Université Panthéon-Assas

école doctorale d'économie, gestion, information et communication[1] (EGIC), ED 455.

Thèse de doctorat en Sciences économiques soutenue le 12 décembre 2014

Thèse de Doctorat / décembre 2014

Cinq essais dans le domaine monétaire, bancaire et financier

Université Panthéon-Assas

Fabien Mercier

Sous la direction de Marianne Guille

Membres du jury :

Damien Besancenot, professeur à l'Université Paris 13, rapporteur

Michel Boutillier, professeur à l'Université Paris Ouest Nanterre, rapporteur

Sébastien Lotz, professeur à l'Université Panthéon-Assas, suffragant

Radu Vranceanu, professeur à l'ESSEC, suffragant

---

[1] Thèse effectuée au sein du laboratoire d'économie mathématiques et de microéconomie appliquée (LEMMA), EA 4442.

Avertissement

La Faculté n'entend donner aucune approbation ni improbation aux opinions émises dans cette thèse ; ces opinions doivent être considérées comme propres à leur auteur.

Remerciements

Je tiens à remercier particulièrement Marianne Guille pour m'avoir orienté à travers le thème très stimulant et toujours actuel de l'illusion monétaire.

Sommaire

Introduction Générale

Le thème initial de la présente thèse était « Illusion monétaire et prix des actifs », l'illusion monétaire étant entendue comme le biais cognitif introduit par la prise en compte des valeurs nominales plutôt que réelles dans les évaluations et les décisions économiques des agents. Deux autres thèmes sont ensuite venus s'ajouter à la thèse, en raison des opportunités dont j'ai bénéficié à la faveur d'un stage de recherche au sein de l'équipe « Target 2 Securities » que j'ai effectué à la Banque Centrale Européenne (BCE) au cours de la thèse. Le premier de ces thèmes s'intéresse à la modélisation de l'industrie post-négociation et plus particulièrement de TARGET2-Securities (T2S). T2S est un projet de l'Eurosystème pour le règlement des actifs financiers. Le second thème aborde le problème de la modélisation du risque de crédit à partir des modèles de défaut.

Les sections suivantes considèrent chacun de ces trois thèmes tour-à-tour. Pour chaque thème le contexte et la motivation des articles qui s'insèrent dans ces thèmes sont présentés. Puis, les articles se rattachant à chacun de ces thèmes sont résumés.

Les cinq articles écrits dans le cadre de ces trois thèmes constituent la thèse. Ces articles sont listés ci-dessous et regroupés suivant le thème auquel ils se rattachent.

Thème 1 : Illusion monétaire et prix des actifs:

- Article [1] "Money-illusion: French evidence", avec Marianne Guille;

- Article [2] "The validity and time-horizon of the Fed model: a co-integration approach" (soumis au *Journal of Empirical Finance*, Septembre 2014, code manuscrit: EMPFIN-S-14-00287), complété par un document séparé contenant des annexes additionnelles.

Thème 2 : L'industrie des services post-négociation :

- Article [3] "Optimal CSD reshaping towards T2S", avec Stephan Sauer, publié dans ECB *Working Paper Series* no 1549, Mai 2013, et au *Journal of Financial Market Infrastructures*, vol. 2, no. 2, pages 3-51, 2014;

- Article [4] "Endogenizing the network effect: how Monte-Carlo simulations and Graph Theory could benefit economic modelling". Une version révisée de cet article a été soumise en Octobre 2014 à la revue Managerial and Decision Economics sous l'intitulé "Network effects and modelling agents' decision to join a network".

Thème 3 : Modèles de défaut:

- Article [5] "The modelling of default rates through hidden Markov chains models", ECB project paper, Juillet 2013.

Thème 1 : illusion monétaire et prix des actifs

1.1. Contexte et motivation

L'illusion monétaire, entendue ici comme un biais induit par la prise en considération des valeurs nominales plutôt que réelles dans les évaluations et les décisions économiques des agents, est sans doute l'un des thèmes les plus anciens de l'étude de la rationalité limitée des comportements en économie. Si Fisher (Fisher, 1928) tout comme Keynes (Keynes, 1935) lui consacre une place importante dans ses écrits, l'illusion monétaire a cependant été négligée à partir des années 1970, en raison du postulat dominant de rationalité de l'homo oeconomicus. Dans un tel cadre d'analyse le comportement des agents économiques se réduit à la simple maximisation d'une fonction d'utilité sous contrainte budgétaire, ou de leur espérance d'utilité en situation de risque, sachant qu'ils utilisent toute l'information disponible et anticipent correctement les conséquences de leurs décisions puisqu'on suppose depuis Lucas (Lucas, 1996) qu'ils forment des anticipations rationnelles et ne peuvent donc se tromper de manière répétée. Cependant, l'étude des nombreux biais comportementaux qui semblent affecter les agents sur les marchés financiers a favorisé un regain d'intérêt empirique pour l'illusion monétaire depuis les années 1990. Différentes études empiriques à la fois quantitatives et qualitatives mettent en évidence une certaine illusion monétaire des agents dans des contextes variés. Ces études sont confortées par des approches plus comportementales, comme la vaste enquête menée par Shafir, Diamond et Tversky (Shafir, Diamond, & Tversky, 1997) ou les jeux expérimentaux développés par Fehr et Tyran (Fehr & Tyran, 2001) pour comprendre la formation des prix, qui tendent à montrer l'existence d'un biais cognitif fort en faveur des grandeurs nominales au niveau individuel. Les neurosciences commencent à apporter un nouvel éclairage à ces travaux. En particulier, une étude récente, réalisée au niveau infra-personnel dans des conditions strictes par Weber *et al.* (Weber, 2009) utilise l'imagerie par résonance magnétique (IRM) pour mesurer l'activation de certaines zones du cortex préfrontal liées au sentiment de satisfaction de sujets mis face à deux situations équivalentes sur le plan réel, mais différentes sur le plan nominal. Les auteurs montrent que dans la situation où les valeurs nominales sont plus grandes, la satisfaction, du moins telle que mesurée par l'activation de ladite zone du cerveau, est en moyenne 30% plus importante que dans la situation où les valeurs nominales sont plus faibles. Ces faisceaux d'indices tendent à confirmer la prégnance de l'illusion monétaire au niveau individuel et, comme le font remarquer Bourgeois-Gironde et Guille (Bourgeois-Gironde & Guille 2011), ses probables répercussions sur les décisions des agents deviennent de plus en plus difficiles à ignorer pour toute théorie économique voulant rendre compte du réel.

