > 0$, in which case $\alpha^T_{1,B} = 1, I^T_{1,B} = 0, q\delta r I^T_0 \geq D^T_{1,B} \geq D^T_0$, and $D^T_{1,B} \leq k$. and $V^T_{1,B} = \delta qr I^T_0 - D^T_0$. This is an optimum if and only if $D^T_0 \leq k$.

b) Or $\nu_{D^T_{1,B}\leq k} > 0, \lambda_1 > 0$, in which case $\alpha^T_{1,B} = 1 - \frac{D^T_0 - k}{\delta qr I^T_0(1-\varepsilon)}, I^T_{1,B} = 0, D^T_{1,B} = k$ and the optimum is so defined if and only if $k \leq D^T_0$. In this case, $V^T_{1,B} = \left(1 - \frac{D^T_0 - k}{\delta qr I^T_0(1-\varepsilon)}\right)\delta qr I^T_0 - k$.

Case $p_{1,B} = \frac{\delta}{1-\varepsilon}$

Again, $\nu_{D^T_{1,B}\geq 0} > 0 = \nu_{D^T_{1,B}\leq k}, \lambda_2 > 0, \nu_{\alpha^T_{1,B}\leq 1} > 0$ cannot occur as $I^T_0 > 0$.

We then have $\nu_{I^T_{1,B}\geq 0} > 0$ and $\nu_{D^T_{1,B}\geq 0} = 0$. In this case,

$$\nu_{\alpha^T_{1,B}\geq 0} = \nu_{\alpha^T_{1,B}\leq 1} + \nu_{D^T_{1,B}\leq k}\delta qr I^T_0$$

and $\nu_{\alpha^T_{1,B}\leq 1} = 0$.

Finally, $I^T_{1,B} = 0$, any $\alpha^T_{1,B} \in [0;1]$, and any $D^T_{1,B} \in [0; \min(k; \delta q \alpha^T_{1,B} r I^T_0)]$ such that $D^T_0 \leq D^T_{1,B} + (1-\alpha^T_{1,B}) p_{1,B} qr I^T_0 (1-\varepsilon)$ is an equilibrium, and $V^T_{1,B} = \delta qr I^T_0 - D^T_0$.

Case $p_{1,B} > \frac{\delta}{1-\varepsilon}$

In this case, we have: $\nu_{I^T_{1,B}\geq 0} + \nu_{D^T_{1,B}\geq 0} p_{1,B} qr = \nu_{D^T_{1,B}\leq k} p_{1,B} qr + (p_{1,B} - \delta)\, qr\, (1+\lambda_2) > 0$ and $\nu_{\alpha^T_{1,B}\geq 0} + \nu_{D^T_{1,B}\geq 0} p_{1,B} qr I^T_0\, (1-\varepsilon) = \nu_{\alpha^T_{1,B}\leq 1} + \nu_{D^T_{1,B}\leq k} p_{1,B} qr I^T_0\, (1-\varepsilon) + qr I^T_0\, (p_{1,B}\,(1-\varepsilon) - \delta)\,(1+\lambda_2) > 0$

Then

1. Either $\nu_{\alpha^T_{1,B}\geq 0} > 0, \nu_{\alpha^T_{1,B}\leq 1} = 0$, in which case we have from equation (5.22) $\nu_{I^T_{1,B}\geq 0} > 0$. Hence, $I^T_{1,B} = 0, \alpha^T_{1,B} = 0, D^T_{1,B} = 0, V^T_{1,B} = p_{1,B}\,(1-\varepsilon)\, qr I^T_0 - D^T_0$.

2. Otherwise $\nu_{\alpha^T_{1,B}\geq 0} = 0, \nu_{D^T_{1,B}\geq 0} > 0, \lambda_2 > 0$ and $\alpha^T_{1,B} = I^T_{1,B} = 0$.

In any case, $D^T_{1,B} = \alpha^T_{1,B} = I^T_{1,B} = 0$ and $V^T_{1,B} = p_{1,B}\,(1-\varepsilon)\, qr I^T_0 - D^T_0$.

If $I_0^T = 0$, the program rewrites

$$V_{1,B}^{T,ND}\left(0, D_0^T, p_{1,B}\right) = \max_{\left(D_{1,B}^T, I_{1,B}^T\right)\in[0;k]\times\mathbb{R}_+} \left(\delta - p_{1,B}\right) qrI_{1,B}^T - D_0^T$$

$$\text{s.t. } D_{1,B}^T \geq D_0^T + p_{1,B}qrI_{1,B}^T$$

$$D_{1,B}^T \leq q\delta\left(rI_{1,B}^T\right)$$

and previous results extend as far as the value function is concerned. □

5.C.3 T-banks: Time 0 program

Let's define $k^* = \frac{\delta qrn}{1-\delta qr}$, $p_{1,L}^T = \frac{\delta}{\delta qr + \frac{p(\delta R-1)}{1-p}}$, $p_{1,H}^{TB} = \frac{p_{1,L}^T}{1-\varepsilon}$

We can now solve for the T-bank's program as of date 0, taking as given n and $p_{1,B}$

Proposition 9 *[T-banks, time 0] The solution to the bank's optimization program is as follows:*

1. *If $0 < k \leq k^*$*

   a) If $0 < p_{1,B} < p_{1,L}^T$, $D_0^T = 0, I_0^T = n$, $V_0^T = p\left(\delta Rn\right) + (1-p)\left(\delta qrn - k + \frac{\delta}{p_{1,B}}k\right)$

   b) If $p_{1,B} = p_{1,L}^T$, any $D_0^T \in [0;k], I_0^T = n + D_0^T$ is an equilibrium solution and $V_0^T = p\left(\delta R(n+k) - k\right) + (1-p)\left(\delta qr(n+k) - k\right)$

   c) If $p_{1,L}^T < p_{1,B} < p_{1,H}^T$, $D_0^T = k, I_0^T = k+n$ and $V_0^T = p\left(\delta R(n+k) - k\right) + (1-p)\left(\delta qr(n+k) - k\right)$

   d) If $p_{1,B} = p_{1,H}^T$, any $D_0^T \in \left[k; \frac{k\left(1-\frac{p_{1,B}(1-\varepsilon)}{\delta}\right) + (1-\varepsilon)p_{1,B}qrn}{1-p_{1,B}(1-\varepsilon)qr}\right]$ is an equilibrium solution, $I_0^T = n + D_0^T$ and $V_0^T = p\left(\delta R(n+k) - k\right) + (1-p)\left(\delta qr(n+k) - k\right)$

   e) If $p_{1,H}^T < p_{1,B} \leq \frac{\delta}{1-\varepsilon}$, $D_0^T = \frac{k\left(1-\frac{p_{1,B}(1-\varepsilon)}{\delta}\right) + (1-\varepsilon)p_{1,B}qrn}{1-p_{1,B}(1-\varepsilon)qr}$, $I_0^T = n + D_0^T$ and $V_0^T = p\left(\delta R\left(D_0 + n\right) - D_0\right)$

   f) If $\frac{\delta}{1-\varepsilon} \leq p_{1,B} < \frac{1}{(1-\varepsilon)qr}$, $D_0^T = \frac{p_{1,B}(1-\varepsilon)qrn}{\left(1-p_{1,B}(1-\varepsilon)qr\right)}$, $I_0^T = D_0^T + n$, and $V_0^T = p\left(\delta R\left(D_0 + n\right) - D_0\right)$

   g) If $p_{1,B} \geq \frac{1}{(1-\varepsilon)qr}$, $D_0^T = +\infty$, $I_0^T = +\infty$ and $V_0^T = +\infty$