Bien que l'interprétation des résultats des études comportementales à la Shafir, Diamond et Tversky (Shafir, Diamond, & Tversky, 1997) puisse toujours faire l'objet de débat, ce type d'analyse présente certains avantages par rapport aux approches à base d'imagerie par résonance magnétique développées récemment. Tout d'abord, et de manière évidente, un coût plus faible, ce qui permet de réaliser l'étude sur des échantillons plus grands et ne nécessite pas l'accès à un laboratoire de neuroscience et à ses machines. Mais également la possibilité de présenter aux sujets des situations plus complexes, ce qui ne peut pas être fait de manière satisfaisante par l'imagerie cérébrale au niveau actuel de la recherche: les zones activées dans le cortex cérébral par un intitulé trop complexe ne permettraient pas de distinguer, ni d'analyser, la zone précise liée à l'illusion monétaire. Pour ces raisons, nous avons choisi d'adopter une démarche proche de celle de Shafir, Diamond et Tversky (Shafir, Diamond, & Tversky, 1997) pour notre premier article [1], mais s'en distinguant par quatre points principaux.

Le premier point concerne le mode opératoire de l'étude. Shafir *et al.* ont réalisé leur enquête avec des étudiants de Princeton, mais aussi avec des personnes interrogées dans un aéroport et dans des centres commerciaux. Notre démarche est différente puisque nous avons choisi d'adopter un protocole de type expérimental, les questions étant posées aux sujets dans des conditions strictement contrôlées de laboratoire, via une interface machine. L'échantillon est constitué d'étudiants français de niveau master qui suivaient différents parcours: sciences, économie, droit et lettres.

Le second point concerne la présentation du contexte et la formulation des questions posées aux sujets. En effet, le questionnaire de Shafir *et al.* (Shafir, Diamond, & Tversky, 1997) comporte certaines questions dont le contexte parfois imprécis nous semble pouvoir donner lieu à des problèmes d'interprétation. Par exemple, le Problème 1 de leur article précise dans l'énoncé que Barbara et Ann, qui se retrouvent dans des conditions d'inflation et de hausses de salaires différentes, ont fait leurs études ensemble. De ce fait l'énoncé inclut dans le même univers les deux personnages. Dès lors, l'évaluation des salaires réels (relativement à la ville où elles vivent) se double d'une probable évaluation comparative entre les deux situations, évaluation qui n'a rien d'irrationnel : les individus peuvent en effet voyager, épargner... En outre, rien n'indique que le pouvoir d'achat (local) plus fort dont dispose Ann ne rende sa situation objectivement supérieure à celle de Barbara. Nous avons reformulé ce type de question en séparant plus clairement les deux cas, c'est-à-dire en donnant le choix entre deux situations correspondant à deux univers possibles distincts et non plus liés. Considérer, par exemple, comme le font Shafir et al, que, pour le Problème 2, la personne qui aurait fait la meilleure affaire immobilière est celle qui a réalisé le profit "réel" (vis-à-vis de l'inflation locale) le plus élevé, ne va pas forcément de soi. Cette optique suppose en effet que les individus non sujets à l'illusion monétaires sont ceux qui auraient considéré que c'est Adam, qui a acheté sa maison dans un contexte déflationniste et l'a revendu juste un peu plus cher que le prix déflaté, qui a fait la bonne affaire, au lieu de Carl qui l'a acheté dans un contexte inflationniste, revendue au même prix réel c'est-à-dire au prix inflaté. Or les sujets peuvent se dire qu'Adam aurait pu simplement conserver de la monnaie et attendre la chute des prix, ou acheter au même endroit que Carl, même si clairement Carl s'est bien prémuni de l'inflation. Là encore, le contexte, parce qu'il précise qu'il s'agit de trois frères, qui ont chacun acheté la maison après avoir reçu un héritage, en même temps, ramène inutilement la comparaison des investissements dans le même univers, favorisant les comparaisons et arbitrages entre régions dans l'esprit des répondants. Il était donc intéressant de repenser et de reformuler clairement ce type de question, afin de démontrer la robustesse - ou non - des résultats obtenus.

Le troisième point de différence s'attache au fait que l'étude de Shafir et. al. comme la plus grande partie des autres études sur le sujet concerne les Etats-Unis. Réaliser une étude expérimentale sur des étudiants français permet d'étudier d'éventuels biais culturels liés à l'illusion monétaire, mais sert aussi de test de robustesse. Même si nous ne nous attendons pas à des divergences importantes de résultats entre pays, une même tendance de fond éliminerait la possibilité d'attribuer à des facteurs purement culturels la propension à l'illusion monétaire des agents (dans le cas où l'on constaterait effectivement un tel biais).

Le dernier point de différence est l'étude plus approfondie des interactions potentielles entre certaines caractéristiques individuelles, mesurant notamment l'aversion au risque et à la perte, le niveau d'éducation financière, les perceptions de justice ou de satisfaction personnelle vis-à-vis de soi-même ou telle que pressentie chez les autres, avec l'illusion monétaire proprement dite. Inclure dans le questionnaire des questions visant à mesurer ces caractéristiques individuelles nous permet en effet de systématiquement tester l'effet de ces caractéristiques sur l'illusion monétaire. Par exemple, le problème posé aux sujets qui concerne l'immobilier sépare clairement la quantification de chaque scénario comme plus ou moins profitable de la quantification de la satisfaction personnelle que le sujet expérimenterait dans chaque scénario[2].

Les résultats expérimentaux que nous obtenons permettent de confirmer la présence de l'illusion monétaire, et d'étayer, entre autres, les résultats obtenus par Shafir *et al.* (Shafir, Diamond, & Tversky, 1997). La comparaison du prix d'achat nominal au prix de vente nominal semble flatter notre attachement aux chiffres, au détriment de l'appréciation de la valeur réelle d'un bien, qui est ce que nous en faisons, autrement dit à la valeur émotionnelle ou hédonique de la monnaie plutôt qu'instrumentale (Kahneman[3] et Camerer, Loewenstein et Prelec (Camerer, Loewenstein, & Prelec, 2005).

L'illusion monétaire semble même être constatée même chez les agents économiques les plus sophistiqués comme les investisseurs institutionnels ainsi que le souligne notamment une étude d'Asness (Asness, 2003) qui étudie la pertinence du modèle appelé "modèle de la Fed". Ce modèle très utilisé par les professionnels du marché compare le taux nominal des obligations du trésor américain à long terme à une grandeur réelle, le ratio des dividendes sur le prix. Asness montre que ledit modèle explique très bien les mouvements corrélés des deux quantités aux Etats-Unis, c'est-à-dire possède un très grand pouvoir descriptif, alors que son pouvoir de prédiction des gains est très faible, inférieur au seul classique quotient des dividendes sur le prix de l'action. Ainsi, les investisseurs, sur un marché aussi important et professionnel que le marché actions, seraient capables de se tromper, et ce, de manière répétée. La caractéristique de cette erreur est d'ailleurs de nature à ne permettre qu'un arbitrage difficile, car nécessitant une position aux gains et pertes très volatils et devant être tenue très longtemps (Campbell et Vuolteenaho (Campbell & Vuolteenaho, Inflation illusion and stock prices, 2004), p.7). Boucher (Boucher, 2006) confirme ces résultats, en utilisant l'extension dynamique du modèle de Gordon proposée dans les anneés 1990 par Campbell et Shiller (Campbell & Shiller, 1989), laquelle, contrairement au modèle de Gordon permet au taux de croissance des dividendes de varier.

Bien sûr, des explications en accord avec la rationalité des agents ont également été avancées pour expliquer que le ratio du dividende sur le prix de l'action et l'inflation soient historiquement

[2] L'illusion monétaire dans le cadre des achats ou ventes de bien immobilier mériterait selon nous une étude spécifique, étant donné ses conséquences en termes de soutiens publique aux politiques de l'immobilier, dont un des effets pourrait être le maintien des prix à des niveaux qui rendent in fine difficile l'accès à la propriété pour une majeure partie de la population, en soutenant la demande face à une offre qui semble plutôt inélastique. L'absence de tout débat public concernant le bien-fondé de ces politiques de soutien à l'immobilier semble témoigner d'une adhésion populaire dont l'illusion monétaire pourrait être une cause. Bien que la plupart des ménages ne possèdent qu'un ou deux logements, ou n'aient pas envie de vendre ni de spéculer car ils utilisent ce ou ces logements comme un bien de consommation (c'est-à-dire pour leur seul usage personnel), une appréciation est en général bien vécue, même parmi les primo accédants qui, ayant financé leur acquisition par le crédit, jugeraient négativement une éventuelle baisse de l'immobilier, puisqu'ils estimeraient qu'ils auraient alors "surpayé" le bien. En revanche, au niveau d'un panier de bien qui inclut le logement, l'appréciation ou la dépréciation ne change rien au pouvoir d'achat réel : si un ménage vend sa maison il doit souvent en racheter une autre ailleurs pour y vivre, et, en moyenne, son prix aura aussi augmenté ou baissé. Le gain purement nominal relève de l'illusion monétaire.

[3] Kahneman, Maps of Bounded Rationality: Psychology for Behavioral Economics, 2003.

positivement corrélés. En particulier un climat d'inflation plus forte pourrait diminuer la tolérance au risque des agents, lesquels demanderaient alors une rémunération réelle plus élevée pour leur investissement en actions. Dans un récent article publié en 2010, Bekaret et Engstrom (Bekaert & Engstrom, 2010) avancent une raison pour laquelle la prime de risque demandée par les investisseurs actions serait plus importante en période de forte inflation : cela tiendrait au fait qu'historiquement, aux Etats-Unis, l'inflation est corrélée à des périodes de plus grande incertitude macroéconomique. Ces auteurs proposent deux manières de mesurer cette incertitude macroéconomique : une première mesure fondée sur les divergences des prévisions du PIB réel par les professionnels, et une seconde mesure fondée sur un modèle où ce sont les variations des habitudes des consommateurs qui reflètent l'incertitude. Leur étude montre en outre que le degré de corrélation du taux nominal des obligations d'Etat est plus important dans les pays ayant connu des périodes de stagflation plus fréquentes (c'est-à-dire un degré de corrélation de l'inflation et des récessions plus important), ce qui tend à accréditer leur théorie explicative. On peut néanmoins objecter, tout comme Boucher (Boucher, 2006), que cette explication ne s'accompagne pas d'un fondement théorique satisfaisant (il est difficile d'expliquer pourquoi une inflation plus élevée rendrait les investisseurs en actions plus adverses au risque sur le long terme). En outre, les résultats de Campbel et Vuolteenaho concernant ladite prime de risque semblent contredire cette explication. En effet, Campbel et Vuolteenaho (Campbell & Vuolteenaho, Inflation illusion and stock prices, 2004) supposent que les agents utilisent le CAPM pour évaluer le risque des valeurs boursières, ce qui leur permet d'estimer la prime de risque dans divers scénarios d'inflation, puis de décomposer le ratio dividende sur prix en trois composantes dont la prime de risque subjective. Ils constatent une insensibilité importante de cette dernière vis-à-vis de l'inflation, et trouvent que l'inflation est au contraire très corrélée à la composante de "mispricing" de leur modèle, confirmant une illusion monétaire dans l'évaluation du prix des actifs. Ainsi ces auteurs semblent, tout comme Asness (Asness, 2003), confirmer la validité de l'hypothèse de Modigliani et Cohn (Modigliani & Cohn, 1979) selon laquelle les investisseurs du marché actions seraient sujets à l'illusion monétaire, et utiliseraient des taux nominaux pour actualiser des quantités réelles. Néanmoins leurs conclusions sont loin de faire l'unanimité.