   Moreover, if $p_{1,B} = 0$, $V_0^T = +\infty$

2. *If* $k > k^*$

   a) *If* $0 < p_{1,B} \leq p^T_{1,L}$, $D^T_0 = 0, I^T_0 = n$, $V^T_0 = p\left(\delta Rn\right) + (1-p)\left(\delta qrn - k + \frac{\delta}{p_{1,B}}k\right)$

   b) *If* $p_{1,B} = p^T_{1,L}$, *any* $D^T_0 \in [0; \frac{k\left(1-\frac{p_{1,B}}{\delta}\right)+p_{1,B}qrn}{1-p_{1,B}qr}]$ *is an equilibrium solution,* $I^T_0 = D^T_0 + n$ *and* $V^T_0 = p\left(\delta R(n+k) - k\right) + (1-p)\left(\delta qr(n+k) - k\right)$

   c) *If* $p^T_{1,L} < p_{1,B} < \delta$, $D_0 = \frac{k\left(1-\frac{p_{1,B}}{\delta}\right)+p_{1,B}qrn}{1-p_{1,B}qr}$, $I^T_0 = D^T_0 + n$ *and* $V^T_0 = p\left(\delta R\left(D^T_0 + n\right) - D^T_0\right)$

   d) *If* $\delta \leq p_{1,B} \leq \frac{\delta}{1-\varepsilon}$, $D^T_0 = \frac{\delta qrn}{1-\delta qr}$, $I^T_0 = D^T_0 + n$ *and* $V^T_0 = p\left(\delta RI - D_0\right) + (1-p)\left(\delta qrI - D^R_0\right)$.

   e) *If* $\frac{\delta}{1-\varepsilon} \leq p_{1,B} < \frac{1}{(1-\varepsilon)qr}$, $D^T_0 = \frac{p_{1,B}(1-\varepsilon)qrn}{\left(1-p_{1,B}(1-\varepsilon)qr\right)}$, $I^T_0 = D^T_0 + n$, *and* $V^T_0 = p\left(\delta R\left(D_0 + n\right) - D_0\right)$

   f) *If* $p_{1,B} \geq \frac{1}{(1-\varepsilon)qr}$, $D^T_0 = +\infty$, $I^T_0 = +\infty$ *and* $V^T_0 = +\infty$

Moreover, if $p_{1,B} = 0$, $V^T_0 = +\infty$

Proof. As S-banks, T-banks can only raise funds in the form of riskless short term debt. They will choose a debt level which ensures that all debt is reimbursed in time 1, with certainty. The T-bank's program for a given level n of own funds and a given price in time 1, B market $p_{1,B}$ writes

$$V^T_0(p_{1,B}, n) = \max_{I^T_0, D^T_0}\left(D^T_0 + n - I^T_0\right) + pV^{T,ND}_{1,G}\left(I^T_0, D^T_0, p_{1,B}\right) + (1-p)V^{T,ND}_{1,B}\left(I^T_0, D^T_0, p_{1,B}\right) \tag{5.23}$$

$$D^T_0 \leq \delta RI^T_0 \tag{5.24}$$

$$D^T_0 \leq \overline{D}^T_0\left(p_{1,B}, I^T_0\right) \tag{5.25}$$

$$D^T_0 + n \geq I^T_0 \tag{5.26}$$

$$I^T_0, D^T_0 \geq 0 \tag{5.27}$$

with

$$\overline{D}_0^T\left(I_0^T, p_{1,B}\right) = \begin{cases} k + \frac{p_{1,B}}{\delta}\left(\delta q r I_0^T - k\right)_- + \frac{p_{1,B}(1-\varepsilon)}{\delta}\left(\delta q r I_0^T - k\right)_+ & \text{if } 0 \leq p_{1,B} \leq \delta \\ k + \left(\delta q r I_0^T - k\right)_- + \frac{p_{1,B}(1-\varepsilon)}{\delta}\left(\delta q r I_0^T - k\right)_+. & \text{if } \delta \leq p_{1,B} \leq \frac{\delta}{1-\varepsilon} \\ p_{1,B} q r I_0^T (1-\varepsilon) & \text{if } p_{1,B} \geq \frac{\delta}{1-\varepsilon} \end{cases}$$

$$V_{1,G}^{T,ND}\left(I_0^T, D_0^T, p_{1,B}\right) = \delta R I_0^T - D_0^T$$

and, as shown in the previous proposition,

$$\begin{aligned} V_{1,B}^{T,ND}\left(I_0^T, D_0^T, p_{1,B}\right) = &\left(\frac{(\delta - p_{1,B})_+}{p_{1,B}}\left(k - D_0^T\right)_+ + \left(k - D_0^T\right)_- \frac{(\delta - p_{1,B}(1-\varepsilon))_+}{p_{1,B}(1-\varepsilon)}\right) \\ &+ (p_{1,B}(1-\varepsilon) - \delta)_+ q r I_0^T + \delta q r I_0^T - D_0^T \end{aligned}$$

The first thing to notice is that, as for S-banks, T-banks are always binding their date 0 funding constraint (equation (5.26)): indeed, they always prefer investing one unit of funds in the investment technology available to them at date 0, which yields at least an expected $\delta(pR + (1-p)qr)$ than consuming it at date 0 and getting 1.

The program of the bank at date 0 can then be rewritten:

$$\begin{aligned} V_0^T(p_{1,B}, n) = \max_{D_0^T \geq 0}\; & p V_{1,G}^{T,ND}\left(I_0^T, D_0^T, p_{1,B}\right) + (1-p) V_{1,B}^{T,ND}\left(I_0^T, D_0^T, p_{1,B}\right) \\ & D_0^T \leq \overline{D}_0^T\left(p_{1,B}, I_0^T\right) \\ & D_0^T + n = I_0^T \end{aligned}$$

We solve this program according to the values of $p_{1,B}$, and split it into three subprograms to ease the resolution: the solution to our program is the maximum of the solution of the three subprograms defined that way.

Case $p_{1,B} = 0$

If $p_{1,B} = 0$, as $V_{1,B}^{T,ND}\left(I_0^T, D_0^T, 0\right) = +\infty$ whatever D_0^T and I_0^T, we have $V_0^T(0, n) = +\infty$

Case $0 < p_{1,B} < \delta$

When $0 < p_{1,B} < \delta$, we get $\overline{D}_0^T\left(I_0^T, p_{1,B}\right) = k + \frac{p_{1,B}}{\delta}\left(\delta qrI_0^T - k\right)_- + \frac{p_{1,B}(1-\varepsilon)}{\delta}\left(\delta qrI_0^T - k\right)_+$ and

$$V_{1,B}^{T,ND}\left(I_0^T, D_0^T, p_{1,B}\right) = \left(\frac{\left(\delta - p_{1,B}\right)_+}{p_{1,B}}\left(k - D_0^T\right)_+ + \left(k - D_0^T\right)_- \frac{\left(\delta - p_{1,B}\left(1-\varepsilon\right)\right)_+}{p_{1,B}\left(1-\varepsilon\right)}\right) + \delta qrI_0^T - D_0^T$$

First subprogram We first focus on the subprogram where we additionally constrain I_0^T to evolve in the set where $\delta qrI_0^T \leq k$.

In this case, $\overline{D}_0^T\left(p_{1,B}, I_0^T\right) \leq k + p_{1,B}\left(qrI_0^T - \frac{k}{\delta}\right) \leq k$ and the program sums up as follows:

$$\begin{aligned}
V_0^T(p_{1,B}, n) &= \max_{D_0^T \geq 0} p\left(\delta RI_0^T - D_0^T\right) + (1-p)\left(\delta qrI_0^T - k + \frac{\delta}{p_{1,B}}\left(k - D_0^T\right)\right) \\
D_0^T &\leq k + p_{1,B}\left(qrI_0^T - \frac{k}{\delta}\right) \\
\delta qrI_0^T &\leq k \\
D_0^T + n &= I_0^T
\end{aligned}$$

Or,

$$\begin{aligned}
V_0^T(p_{1,B}, n) &= \max_{D_0^T \geq 0}\left(p\delta R + (1-p)\delta qr - 1 - (1-p)\frac{\delta - p_{1,B}}{p_{1,B}}\right)D_0^T + \left(p\delta R + (1-p)\delta qr\right)n + (1-p)\frac{\delta - p_{1,B}}{p_{1,B}}k \\
D_0^T &\leq k + p_{1,B}\left(qrI_0^T - \frac{k}{\delta}\right) \\
\delta qrI_0^T &\leq k \\
D_0^T + n &= I_0^T
\end{aligned}$$

This sub-program has a non-empty set of solutions if and only if $\delta qrn \leq k$. In which case:

1. $p\delta R+(1-p)\delta qr-1-(1-p)\frac{\delta-p_{1,B}}{p_{1,B}}<0$, $D_0^T=0$ and $V_0^T(p_{1,B},n)=(p\delta R+(1-p)\delta qr)\,n+(1-p)\frac{\delta-p_{1,B}}{p_{1,B}}k$

2. $p\delta R+(1-p)\delta qr-1-(1-p)\frac{\delta-p_{1,B}}{p_{1,B}}=0$ and any D_0^T such that $0\leq D_0^T\leq\min\left(\frac{k\left(1-\frac{p_{1,B}}{\delta}\right)+p_{1,B}qrn}{1-p_{1,B}qr};\frac{k-\delta qrn}{\delta qr}\right)$ is an optimum. In this case, $V_0^T(p_{1,B},n)=(p\delta R+(1-p)\delta qr)\,n+(1-p)\frac{\delta-p_{1,B}}{p_{1,B}}k$

3. $p\delta R+(1-p)\delta qr-1-(1-p)\frac{\delta-p_{1,B}}{p_{1,B}}>0$ and $D_0^T{=}\min\left(\frac{k\left(1-\frac{p_{1,B}}{\delta}\right)+p_{1,B}qrn}{1-p_{1,B}qr};\frac{k-\delta qrn}{\delta qr}\right)$. In this case, $V_0^T(p_{1,B},n)=p\left(\delta R(D_0^T+n)-D_0^T\right)$

Notice that $\frac{k\left(1-\frac{p_{1,B}}{\delta}\right)+p_{1,B}qrn}{1-p_{1,B}qr}\leq\frac{k-\delta qrn}{\delta qr}$ if and only if $k\geq k^\star$

Second subprogram Let's now turn to the subprogram where we additionally constrain D_0^T and I_0^T to evolve in the set where $\delta qrI_0^T\geq k$ and $D_0^T\leq k$.

In this case, the program writes

$$\begin{aligned}V_0^T(p_{1,B},n)=\max_{D_0^T\geq 0}\ & p\left(\delta RI_0^T-D_0^T\right)+(1-p)\left(\delta qrI_0^T-k+\frac{\delta}{p_{1,B}}\left(k-D_0^T\right)\right)\\ D_0^T&\leq k\\ \delta qrI_0^T&\geq k\\ D_0^T+n&=I_0^T\end{aligned}$$

This program has a non-empty set of solutions if and only if $\frac{k-\delta qrn}{\delta qr}\leq k$. or $k\leq k^\star$

And, in this case, either,

1. $p\delta R+(1-p)\delta qr-1-(1-p)\frac{\delta-p_{1,B}}{p_{1,B}}<0$ and $D_0^T=\max(0;\frac{k-\delta qrn}{\delta qr})$ is the solution to this program. We also have $V_0^T(p_{1,B},n)=(p\delta R+(1-p)\delta qr)\,n+(1-p)\frac{\delta-p_{1,B}}{p_{1,B}}k$ if $k\leq\delta qrn$ and $V_0^T(p_{1,B},n)=p\left(\delta R(\frac{k-\delta qrn}{\delta qr}+n)-\frac{k-\delta qrn}{\delta qr}\right)+(1-p)\left(\delta qr(\frac{k-\delta qrn}{\delta qr}+n)-k+\frac{\delta}{p_{1,B}}\left(k-\frac{k-\delta qrn}{\delta qr}\right)\right)$ otherwise.

2. $p\delta R+(1-p)\delta qr-1-(1-p)\frac{\delta-p_{1,B}}{p_{1,B}}=0$ and any $D_0^T\in[\max(0;\frac{k-\delta qrn}{\delta qr});k]$ is a solution to this program. We also have $V_0^T(p_{1,B},n)=(p\delta R+(1-p)\delta qr)\,n+(1-p)\frac{\delta-p_{1,B}}{p_{1,B}}k$ if

$k \leq \delta qrn$ and $V_0^T(p_{1,B},n) = p\left(\delta R(\frac{k-\delta qrn}{\delta qr}+n) - \frac{k-\delta qrn}{\delta qr}\right) + (1-p)\left(\delta qr(\frac{k-\delta qrn}{\delta qr}+n) - k + \frac{\delta}{p_{1,B}}\left(k - \frac{k-\delta qrn}{\delta qr}\right)\right)$ otherwise.

3. $p\delta R + (1-p)\delta qr - 1 - (1-p)\frac{\delta - p_{1,B}}{p_{1,B}} > 0$ and $D_0^T = k$ is the solution to this program. We also have $V_0^T(p_{1,B},n) = \delta(pR + (1-p)qr)(n+k) - k$

Third subprogram Finally, in the last subprogram we additionally constrain D_0^T and I_0^T to evolve in the set where $\delta qrI_0^T \geq k$ and $D_0^T \geq k$.

In this case,

$$
\begin{aligned}
V_0^T(p_{1,B},n) = &\max_{D_0^T \geq 0} p\left(\delta RI_0^T - D_0^T\right) + (1-p)\left(\delta qrI_0^T - k + \frac{\delta}{p_{1,B}(1-\varepsilon)}\left(k - D_0^T\right)\right) \\
&k \leq D_0^T \leq p_{1,B}(1-\varepsilon)\left(qrI_0^T - \frac{k}{\delta}\right) + k \\
&\delta qrI_0^T \geq k \\
&D_0^T + n = I_0^T
\end{aligned}
$$

It has a non-empty set of solutions if and only if $\frac{k-\delta qrn}{\delta qr} \leq \frac{k\left(1-\frac{p_{1,B}(1-\varepsilon)}{\delta}\right)+(1-\varepsilon)p_{1,B}qrn}{1-p_{1,B}(1-\varepsilon)qr}$ (or $k \leq k^\star$). In this range

1. Either $p\delta R + (1-p)\delta qr - 1 - (1-p)\frac{\delta-(1-\varepsilon)p_{1,B}}{p_{1,B(1-\varepsilon)}} < 0$, $D_0^T = \max(k; \frac{k-\delta qrn}{\delta qr}) = k$, and $V_0^T(p_{1,B},n) = \delta(pR + (1-p)qr)(n+k) - k$

2. Or $p\delta R + (1-p)\delta qr - 1 - (1-p)\frac{\delta-(1-\varepsilon)p_{1,B}}{(1-\varepsilon)p_{1,B}} = 0$ and any $D_0 \in [\max(k; \frac{k-\delta qrn}{\delta qr}); \frac{k\left(1-\frac{p_{1,B}(1-\varepsilon)}{\delta}\right)+(1-\varepsilon)p_{1,B}qrn}{1-p_{1,B}(1-\varepsilon)qr}] = [k; \frac{k\left(1-\frac{p_{1,B}(1-\varepsilon)}{\delta}\right)+(1-\varepsilon)p_{1,B}qrn}{1-p_{1,B}(1-\varepsilon)qr}]$ is an equilibrium, and $V_0^T(p_{1,B},n) = \delta(pR + (1-p)qr)(n+k) - k$

3. Or $p\delta R + (1-p)\delta qr - 1 - (1-p)\frac{\delta-(1-\varepsilon)p_{1,B}}{(1-\varepsilon)p_{1,B}} > 0$ and $D_0^T = \frac{k\left(1-\frac{p_{1,B}(1-\varepsilon)}{\delta}\right)+(1-\varepsilon)p_{1,B}qrn}{1-p_{1,B}(1-\varepsilon)qr}$. In this case, $V_0^T(p_{1,B},n) = p\left(\delta R(D_0^T + n) - D_0^T\right)$

Now, let's distinguish three cases.

Case $\delta qrn \geq k$ In this case, the first program delivers an empty set of solutions. The overall maximum is the maximum of the two other subprograms.

1. If $p\delta R + (1-p)\delta qr - 1 - (1-p)\frac{\delta - p_{1,B}}{p_{1,B}} < 0$, $D_0^T = 0$, $V_0^T(p_{1,B}, n) = p(\delta Rn) + (1-p)\left(\delta qrn - k + \frac{\delta}{p_{1,B}}k\right)$

2. If $p\delta R + (1-p)\delta qr - 1 - (1-p)\frac{\delta - p_{1,B}}{p_{1,B}} = 0$, any $D_0^T \in [0; k]$ is an optimum debt level and $V_0^T(p_{1,B}, n) = p(\delta R(n+k) - k) + (1-p)(\delta qr(n+k) - k)$

3. If $p\delta R + (1-p)\delta qr - 1 - (1-p)\frac{\delta - p_{1,B}}{p_{1,B}} > 0$, and $p_{1,B} < \delta$, and $p\delta R + (1-p)\delta qr - 1 - (1-p)\frac{\delta - (1-\varepsilon)p_{1,B}}{p_{1,B(1-\varepsilon)}} < 0$, $D_0^T = k$,and $V_0^T(p_{1,B}, n) = p(\delta R(n+k) - k) + (1-p)(\delta qr(n+k) - k)$

4. If $p\delta R + (1-p)\delta qr - 1 - (1-p)\frac{\delta - p_{1,B}}{p_{1,B}} > 0$, and $p_{1,B} < \delta$, and $p\delta R + (1-p)\delta qr - 1 - (1-p)\frac{\delta - (1-\varepsilon)p_{1,B}}{(1-\varepsilon)p_{1,B}} = 0$, any $D_0^T \in [k; \frac{k\left(1 - \frac{p_{1,B}(1-\varepsilon)}{\delta}\right) + (1-\varepsilon)p_{1,B}qrn}{1 - p_{1,B}(1-\varepsilon)qr}]$ is an optimum debt level and $V_0^T(p_{1,B}, n) = p(\delta R(n+k) - k) + (1-p)(\delta qr(n+k) - k)$

5. If $p\delta R + (1-p)\delta qr - 1 - (1-p)\frac{\delta - (1-\varepsilon)p_{1,B}}{(1-\varepsilon)p_{1,B}} > 0$, and $p_{1,B} < \delta$, and $D_0^T = \frac{k\left(1 - \frac{p_{1,B}(1-\varepsilon)}{\delta}\right) + (1-\varepsilon)p_{1,B}qrn}{1 - p_{1,B}(1-\varepsilon)qr}$, $V_0^T(p_{1,B}, n) = p\left(\delta R\left(D_0^T + n\right) - D_0^T\right)$

Case $\delta qrn < k \leq \frac{\delta qrn}{1-\delta qr}$ In this case the three sub-programs admit a non-empty set of solutions. The solution to the program is identical to the case above.

Case $k > \frac{\delta qrn}{1-\delta qr}$ In this case, the last two programs always admit empty set of solutions. Indeed $(k > \delta qr(n+k))$. We end up in the first program, and:

1. If $p\delta R + (1-p)\delta qr - 1 - (1-p)\frac{\delta - p_{1,B}}{p_{1,B}} < 0$, $D_0^T = 0$, $V_0^T(p_{1,B}, n) = p(\delta Rn) + (1-p)\left(\delta qrn - k + \frac{\delta}{p_{1,B}}k\right)$

2. If $p\delta R + (1-p)\delta qr - 1 - (1-p)\frac{\delta - p_{1,B}}{p_{1,B}} = 0$, any $D_0^T \in [0; \frac{k\left(1 - \frac{p_{1,B}}{\delta}\right) + p_{1,B}qrn}{1 - p_{1,B}qr}]$ is optimal and $V_0^T(p_{1,B}, n) = p(\delta R(n+k) - k) + (1-p)(\delta qr(n+k) - k)$

3. If $p\delta R + (1-p)\delta qr - 1 - (1-p)\frac{\delta - p_{1,B}}{p_{1,B}} > 0$, and $p_{1,B} < \delta$, $D_0^T = \frac{k\left(1 - \frac{p_{1,B}}{\delta}\right) + p_{1,B}qrn}{1 - p_{1,B}qr}$,and $V_0^T(p_{1,B}, n) = p\left(\delta R\left(D_0^T + n\right) - D_0^T\right)$

Case $\delta \leq p_{1,B} \leq \frac{\delta}{(1-\varepsilon)}$

Let's now focus on the case where $\delta \leq p_{1,B} \leq \frac{\delta}{(1-\varepsilon)}$.

In this case, $\overline{D}_0^T\left(I_0^T, p_{1,B}\right) = k + \left(\delta qr I_0^T - k\right)_- + \frac{p_{1,B}(1-\varepsilon)}{\delta}\left(\delta qr I_0^T - k\right)_+$

$$V_{1,B}^{T,ND}\left(I_0^T, D_0^T, p_{1,B}\right) = \left(\left(k - D_0^T\right)_- \frac{\left(\delta - p_{1,B}\left(1-\varepsilon\right)\right)_+}{p_{1,B}\left(1-\varepsilon\right)}\right) + \delta qr I_0^T - D_0^T$$

The program rewrites again in three subprograms.

First subprogram We first focus on the subprogram where we additionally constrain I_0^T to evolve in the set where $\delta qr I_0^T \leq k$. In this case

$$\begin{aligned} V_0^T(p_{1,B}, n) &= \max_{D_0^T \geq 0} p\left(\delta R I_0^T - D_0^T\right) + (1-p)\left(\delta qr I_0^T - D_0^T\right) \\ D_0^T &\leq \delta qr I_0^T \\ \delta qr I_0^T &\leq k \\ D_0^T + n &= I_0^T \end{aligned}$$

This program has a non-empty set of solution when $q\delta rn \leq k$, in which case $D_0^T = \min\left(\frac{k-q\delta rn}{q\delta r}; \frac{\delta qrn}{1-\delta qr}\right)$

Second subprogram Let's now turn to the subprogram where we additionally constrain D_0^T and I_0^T to evolve in the set where $\delta qr I_0^T \geq k$ and $D_0^T \leq k$.

In this case, the program writes

$$\begin{aligned} V_0^T(p_{1,B}, n) &= \max_{D_0^T \geq 0}\ p\left(\delta R I_0^T - D_0^T\right) + (1-p)\left(\delta qr I_0^T - D_0^T\right) \\ D_0^T &\leq k \\ \delta qr I_0^T &\geq k \\ D_0^T + n &= I_0^T \end{aligned}$$

This program has a non-empty set of solutions if and only if $\frac{k-q\delta rn}{q\delta r} \leq k$, in which case $D_0^T = k$.

Third subprogram Finally, in the last subprogram we additionally constrain D_0^T and I_0^T to evolve in the set where $\delta qrI_0^T \geq k$ and $D_0^T \geq k$.

In this case,

$$
\begin{aligned}
V_0^T(p_{1,B},n) = &\max_{D_0^T \geq 0} p\left(\delta RI_0^T - D_0^T\right) + (1-p)\left(\delta qrI_0^T - k + \frac{\delta}{p_{1,B}(1-\varepsilon)}\left(k - D_0^T\right)\right)\\
&k \leq D_0^T \leq p_{1,B}(1-\varepsilon)\left(qrI_0^T - \frac{k}{\delta}\right) + k\\
&\delta qrI_0^T \geq k\\
&D_0^T + n = I_0^T
\end{aligned}
$$

Again, this program has a non-empty set of solutions if and only if $k \leq k^*$.

In this range

1. Either $p\delta R + (1-p)\delta qr - 1 - (1-p)\frac{\delta-(1-\varepsilon)p_{1,B}}{p_{1,B(1-\varepsilon)}} < 0$, $D_0^T = \max(k; \frac{k-\delta qrn}{\delta qr}) = k$, and $V_0^T(p_{1,B},n) = \delta(pR + (1-p)qr)(n+k) - k$

2. Or $p\delta R + (1-p)\delta qr - 1 - (1-p)\frac{\delta-(1-\varepsilon)p_{1,B}}{(1-\varepsilon)p_{1,B}} = 0$ and any $D_0 \in [\max(k; \frac{k-\delta qrn}{\delta qr}); \frac{k\left(1-\frac{p_{1,B}(1-\varepsilon)}{\delta}\right)+(1-\varepsilon)p_{1,B}qrn}{1-p_{1,B}(1-\varepsilon)qr}] = [k; \frac{k\left(1-\frac{p_{1,B}(1-\varepsilon)}{\delta}\right)+(1-\varepsilon)p_{1,B}qrn}{1-p_{1,B}(1-\varepsilon)qr}]$ is an equilibrium, and $V_0^T(p_{1,B},n) = \delta(pR + (1-p)qr)(n+k) - k$

3. Or $p\delta R + (1-p)\delta qr - 1 - (1-p)\frac{\delta-(1-\varepsilon)p_{1,B}}{(1-\varepsilon)p_{1,B}} > 0$ and $D_0^T = \frac{k\left(1-\frac{p_{1,B}(1-\varepsilon)}{\delta}\right)+(1-\varepsilon)p_{1,B}qrn}{1-p_{1,B}(1-\varepsilon)qr}$. In this case, $V_0^T(p_{1,B},n) = p\left(\delta R(D_0^T + n) - D_0^T\right)$

We can now solve for the time 0 bank program.