Le second article de la thèse [2] a donc pour objet principal d'étudier la pertinence empirique du modèle de la Fed, ou plus précisément de déterminer son domaine de validité, aussi bien sur le plan géographique (pays et indice de marché), sur le plan temporel (période d'étude), que sur l'horizon possible du modèle, ou du moins de sa version linéaire. Si le modèle de la Fed fonctionne à moyen terme mais pas à long terme, cela serait cohérent avec l'hypothèse d'une illusion monétaire des agents, que ceux-ci corrigent dans le temps. Néanmoins, nos résultats montrent que le modèle de la Fed ne s'applique qu'à une période restreinte (1980-2000), à un pays particulier (les Etats-Unis) et un indice boursier donné (le Standard and Poor's 500). L'approche fréquentielle employée permet, pour ce pays et cette période temporelle, de déterminer l'horizon du modèle, étude qui n'a à notre connaissance jamais été tentée de manière systématique.

1.2 Résumés détaillés des articles

Article [1]: Money-illusion: French evidence

L'article a pour objectif d'étudier la sensibilité individuelle à l'illusion monétaire d'étudiants français de niveau master à partir d'un questionnaire du type de celui de Shafir *et al.* (Shafir, Diamond,

& Tversky, 1997). Le questionnaire a été posé dans un cadre expérimental sur machine permettant d'isoler les sujets et de contrôler leur information ainsi que le report de leurs réponses de manière rigoureuse, homogène et indépendante. Les 107 étudiants interrogés provenaient de différentes formations (économie, sciences, droit et lettres). Les résultats de l'article confirment une forte illusion monétaire chez les sondés, économistes inclus, mais moins importante néanmoins que dans l'étude de Shafir (Shafir, Diamond, & Tversky, 1997). Des asymétries de perceptions de justice sociale sont constatées entre les scénarios de hausse de salaire et d'inflation, et de baisse des prix et des salaires, même dans le cas de contrats indexés sur l'inflation, et sur ce point les résultats du groupe d'étudiants économistes ne diffèrent pas de ceux des non-économistes. En général les résultats diffèrent selon le type de formation, le groupe le plus sujet à l'illusion monétaire étant le groupe d'étudiants en lettres et en droit, suivi des étudiants en sciences, et, en dernier, le groupe d'étudiants en économie. Cela atteste d'un certain effet apprentissage ou de dé-biaisement, même si cet effet n'est que partiel, la majorité des étudiants, quelle que soit leur formation, étant victime d'illusion monétaire. Ce résultat conforte l'hypothèse d'une illusion monétaire fortement ancrée dans les comportements.

L'article met aussi en évidence un biais heuristique dans l'évaluation subjective des conditions qui ne relève pas de l'illusion monétaire proprement dite.

Article [2]: The validity and time-horizon of the Fed model: a co-integration approach

L'article propose une analyse économétrique de la pertinence du « modèle de la Fed » qui se fonde sur la comparaison d'une grandeur nominale, le taux des obligations d'Etat, et d'une grandeur réelle, le rendement des actions (ratios bénéfice/prix), pour identifier une sous-évaluation ou surévaluation du marché actions, ce qui suppose une certaine illusion monétaire des agents. Plus précisément, cet article a pour but de déterminer le domaine de validité de ce modèle en termes d'étendue géographique, de période historique et d'horizon temporel en utilisant principalement des méthodes provenant de la théorie de la co-intégration. En effet, puisqu'une classe d'actifs ne peut par définition pas rester constamment sous-évaluée, on obtient comme condition nécessaire du modèle de la Fed la stationnarité de certains indicateurs, ou de co-intégration des deux séries.

Dans cet objectif, l'article considère les indices boursiers et les taux nominaux des obligations d'Etat de 21 pays et montre que seuls les indicateurs de trois pays répondent à ces critères de stationnarité: les Etats-Unis, l'Italie et le Mexique. Ainsi, pour la plupart des marchés, le modèle de la Fed ne passe pas le test le plus basique de stationnarité, et un indicateur qui devrait être stationnaire possède au contraire une tendance stochastique, son niveau actuel dépendant de l'accumulation de chocs antérieurs qui ne montrent pas de tendance particulière à s'auto-corriger. En outre, même pour les Etats-Unis et le Standard and Poor's, pays et indice pour lequel le modèle a été formulé, il faut restreindre la période d'études de 1980 à 2000 pour obtenir un indicateur stationnaire. L'étude se focalise ensuite sur les Etats-Unis et sur cette période, et utilise une comparaison entre modèles similaires avec ou sans composante de long terme linéaire (la relation de la Fed estimée) pour déterminer l'horizon du modèle. Les résultats montrent qu'à fréquence journalière, le pouvoir prédictif du modèle est nul. Cependant, la relation estimée améliore la prévision du rendement des actions dès la fréquence hebdomadaire, tandis que pour le taux nominal des obligations elle permet d'obtenir des prédictions supérieures seulement à partir d'une fréquence mensuelle. Finalement, on peut conclure que le modèle de la Fed dans lequel une comparaison entre les niveaux des deux grandeurs (rendement des actions et taux nominal des obligations) est prise en compte ne produit des prévisions que très faiblement supérieures à un simple modèle dynamique incorporant seulement le feedback entre les variations des deux variables[4].