Case $\delta qrn \geq k$ In this case, as before, we necessarily end up in the last two programs.

Hence

1. If $p\delta R+(1-p)\delta qr-1-(1-p)\frac{\delta-(1-\varepsilon)p_{1,B}}{p_{1,B(1-\varepsilon)}}<0$, $D_0^T=k$,and $V_0^T(p_{1,B},n)=p(\delta R(n+k)-k)+(1-p)(\delta qr(n+k)-k)$

2. If $p\delta R+(1-p)\delta qr-1-(1-p)\frac{\delta-(1-\varepsilon)p_{1,B}}{p_{1,B(1-\varepsilon)}}=0$, any $D_0^T\in[k;\frac{k\left(1-\frac{p_{1,B}(1-\varepsilon)}{\delta}\right)+(1-\varepsilon)p_{1,B}qrn}{1-p_{1,B}(1-\varepsilon)qr}]$ is an optimum and $V_0^T(p_{1,B},n)=p(\delta R(n+k)-k)+(1-p)(\delta qr(n+k)-k)$

3. If $p\delta R+(1-p)\delta qr-1-(1-p)\frac{\delta-(1-\varepsilon)p_{1,B}}{(1-\varepsilon)p_{1,B}}>0$, $D_0^T=\frac{k\left(1-\frac{p_{1,B}(1-\varepsilon)}{\delta}\right)+(1-\varepsilon)p_{1,B}qrn}{1-p_{1,B}(1-\varepsilon)qr}$, $V_0^T(p_{1,B},n)=p(\delta R(D_0+n)-D_0)$

Case $\delta qrn<k\leq\frac{\delta qrn}{1-\delta qr}$ The three subprograms have a non-empty set of solutions, and the equilibria of the T-bank at time 0 are the same as in the case above. Indeed, on this range, the value function of the first subprogram is weakly dominated by the one of the second subprogram (strictly if $k<k^*$), and we can disregard it and perform the same analysis as in the case above.

Case $k>\frac{\delta qrn}{1-\delta qr}$ Here, the only subprogram with non-empty solution set is the first one. In this case the optimum is such that $D_0^T=\frac{\delta qrn}{1-\delta qr}$, $V_0^T(p_{1,B},n)=p(\delta RI-D_0)+(1-p)\left(\delta qrI-D_0^R\right)$.

Case $\frac{\delta}{(1-\varepsilon)}\leq p_{1,B}<\frac{1}{(1-\varepsilon)qr}$

Here, the program rewrites

$$V_0^T(p_{1,B},n)=\max_{D_0^T\geq0}p\left(\delta RI_0^T-D_0^T\right)+(1-p)\left(p_{1,B}(1-\varepsilon)qrI_0^T-D_0^T\right)$$

$$D_0^T\leq p_{1,B}(1-\varepsilon)qrI_0^T$$

$$D_0^T+n=I_0^T$$

And $D_0^T=\frac{p_{1,B}(1-\varepsilon)qrn}{1-p_{1,B}((1-\varepsilon)qr)}$ and $V_0^T(p_{1,B},n)=p\left(\delta RI_0^T-D_0^T\right)$

Case $p_{1,B} \geq \frac{1}{(1-\varepsilon)qr}$

At last, if $p_{1,B} \geq \frac{1}{(1-\varepsilon)qr}$, $D_0^T = +\infty$ and $V_0^T(p_{1,B}, n) = +\infty$

□

5.D Market equilibria

Taking all the above elements together, we obtain the following

If $k \leq k^*$

1. If $p_{1,B} = 0$, $V_0^T(p_{1,B}, n) = +\infty$,
2. If $0 < p_{1,B} < p_{1,L}^T$, $D_0^T = 0$, $V_0^T(p_{1,B}, n) = p(\delta Rn) + (1-p)\left(\delta qrn - k + \frac{\delta}{p_{1,B}}k\right)$, $I_{1,B}^T = \frac{k}{p_{1,B}qr}, \alpha_{1,B}^T = 1$
3. If $p_{1,B} = p_{1,L}^T$, any $D_0^T \in [0; k]$ is an equilibrium debt level, and $V_0^T(p_{1,B}, n) = p(\delta R(n+k) - k) + (1-p)(\delta qr(n+k) - k)$,$I_{1,B}^T = \frac{k-D_0^T}{p_{1,B}qr} \in [0; \frac{k}{p_{1,B}qr}], \alpha_{1,B}^T = 1$
4. If $p_{1,L}^T < p_{1,B} < p_{1,H}^T$, $D_0^T = k$,and $V_0^T(p_{1,B}, n) = p(\delta R(n+k) - k) + (1-p)(\delta qr(n+k) - k)$,$I_{1,B}^T = 0, \alpha_{1,B}^T = 1$
5. If $p_{1,B} = p_{1,H}^T$, any $D_0^T \in [k; \frac{k\left(1-\frac{p_{1,B}(1-\varepsilon)}{\delta}\right)+(1-\varepsilon)p_{1,B}qrn}{1-p_{1,B}(1-\varepsilon)qr}]$ is an optimal debt level, $V_0^T(p_{1,B}, n) = p(\delta R(n+k) - k) + (1-p)(\delta qr(n+k) - k)$, $I_{1,B}^T = 0, (1-\alpha_{1,B}^T) = \frac{D_0^T - k}{p_{1,B}qrI_0^T(1-\varepsilon)}$ hence $\alpha_{1,B}^T \in [1 - \frac{qr(n+k)-\frac{k}{\delta}}{qr\left(k\left(1-\frac{p_{1,B}(1-\varepsilon)}{\delta}\right)+n\right)}; 1]$,
6. If $p_{1,H}^T < p_{1,B} < \frac{\delta}{1-\varepsilon}$, $D_0^T = \frac{k\left(1-\frac{p_{1,B}(1-\varepsilon)}{\delta}\right)+(1-\varepsilon)p_{1,B}qrn}{1-p_{1,B}(1-\varepsilon)qr}$, $V_0^T(p_{1,B}, n) = p(\delta R(D_0+n) - D_0)$, $I_{1,B}^T = 0$, $\alpha_{1,B}^T = 1 - \frac{qr(n+k)-\frac{k}{\delta}}{qr\left(k\left(1-\frac{p_{1,B}(1-\varepsilon)}{\delta}\right)+n\right)}$
7. If $p_{1,B} = \frac{\delta}{1-\varepsilon}$, $D_0^T = \frac{(1-\varepsilon)p_{1,B}qrn}{1-p_{1,B}(1-\varepsilon)qr}$, $V_0^T(p_{1,B}, n) = p(\delta R(D_0+n) - D_0)$, $I_{1,B}^T = 0$, $\alpha_{1,B}^T \in \left[0; \frac{k(1-\delta qr)}{qrn}\right]$
8. If $p_{1,B} > \frac{\delta}{1-\varepsilon}$, $D_0^T = \frac{p_{1,B}(1-\varepsilon)qrn}{(1-p_{1,B}(1-\varepsilon)qr)_+}$, $V_0^T(p_{1,B}, n) = p\left(\delta R\left(D_0^T+n\right) - D_0^T\right)$, $I_{1,B}^T = 0$, $\alpha_{1,B}^T = 0$

The net aggregate demand of assets by T-banks on the time 1, G asset market is

$$D^{TB}(p_{1,B}) = \frac{k - D_0^T(p_{1,B}, n)}{p_{1,B}qr} \times (1 - \chi_S). \tag{5.28}$$

This is strictly negative when $p_{1,B} > p_{1,H}^T$. As S-banks net supply is always positive, no market clearing can occur for such price levels.