L'article est complété par un document de 65 pages réunissant des Annexes supplémentaires contenant d'autres résultats quantitatifs qui confortent les conclusions de notre étude; les annexes fournies dans ce document ne sont pas nécessaires à la compréhension de l'article.

Thème 2 : l'industrie post-négociation

2.2 Contexte et motivation

TARGET2-Securities (T2S) est un projet de l'Eurosystème pour le règlement des actifs financiers. Développé par les banques centrales de la zone euro (Eurosystème), T2S offrira aux dépositaires centraux nationaux (CSDs) un système harmonisé de règlement en monnaie centrale qui participe à l'unification du marché financier européen. L'introduction de T2S aura un impact significatif sur l'industrie des services post-négociation en Europe. L'entrée en opération de T2S est prévue pour début 2015, mais les principaux CSD utilisent déjà la plateforme en mode de test. Aujourd'hui, environ 40 CSDs opèrent en Europe dans ce qui demeure un ensemble de marchés fragmentés, domestiques, et donc largement monopolistes. En 2001, un groupe de travail dirigé par l'économiste Alberto Giovannini publia un rapport affirmant que "l'inefficience de la fonction de compensation et de règlement représente l'obstacle le plus important à l'intégration des marchés financiers en Europe". Les inefficiences concernent, entre autres, les interfaces techniques, les formats des messages électroniques, et les règles de finalité des règlements, ainsi que les jours ouvrés et les horaires journaliers. Le rapport identifie 15 obstacles précis à l'efficience du marché, qui ont été nommés les 15 barrières de Giovannini; la plupart de ces barrières n'ont pas été levées depuis 2001. Ces obstacles contribuent aux coûts plus élevés pour le règlement transfrontalier en Europe comparé, par exemple, aux Etats-Unis[5]. Une des conséquences d'un marché pour le règlement moins efficace en Europe qu'aux Etats-Unis est une perte relative d'attractivité des marchés financiers européens vis-à-vis des investisseurs internationaux. Le but de T2S est essentiellement de créer un marché unique pour le règlement en Europe[6].

Le règlement d'actifs consiste en le transfert final des actifs du vendeur vers l'acheteur et du paiement de l'acheteur vers le vendeur, que ce soit sur un marché boursier ou en commerce bilatéral. Il existe plusieurs définitions d'un règlement transfrontalier. Par exemple, Oxera (Oxera, Methodology for monitoring prices, costs and volumes of trading and post-trading services, 2001), (Oxera, Monitoring prices, costs and volumes of trading and post-trading services, 2009) et (Oxera, Monitoring prices, costs and volumes of trading and post-trading services, 2011)) définit une transaction domestique comme une transaction où le domicile de l'investisseur et celui de l'actif sont les mêmes, et parle de règlement transfrontalier lorsque ces deux domiciles sont différents. Diverses

[4] Le modèle correspondant étant un VAR en première différence.
[5] Voir, par exemple, ECB (2007), NERA (2004) et Oxera (2009, 2011).
[6] La banque centrale européeenne fournit des informations concernant le projet sur le site www.t2s.eu.

sources corroborent que la proportion des règlements transfrontaliers serait proche de 40% du volume total de transaction en Europe. Tout d'abord, 37% des compagnies européennes listées seraient possédées par des investisseurs étrangers en 2007 d'après la Federation of European Securities Exchange (FESE). Ensuite, Clearstream (Clearstream, 2002), page 7 et 15) suggère qu'entre 35% et 40% du commerce d'actions serait transfrontalier[7].

Enfin, Oxera (Oxera, Monitoring prices, costs and volumes of trading and post-trading services, 2011) rapporte que les proportions des transactions transfrontalières d'actions se répartiraient entre 57% pour les brokers locaux et 27% pour les brokers globaux en 2009[8].

En termes de règlement cela se traduit par une utilisation de liens presque négligeable représentant 1 à 2% du règlement total. Les liens sont créés entre deux CSDs pour permettre le transfert des actifs d'un CSD à l'autre. Dans la majorité des cas l'un de ces CSDs est le CSD où les actifs ont été émis, et l'autre est appelé CSD investisseur[9]. Le tableau fournit en Annexe de ce document indique la part des règlements transfrontaliers dans les différents Etats de l'Union Européenne via les liens des CSDs.

Les seules institutions où l'on observe une utilisation significative des liens, (i.e. plus de 25% du volume total de transactions) sont les trois CSDs Baltiques (Esthonie, Lettonie et Lituanie), le CSD Autrichien et les deux CSDs internationaux, Euroclear Bank et Clearstream Banking Luxembourg. Les CSDs Baltiques ont une quantité relative de transactions transfrontalières importante car ils émettent des actifs qui sont échangés sur une place boursière commune. Le CSD Autrichien OeKB a établi un lien très utilisé avec le CSD allemand German Clearstream Banking Frankfurt. Les deux CSDs internationaux ont un business model qui est un mélange de celui d'une banque dépositaire et d'un CSD. Ils ont été créés en 1970 pour le règlement des eurobonds, c'est-à-dire des obligations libellées dans une devise différente de celle du pays émetteur. Bien que les CSDs internationaux couvrent tout type d'instruments financiers, comme les actions et les parts de fonds d'investissement, les obligations sont restées leur principal focus et comptent pour plus de 80% du volume de leur règlement. Pour tous les autres CSDs, l'utilisation des liens est marginale, presque inexistante.

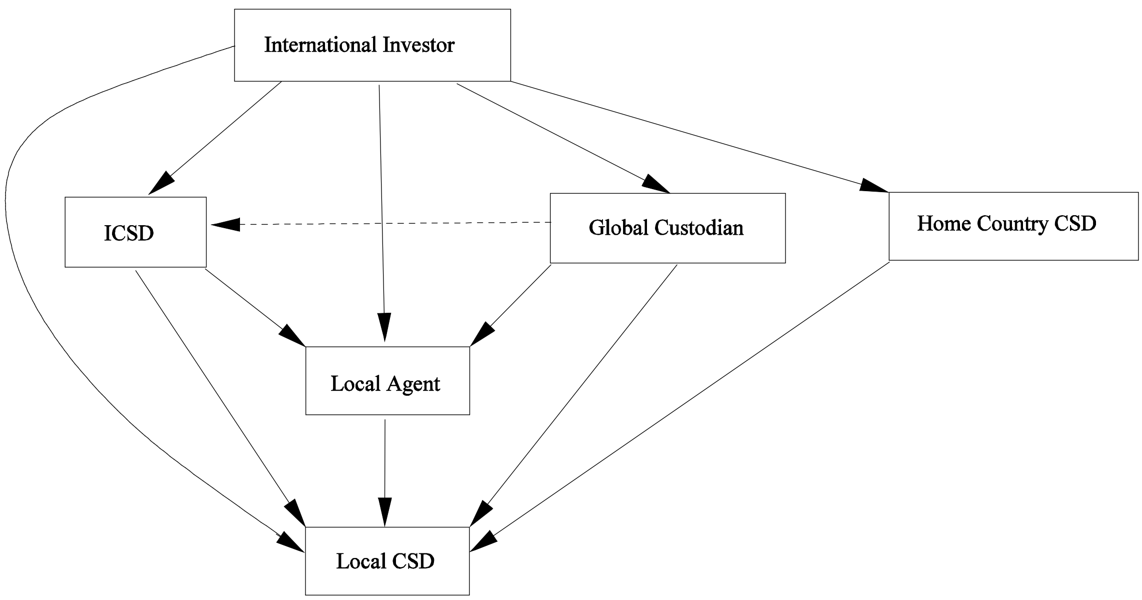
Cette faible utilisation suggère que le règlement des transactions transfrontalières via les liens établis par les CSDs n'est pas attractif dans la configuration actuelle du marché, probablement parce qu'utiliser les liens est à la fois complexe et coûteux. Giovannini (Giovannini, 2001) souligne l'inefficience de la partie cash des règlements qui manque parfois des services que savent par ailleurs

[7] Les actions constituent la plus grande part du volume échangé en nombre de transactions. Au niveau du règlement, environ 80% du volume de transaction dans les CSDs Européeens provient des actions, d'après le Livre Bleu de la BCE (données de 2005-2010). Les obligations sont échangées moins fréquemment mais la valeur de leur transactions est significativement plus élevée.

[8] Les brokers globaux sont présents dans de multiples centres financiers, tandis que les brokers locaux sont présent dans un unique centre financier, et par conséquent tous leurs clients établis dans un autre centre financier que le leur sont comptabilisés comme des clients transfrontaliers. Cela explique la part moindre des brokers globaux.

[9] L'issuer CSD ouvre un compte permettant aux investisseurs (dans un direct holding system) et/ou à leurs intermédiaires (par exemple leurs investors CSDs) de posséder des titres. Un investor CSD -- ou une tierce partie agissant en la qualité de l'investor CSD -- ouvre un compte avec un autre CSD (l'issuer CSD) afin de permettre un réglement des titres. Le site internet de la BCE possède un glossaire des termes: http://www.ecb.int/home/glossary/html/glossa.en.html.

procurer les banques dépositaires. Les investisseurs préfèrent utiliser le service de leur réseau de banques dépositaires ou de CSDs internationaux[10]. D'après Clearstream (Clearstream, 2002), les banques dépositaires offrent un meilleur niveau de service ainsi qu'une gamme plus large de services, tels que le calcul des éventuelles taxes, un reporting extensif et personnalisé, la gestion des droits de votes liés aux actions, etc. tandis que les CSDs se focalisent sur des processus automatiques et moins complexes liés aux règlements domestiques. La figure ci-dessous résume les différents canaux permettant de réaliser des règlements transfrontaliers:



Les différentes manières d'accéder à un marché étranger en Europe. Source: Giovannini (2001, p. 8).

L'introduction de T2S permettra d'éliminer les obstacles à l'utilisation des liens entre CSDs, en particulier ceux dus à la différence des interfaces techniques, au format des messages, au règlement intra-journalier de finalité, aux horaires d'opérations. T2S sera aussi probablement un catalyseur pour une harmonisation plus en profondeur de l'industrie des services post-négociation. T2S favorisera aussi la compétition, non seulement en diminuant les coûts d'entrée du marché de manière drastique pour les nouveaux entrants (puisque ceux-ci pourront simplement utiliser la plateforme IT de T2S pour les règlements au lieu de devoir construire la leur propre), mais aussi en poussant les CSDs à se

[10] Un correspondant titre ("custodian") est une entité, souvent une banque, qui pourvoit des "custody services", tels que le "safekeeping", et l'administration des instruments financiers pour ses clients. Les correspondants titres ou "custodians" pour reprendre le mot anglais, peuvent pourvoir d'autres services à leurs clients, tels le "clearing" et le règlement ("settlement"), le "cash management", des services liés aux conversions de devises, ainsi que des prêts d'instruments financiers. Un "sub-custodian" est un custodian pourvoyant ses services à un autre custodian. Clearstream (2005) définit un "agent bank" comme une banque, soit locale soit branche d'une banque étrangère, qui agit de la part du correspondant titre d'un investisseur étranger. Un "local agent" est un correspondant titre qui pourvoit ses services pour des instruments financiers échangé et réglés dans le pays où il se trouve à des intermédiaires et contreparties d'un autre pays.

déplacer vers le haut de la chaîne des valeurs ajoutées des valeur ajoutée et en favorisant ainsi la compétition entre CSDs et banques dépositaires.