For $0 \leq p_{1,B} \leq p_{1,H}^T$, the aggregate demand for capital goods is

$$D(p_{1,B}) = \begin{cases} \frac{k}{p_{1,B}qr}(1 - \chi^S) & \text{if } 0 \leq p_{1,B} < \overline{p}_{1,L}^T \\ \in \left[0; \frac{k}{p_{1,B}qr}(1 - \chi^S)\right] & \text{if } p_{1,B} = \overline{p}_{1,L}^T \\ 0 & \text{if } p_{1,B} > \overline{p}_{1,L}^T \end{cases}$$

while the aggregate supply of capital goods is:

$$S(p_{1,B}) = \begin{cases} 0 & \text{if } 0 \leq p_{1,B} < \overline{p}_1^S \\ \in \left[0; \frac{n(1-\varepsilon)}{1-(1-\varepsilon)p_{1,B}qr}\chi^S\right] & \text{if } p_{1,B} = \overline{p}_1^S \\ \frac{n(1-\varepsilon)}{1-(1-\varepsilon)p_{1,B}qr}\chi^S & \text{if } \overline{p}_1^S < p_{1,B} < \overline{p}_{1,H}^T \\ \in \left[\frac{n(1-\varepsilon)}{1-(1-\varepsilon)p_{1,B}qr}\chi^S; \frac{n(1-\varepsilon)}{1-(1-\varepsilon)p_{1,B}qr}\chi^S + \frac{(1-\varepsilon)}{qr}\frac{\left(-\frac{k}{\delta}\right)+qr(n+k)}{1-p_{1,B}(1-\varepsilon)qr}\left(1 - \chi^S\right)\right] & \text{if } p_{1,B} = \overline{p}_{1,H}^T \end{cases}$$

According to the value of χ_S the market clears at different values, detailed in proposition 1.

Bibliography

[1] Puriya Abbassi, Rajkamal Iyer, José-Luis Peydró, and Francesc R Tous. Securities trading by banks and credit supply: Micro-evidence. *Journal of Financial Economics*, 2015.

[2] Viral V. Acharya and Nada Mora. A Crisis of Banks as Liquidity Providers. *Journal of Finance*, 70 (1):1–43, 02 2015.

[3] Viral V. Acharya, Denis Gromb, and Tanju Yorulmazer. Imperfect Competition in the Interbank Market for Liquidity as a Rationale for Central Banking. *American Economic Journal: Macroeconomics*, 4(2):184–217, April 2012.

[4] Viral V. Acharya, Philipp Schnabl, and Gustavo Suarez. Securitization without risk transfer. *Journal of Financial Economics*, 107(3):515–536, 2013.

[5] Tobias Adrian and Nina Boyarchenko. Intermediary balance sheets. *Society for Economic Dynamics*, 2015 Meeting Papers, 2015.

[6] Tobias Adrian and Hyun Song Shin. Liquidity and leverage. *Journal of Financial Intermediation*, 19 (3):418–437, July 2010.

[7] Andrew Ang, Sergiy Gorovyy, and Gregory B. van Inwegen. Hedge fund leverage. *Journal of Financial Economics*, 102(1):102–126, October 2011.

[8] Matthew Baron. Countercyclical bank equity issuance. 2016.

[9] Saki Bigio, Juliane Begenau, and Jeremy Majerovitz. What can we learn from the financial flows of the 2008-2009 crisis? *mimeo, Harvard University*, 2016.

[10] John H. Boyd and Mark Gertler. Are banks dead? Or are the reports greatly exaggerated? *Quarterly Review*, (Sum):2–23, 1994. URL https://ideas.repec.org/a/fip/fedmqr/y1994isump2-23nv.18no.3.html.

[11] Markus K. Brunnermeier and Yuliy Sannikov. A Macroeconomic Model with a Financial Sector. *American Economic Review*, 104(2):379–421, February 2014.

[12] Ricardo J Caballero and Emmanuel Farhi. The Safety Trap. Working Paper 233766, Harvard University OpenScholar, January 2016. URL https://ideas.repec.org/p/qsh/wpaper/233766.html.

[13] Ricardo J. Caballero and Arvind Krishnamurthy. Collective Risk Management in a Flight to Quality Episode. *Journal of Finance*, 63(5):2195–2230, October 2008. URL https://ideas.repec.org/a/bla/jfinan/v63y2008i5p2195-2230.html.

[14] Sergey Chernenko, Samuel G. Hanson, and Adi Sunderam. The Rise and Fall of Demand for Securitizations. NBER Working Papers 20777, National Bureau of Economic Research, Inc, December 2014.

[15] Adam Copeland, Antoine Martin, and Michael Walker. Repo Runs: Evidence from the Tri-Party Repo Market. *Journal of Finance*, 69(6):2343–2380, December 2014. URL https://ideas.repec.org/a/bla/jfinan/v69y2014i6p2343-2380.html.

[16] Harry DeAngelo and René M. Stulz. Liquid-claim production, risk management, and bank capital structure: Why high leverage is optimal for banks. *Journal of Financial Economics*, 116(2):219–236, 2015.

[17] Douglas W Diamond and Philip H Dybvig. Bank runs, deposit insurance, and liquidity. *Journal of Political Economy*, 91(3):401–19, June 1983.

[18] Douglas W. Diamond and Raghuram G. Rajan. Fear of fire sales, illiquidity seeking, and credit freezes. *The Quarterly Journal of Economics*, 126(2):557–591, 2011.

[19] European Systemic Risk Board ESRB. Eu shadow banking eu shadow banking monitor. *ESRB Reports*, 1, July 2016.

[20] Financial Stability Board FSB. Global shadow banking monitoring report 2015. *Technical Report*, November 2015.

[21] Joseph E. Gagnon, Matthew Raskin, Julie Remache, and Brian P. Sack. Large-scale asset purchases by the Federal Reserve: did they work? *Economic Policy Review*, (May):41–59, 2011.

[22] Evan Gatev and Philip E. Strahan. Banks' Advantage in Hedging Liquidity Risk: Theory and Evidence from the Commercial Paper Market. *Journal of Finance*, 61(2):867–892, 04 2006.

[23] Nicola Gennaioli, Andrei Shleifer, and Robert W. Vishny. A Model of Shadow Banking. *Journal of Finance*, 68(4):1331–1363, 08 2013.

[24] Lucyna Gornicka. Banks and shadow banks: Competitors or complements? *Journal of Financial Intermediation*, forthcoming, 2016.

[25] Gary Gorton and Andrew Metrick. Securitized banking and the run on repo. *Journal of Financial Economics*, 104(3):425–451, 2011.

[26] Gary Gorton and George Pennacchi. Financial Intermediaries and Liquidity Creation. *Journal of Finance*, 45(1):49–71, March 1990.

[27] Denis Gromb and Dimitri Vayanos. Equilibrium and welfare in markets with financially constrained arbitrageurs. *Journal of Financial Economics*, 66(2-3):361–407, 2002.

[28] Samuel G. Hanson, Andrei Shleifer, Jeremy C. Stein, and Robert W. Vishny. Banks as patient fixed-income investors. *Journal of Financial Economics*, 117(3):449–469, 2015.

[29] Milton Harris, Christian C. Opp, and Marcus M. Opp. The system-wide effects of bank capital regulation on credit supply and risk-taking. *mimeo*, 2015.

[30] Zhiguo He and Arvind Krishnamurthy. Intermediary Asset Pricing. *American Economic Review*, 103 (2):732–70, April 2013.

[31] Zhiguo He, In Gu Khang, and Arvind Krishnamurthy. Balance sheet adjustments during the 2008 crisis. *IMF Economic Review*, 58(1):118–156, August 2010.

[32] IMF. Shadow banking around the globe: How large, and how risky? Global financial stability report, International Monetary Fund, Oct 2014.

[33] Christopher James and Joel Houston. Evolution Or Extinction: Where Are Banks Headed? *Journal of Applied Corporate Finance*, 9(2):8–23, 1996. URL `https://ideas.repec.org/a/bla/jacrfn/v9y1996i2p8-23.html`.

[34] Anil K. Kashyap, Raghuram Rajan, and Jeremy C. Stein. Banks as Liquidity Providers: An Explanation for the Coexistence of Lending and Deposit-Taking. *Journal of Finance*, 57(1):33–73, 02 2002.