La contribution clef de T2S, qui se traduit au niveau des hypothèses du modèle général présenté dans [4], est que le processus technique pour un règlement domestique dans T2S est fondamentalement le même que pour un règlement transfrontalier avec n'importe quel autre CSD ayant rejoint T2S. Ainsi, l'ensemble des liens vers l'intégralité des CSDs participant à T2S se trouve entièrement établi pour un CSD joignant T2S, du moins du point de vue technique. D'un point de vue légal les CSDs devront également signer un contrat concernant l'utilisation des liens. Le critère d'éligibilité au projet (voir (ECB, T2S Economic Feasibility Study, 2007), en particulier, le critère 3) assure que chaque CSD joignant T2S a le droit et le devoir de rendre ses actifs disponibles pour un transfert via un lien vers tout autre CSD joignant T2S. T2S chargera le même prix par transaction, que cette transactions soit transfrontalière et implique un lien entre deux CSDs ou non. Ainsi T2S traitera toute transaction comme domestique, et permettra aux agents économiques de bénéficier de réductions de coûts dues aux économies d'échelles, ainsi qu'à une plus grande compétition entre CSDs. Néanmoins, pour que le projet soit efficace, il est nécessaire que les CSDs adaptent leur propre architecture IT à la nouvelle plateforme. Tous les CSDs en Europe sont effet incités à déléguer la fonction "règlement" à la plateforme T2S. Ainsi, même si les CSDs continuent à maintenir la relation légale avec leurs clients, et à fournir à leurs clients d'autres services que le règlement, techniquement toutes les opérations de règlement seront réalisées dans T2S. Un dialogue informatique constant est donc requis entre chaque CSD et T2S, d'où l'importance, pour éviter la duplication des messages et les coûts par transaction qu'elle provoque, d'adapter la plateforme IT des CSDs. Le succès final du projet de l'Eurosystème dépend donc en partie de l'adaptation et de la coopération des CSDs, qui sont la plupart du temps des acteurs du secteur privé.

Le troisième article de la thèse [3] s'insère dans cette problématique. Il a pour objet d'étudier, à l'aide d'outils empruntés à la théorie des jeux, la possibilité, ainsi que la persistance possible, d'un scénario adverse en termes de coûts pour l'utilisateur final, à savoir une collusion tacite des CSDs pour ne pas s'adapter à T2S après avoir rejoint le système, en faisant supporter le coût plus grand des transactions unitaires à leurs clients[11]. Il s'agit de la première étude théorique de la décision de s'adapter techniquement ("reshaping decision"). L'Eurosystème a publié deux études, purement descriptives (c'est-à-dire sans modélisation des décisions stratégiques des partis concernés) sur T2S et les bénéfices économiques potentiels liés au projet. L'étude de faisabilité économique de T2S (ECB, T2S Economic Feasibility Study, 2007) a été publiée en 2007. Ensuite, l'Eurosystème a publié une évaluation de l'impact économique de T2S (ECB, T2S Economic Impact Assessment, 2008), réalisée avec la participation des acteurs de marché[12]. Sur la base de la méthodologie sélectionnée, la BCE a demandé à tous les CSDs Européens, et à un sous-ensemble de leurs clients, des données spécifiques. Les résultats sont résumés dans le document (ECB, T2S Economic Impact Assessment, 2008) et montrent que T2S devrait permettre de réaliser des économies de coûts de l'ordre de 145 à 223 millions d'euros par an, suivant l'hypothèse de couverture de devises retenue[13]. Le focus de ces deux

[11] Les clients des CSDs sont la plupart du temps les banques commerciales, mais faire passer des frais de transactions plus élevés peut avoir des conséquences sur les clients de ces banques (investisseurs privés, particuliers, fonds de pensions).

[12] La méthodologie fut inter alia sujette à un appel d'offre publique, voir http://www.ecb.int/paym/cons/html/t2s-2.en.html.

[13] En effet d'autres devises que l'euro seront éligible pour T2S.

études quantitatives de la BCE a été l'évaluation des coûts associés à T2S et la réduction des coûts des CSDs associée, en termes des coûts de back-office et en termes d'économie de collatéral et de liquidité supplémentaire[14]. Aucune étude n'avait tenté jusqu'à présent de modéliser directement l'impact positif sur les bénéfices liés à T2S qu'auraient les effets de réseaux. Un aspect fondamental de l'industrie des services post-négociations en général est pourtant bien son effet de réseau inhérent: la participation d'acteurs de marché supplémentaires augmente l'utilité de tous les participants du réseau de règlement. En effet, plus de participants implique que chaque participant peut établir des liens avec plus de partenaires, donc que la liquidité s'accroît (voir par exemple (ECB, The Payment System. Payments, Securities and Derivatives, and the Role of the Eurosystem, 2010), chapitre 5). L'industrie des services post-négociations se caractérise également par de grandes économies d'échelles (voir (P. Van Cayseele, 2007) pour une estimation quantitative). Ces effets expliquent que le coût de transaction moyen tend à décroître quand le nombre d'utilisateurs du réseaux augmente. Ces effets potentiels sont importants pour expliquer la décision future des acteurs économiques de joindre ou non un réseau donné. En effet, bien que la participation à T2S ne soit pas obligatoire, T2S permettra à chaque CSD joignant son réseau d'être un point d'entrée unique pour ses clients sur toute l'intégralité du marché des autres CSDs participants au projet.