[35] Arvind Krishnamurthy. Fundamental value and limits to arbitrage. December 4 2008.

[36] Arvind Krishnamurthy. Amplification Mechanisms in Liquidity Crises. *American Economic Journal: Macroeconomics*, 2(3):1–30, July 2010. URL `https://ideas.repec.org/a/aea/aejmac/v2y2010i3p1-30.html`.

[37] Arvind Krishnamurthy and Annette Vissing-Jorgensen. The impact of Treasury supply on financial sector lending and stability. *Journal of Financial Economics*, 118(3):571–600, 2015.

[38] Stephan Luck and Paul Schempp. Banks, shadow banking, and fragility. Working Paper Series 1726, European Central Bank, August 2014. URL https://ideas.repec.org/p/ecb/ecbwps/20141726.html.

[39] Craig B. Merrill, Taylor Nadauld, Rene M. Stulz, and Shane M. Sherlund. Did Capital Requirements and Fair Value Accounting Spark Fire Sales in Distressed Mortgage-Backed Securities? Working Paper Series 2012-12, Ohio State University, Charles A. Dice Center for Research in Financial Economics, July 2012.

[40] Robert C. Merton. Financial innovation and the management and regulation of financial institutions. *Journal of Banking & Finance*, 19(3-4):461–481, June 1995. URL https://ideas.repec.org/a/eee/jbfina/v19y1995i3-4p461-481.html.

[41] Guillermo Ordonez. Sustainable Shadow Banking. NBER Working Papers 19022, National Bureau of Economic Research, Inc, May 2013.

[42] Guillaume Plantin. Shadow Banking and Bank Capital Regulation. *Review of Financial Studies*, 28 (1):146–175, 2015.

[43] Zoltan Pozsar, Tobias Adrian, Adam B. Ashcraft, and Hayley Boesky. Shadow banking. *Economic Policy Review*, (Dec):1–16, 2013.

[44] Raghuram G. Rajan. Do we still need commercial banks? 1998.

[45] Raghuram G Rajan. The Past and Future of Commercial Banking Viewed through an Incomplete Contract Lens. *Journal of Money, Credit and Banking*, 30(3):524–50, August 1998. URL https://ideas.repec.org/a/mcb/jmoncb/v30y1998i3p524-50.html.

[46] Hyun Song Shin. *Risk and Liquidity*. Number 9780199546367 in OUP Catalogue. Oxford University Press, March 2010.

[47] Andrei Shleifer and Robert W Vishny. Liquidation values and debt capacity: A market equilibrium approach. *Journal of Finance*, 47(4):1343–66, September 1992.

[48] Andrei Shleifer and Robert W Vishny. The limits of arbitrage. *Journal of Finance*, 52(1):35–55, March 1997.

[49] Andrei Shleifer and Robert W. Vishny. Unstable banking. *Journal of Financial Economics*, 97(3): 306–318, September 2010.

[50] Jeremy C. Stein. Monetary Policy as Financial Stability Regulation. *The Quarterly Journal of Economics*, 127(1):57–95, 2012.

[51] Jeremy C. Stein. The fire-sales problem and securities financing the fire-sales problem and securities financing transactions. *Speech by Governor Jeremy C. Stein at the Federal Reserve Bank of New York during the Workshop on Fire Sales as a Driver of Systemic Risk in Triparty Repo and other Secured Funding Markets, New York.*, October 2013.

[52] S. Viswanathan and Adriano A. Rampini. Financial intermediary capital. *Society for Economic Dynamics*, 2010 Meeting Papers, 2010.

Chapter

6 Conclusion

This last chapter offers conclusive remarks about this dissertation, and opens towards new paths of research that I am currently exploring

.

6.1 Conclusive remarks

This dissertation has focused on three distinct works, dealing with theoretically oriented financial economics question.

The first one has emphasized the importance of order choices in the amount of information embedded in prices, in a large market with privately informed agents, showing how the choice to protect oneself against price execution risk might ultimately not be the most frequent one when cheaper market orders exist in the market.

Taking this insight into a two-stage global game, the second work shows how self-fulfiliing equilibria might disappear when frictions impact the trading of an asset whose price is used as a coordinating device for a speculative attack (bank runs, currency pegs, etc.). An interesting extension of this work would be to consider empirical questions related to this point, be it regarding currency attacks or bank runs. The link between the informationnal content of the price of bailinable securities or currencies and the risks of self-fulfilling crises could be further analysed with this in mind.

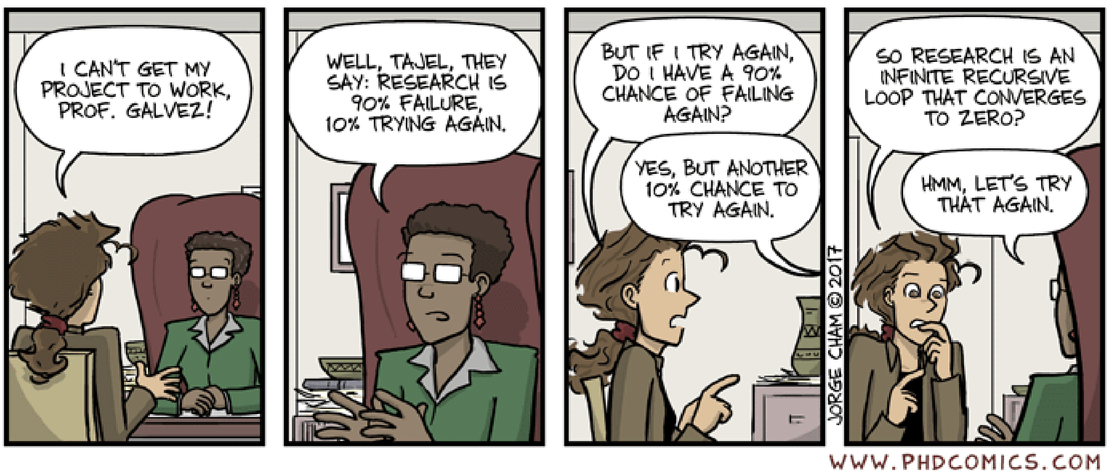
Finally, the last work analyses the structure of the financial intermediation system, and the allocations of agents between traditional and shadow banks, when considering the way they interact in times of crises. A deeper emphasis should be later put onto the normative part of the analysis. Among others, one interesting question that need to be deepened is the question of the optimal level of support the government should be providing to the traditional banking sector, assuming some functionnal form between the size of the support and the cost of managing a traditional bank. This is an ongoing work.

6.2 Future research

On top of the aforementioned extensions and improvements of the chapters developped in this dissertation, I have been recently pursuing different projects in related fields, which are still at their initial stages of development. I will briefly expose the most advanced one

A project currently under development with Rajkamal Iyer and Victor Lyonnet is seeking to bring additional insights on the reasons why people choose to allocate such a high share of their financial assets into deposits. To give an example, the second wave of the Household Finance and Consumption Survey, led on a European level, shows that 44.2% of households financial assets are kept in the form of deposits, be it savings deposits or current accounts. These types of assets are very safe and very liquid, and a need for such a high amount of demandable assets, seems hard to reconcile with a willingness to hedge against liquidity needs. The aim of this project is twofold. On the one hand, an experimental work will be done to try and deepen the understanding of the determinants of the demand for this demandability feature. A second part of this project is more theoretically oriented, and takes a financial intermediation perspective, proposing a supply driven theory for this holding of demandable claims by households. The study of this theoretical approach will have regulatory implications. This project has benefited from a grant of the Banque de France Foundation, to which we are grateful.

Together with this project, others are in development, at different levels of progress. In these, I explore with coauthors both theoretical and empirical aspects, ranging from questions of shadow banking regulation, to the impact of notation on bank lending and real firms outcomes, or the potential governance-induced biases in bank lending. Developing these ideas constitute my current research agenda.


I CAN'T GET MY PROJECT TO WORK, PROF. GALVEZ!
WELL, TAJEL, THEY SAY: RESEARCH IS 90% FAILURE, 10% TRYING AGAIN.
BUT IF I TRY AGAIN, DO I HAVE A 90% CHANCE OF FAILING AGAIN?
YES, BUT ANOTHER 10% CHANCE TO TRY AGAIN.
JORGE CHAM © 2017
SO RESEARCH IS AN INFINITE RECURSIVE LOOP THAT CONVERGES TO ZERO?
HMM, LET'S TRY THAT AGAIN.
WWW.PHDCOMICS.COM

NEC virtute foret clarisve potentius armis
Quam lingua Latium, si non offenderet unum
Quemque poetarum limae labor et mora. Vos, o
Pompilius sanguis, carmen reprehendite quod non
Multa dies et multa litura coercuit atque
Praesectum deciens non castigauit ad unguem.

Ars Poetica , Horatius

LE Latium ne serait pas moins puissant par la langue
que par le courage et ses armes glorieuses, si la lenteur
et le travail de la lime ne rebutaient la plupart de nos poètes.
Ô vous, sang de Pompilius, blâmez le poème qui n'a point été épuré
par de nombreux jours de travail et par de nombreuses ratures,
auquel dix corrections n'ont point donné le poli de l'ongle.

L'art poétique , Horace

AND Latium would be no less supreme in letters
Than in courage and force of arms, if all her poets
Weren't deterred by revision's time and effort.
O scions of Numa, condemn that work that many
A day, and many erasures, have not corrected,
Improving it ten times over, smoothed to the touch.

The Art of Poetry , Horace

Thèse pour l'obtention du grade de docteur de l'Université Paris-Saclay en Sciences Économiques, soutenue le 23 Mai 2017 par Edouard Chrétien. Directeur de thèse : Édouard Challe. Rapporteurs: Guillaume Plantin et Andrei Shleifer

Titre : Essais en Économie Financière

Résumé : Cette thèse est composée de trois chapitres distincts. Dans le premier chapitre, coécrit avec Edouard Challe, nous analysons la détermination jointe de l'information incorporée dans les prix, et la composition du marché par type d'ordres sur un marché d'actifs avec information dispersée. La microstructure du marché est telle que les agents informés peuvent placer soit des ordres de marché simples, soit un ensemble d'ordres limites. Les market-makers établissent le prix. Les agents utilisant des ordres de marché simple négocient moins agressivement sur leur information et réduisent ainsi le contenu informationnel du prix; dans un marché où seul ce type d'ordre est présent, l'information incorporée dans le prix est limitée, quelle que soit la qualité de l'information des agents sur le dividende de l'actif. Lorsque les agents peuvent choisir leur type d'ordre et les ordres limites sont plus coûteux que les ordres de marché, alors les agents choisissent majoritairement les ordres de marché lorsque la précision des signaux privés tend vers l'infini. Les ordres limites sont des substituts: à des niveaux élevés de précision, une fraction résiduelle d'agents plaçant des ordres limites est suffisante pour aligner le prix aux signaux des agents, et donc au dividende. Ainsi le gain à conditionner ses ordres au prix (via des ordres limites) en plus de son propre signal (comme le font tous les agents) disparaît. Nous appliquons ensuite ce mécanisme dans le deuxième chapitre de cette thèse. Les spéculateurs envisageant une attaque (comme dans le cas des crises de change) doivent deviner les croyances des autres spéculateurs, ce qu'ils peuvent faire en regardant le marché boursier. Ce chapitre examine si ce processus de collecte d'informations est stabilisateur, en ancrant mieux les attentes ou déstabilisateur en générant des équilibres multiples. Pour ce faire, nous étudions les résultats d'un jeu global en deux étapes où un prix d'actif déterminé au stade de négociation du jeu fournit un signal public endogène sur le fondamental qui affecte la décision des agents d'attaquer dans la phase de coordination du jeu. La microstructure du marché d'actif reprend celle étudiée dans le premier chapitre. Les frictions de microstructure qui conduisent à une plus grande exposition individuelle (au risque d'exécution des prix) peuvent réduire l'incertitude agrégée (en fixant un résultat d'équilibre unique). Enfin, dans le troisième chapitre, en collaboration avec Victor Lyonnet, nous présentons un modèle des interactions entre les banques traditionnelles et les shadow banks qui parle de leur coexistence. Au cours de la crise financière de 2007, certains actifs et passifs des shadow banks sont passées aux banques traditionnelles et les actifs ont été vendus à des prix de fire sale. Notre modèle réplique ces faits stylisés. La différence entre les banques traditionnelles et les shadow banks est double. Premièrement, les banques traditionnelles ont accès à un fonds de garantie qui leur permet de se financer sans risque en période de crise. Deuxièmement, les banques traditionnelles doivent respecter une réglementation coûteuse. Nous montrons qu'en cas de crise, les shadow banks liquident les actifs pour rembourser leurs créanciers, alors que les banques traditionnelles achètent ces actifs à des prix de fire sale. Cet échange d'actifs en temps de crise génère une complémentarité entre les banques traditionnelles et les shadow banks, où chaque type d'intermédiaire profite de la présence de l'autre. Nous constatons deux effets concurrents d'une petite diminution du soutien des banques traditionnelles en période de crise, que nous appelons effet de substitution et effet de revenu. Ce dernier effet domine le premier, de sorte qu'un niveau de soutien anticipé plus faible aux banques traditionnelles en temps de crise induit plus de banquiers à s'orienter vers le secteur traditionnel ex-ante.

Mots clés : Microstructure de marché, Jeux globaux, Théorie de l'intermédiation

Title : Essays in Financial Economics

Summary : This dissertation is made of three distinct chapters. In the first chapter, which is joint with Edouard Challe, we analyse the joint determination of price informativeness and the composition of the market by order type in a large asset market with dispersed information. The market microstructure is one in which informed traders may place market orders or full demand schedules and where market makers set the price. Market-order traders trade less aggressively on their information and thus reduce the informativeness of the price; in a full market-order market, price informativeness is bounded, whatever the quality of traders' information about the asset's dividend. When traders can choose their order type and demand schedules are (even marginally) costlier than market orders, then market-order traders overwhelm the market when the precision of private signals goes to infinity. This is because demand schedules are substitutes: at high levels of precision, a residual fraction of demand-schedule traders is sufficient to take the trading price close to traders' signals, while the latter is itself well aligned with the dividend. Hence, the gain from trading conditional on the price (as demand-schedule traders do) in addition to one's own signal (as all informed traders do) vanishes. We then apply this idea in the second chapter of this dissertation. Speculators contemplating an attack (e.g., on a currency peg) must guess the beliefs of other speculators, which they can do by looking at the stock market. This chapter examines whether this information-gathering process is stabilizing by better anchoring expectations or destabilizing by creating multiple self-fulfilling equilibria. To do so, we study the outcome of a two-stage global game wherein an asset price determined at the trading stage of the game provides an endogenous public signal about the fundamental that affects traders' decision to attack in the coordination stage of the game. The trading stage follows the microstructure of the first chapter. Price execution risk reduces traders' aggressiveness and hence slows down information aggregation, which ultimately makes multiple equilibria in the coordination stage less likely. In this sense, microstructure frictions that lead to greater individual exposure (to price execution risk) may reduce aggregate uncertainty (by pinning down a unique equilibrium outcome). Finally, in the third chapter, joint with Victor Lyonnet, we present a model of the interactions between traditional and shadow banks that speaks to their coexistence. In the 2007 financial crisis, some of shadow banks' assets and liabilities have moved to traditional banks, and assets were sold at fire sale prices. Our model is able to accommodate these stylized facts. The difference between traditional and shadow banks is twofold. First, traditional banks have access to a guarantee fund that enables them to issue claims to households in a crisis. Second, traditional banks have to comply with costly regulation. We show that in a crisis, shadow banks liquidate assets to repay their creditors, while traditional banks purchase these assets at fire-sale prices. This exchange of assets in a crisis generates a complementarity between traditional and shadow banks, where each type of intermediary benefits from the presence of the other. We find two competing effects from a small decrease in traditional banks' support in a crisis, which we dub a substitution effect and an income effect. The latter effect dominates the former, so that lower anticipated support to traditional banks in a crisis induces more bankers to run a traditional bank ex-ante.

Keywords: Market Microstructure, Global Games, Theory of Intermediation

université PARIS-SACLAY | ÉCOLE DOCTORALE Sciences de l'homme et de la société (SHS)

l'X ÉCOLE POLYTECHNIQUE UNIVERSITÉ PARIS-SACLAY