Le quatrième article de la thèse [4] propose une méthodologie à base de modélisation des réseaux pour produire des simulations illustrant cet aspect réseau négligé. L'article est présenté de manière abstraite. Trois arguments nous ont conduit à faire ce choix. Le premier tient à la nature des données utilisées pour les simulations. La précision des données indiquant les flux transfrontaliers des CSDs s'étant révélée insuffisante, présenter des simulations chiffrées pourrait conduire à des interprétations erronées, surtout lorsque l'on inclut le risque de modélisation. Le second argument relève de la confidentialité de ces informations. Le troisième argument est que la méthodologie utilisée pourrait potentiellement s'appliquer à d'autres domaines. Ainsi notre article se présente comme un article de méthodologie pure, se focalisant sur la modélisation de l'effet réseau en général. Economides (Economides, The economics of networks, 1996) présente une description abstraite générale de l'effet de réseau en organisation industrielle, de même que les articles de Economides et Salop (Economides & C., Competition and integration among complements and network market structure, 1992), et de Park et Ahan (Park & Ahn, 1999) qui traitent de l'inter connectivité de deux réseaux du point de vue de l'établissement d'un prix d'équilibre. Une introduction informelle aux effets de réseaux dans l'industrie des services post-négociations peut se trouver chez Knieps (Knieps, 2006). Jusqu'à présent, la littérature concernant l'industrie des services post-négociation s'est contentée d'emprunter les techniques de modélisation de l'effet réseau à l'organisation industrielle, qui suppose une forme a priori de l'utilité marginale dérivée de la participation d'un membre de plus au réseau. Ainsi, cette utilité n'est pas dérivée du modèle. Dans notre approche au contraire, nous donnons un fondement micro-économique à la forme de l'utilité de l'effet de réseau dérivée, car nous modélisons explicitement le réseau. Pour des réseaux réels cependant, dénués des formes de symétrie supposées dans l'article, recourir à des simulations de Monte-Carlo serait nécessaire. L'article se rattache aussi à la littérature des réseaux économiques, car il tente, dans une seconde partie, de caractériser les réseaux stables et les équilibres de Nash sur des réseaux dérivés de la première partie. La littérature des réseaux économiques a été appliquée avec succès à des cas aussi divers que l'étude de la transmission sociale des opportunités d'emploi (Calvo-Armengol et Jackson (Calvo-Armengol & Jackson, 2004)), la conclusion des accords de libre-échange (Goyal & Joshi, 2003), les liens de co-auteurs (Jackson &

[14] Ces économies tiennent au fait qu'un seul compte d'actifs permettra de contenir le collatéral nécessaire pour les prêts en monnaie centrale nécessaire au règlement cash de la transaction.

Wolinsky, 1996). De plus ce courant de la littérature économique contient de nombreux articles formulés de manière abstraite et à même de s'appliquer à plusieurs situations (voir principalement: Bala et Goyal (Bala & S., 2000), Belleflamme et Bloch (Belleflamme & Bloch, 2004), Dutta *et al.* (Dutta, Ghosal, & Ray, 2004), Goyal et Vega-Redondo (V. & Goyal, 2000), ainsi que Jackson (Jackson M. , A strategic model of social and economic networks, 1996), (Jackson M. , The evolution of social and economic networks, 2002)). A notre connaissance aucun de ces modèles n'a encore été formulé dans l'objectif de modéliser l'industrie du règlement. C'est un fait d'autant plus étonnant que le vocabulaire même des utilisateurs suggère une telle application (le vocable de « lien », ou « link » en Anglais, couramment utilisée par les participants de l'industrie dans le contexte des règlements transfrontaliers, et aussi le terme générique employé par les théoriciens de la littérature de réseau pour désigner une arête orienté d'un graphe). En outre, les règles de partage des profits au long d'une chaîne d'intermédiaires proposées dans l'article n'ont pas été étudiées à notre connaissance, même de façon purement théorique. La question de détermination de la totalité des équilibres de Nash, ou même des réseaux stables, reste à ce jour encore ouverte.

2.2 Résumés détaillés des articles

Article [3] : Optimal CSD reshaping towards T2S

Les CSDs ont besoin de réorganiser leur achitecture actuelle afin d'accéder à T2S. Cette adaptation à T2S demande des investissements, souvent conséquents, de leur part. En pratique il existe une infinité de stratégies possibles entre les deux approches extrêmes suivantes. Dans la première approche le CSD cherche à minimiser ses coûts d'adaptation au maximum en ne s'adaptant que le minimum requis pour se connecter à T2S. C'est la "T2S-fees-on-top approach". Le CSD garde simplement son architecture IT actuelle et construit une interface pour dialoguer avec T2S. En particulier, le CSD devra répliquer chaque message, donnée, etc. provenant de T2S dans son propre langage avant de pouvoir traiter l'information et d'offrir ses services spécifiques. Une telle approche ne permet pas d'obtenir des réductions de coûts au niveau du traitement de chaque instruction individuelle. Au contraire, les frais opérationnels de T2S doivent être ajoutés aux coûts du CSD. Cette approche permet cependant de contenir les coûts d'adaptation immédiats au maximum. Dans la seconde approche extrême, le CSD adopte une "green field approach" dans laquelle il établit une architecture IT entièrement nouvelle et prévue pour fonctionner avec le plus de synergies possible avec T2S. Dans ce cas, le CSD minimisera ses coûts de traitement des transactions individuelles, mais devra faire face à des coûts d'adaptations importants.

Les CSDs Européens font face à une décision cruciale: comme la participation à T2S est volontaire, doivent-ils joindre T2S? Et, pour ceux qui décident de joindre le réseau, quel est le degré optimal d'adaptation à la plateforme? C'est-à-dire quel degré de réorganisation de leur architecture IT, mais aussi de leurs ressources humaines et de leur business modèle est le plus souhaitable? En particulier, les CSDs doivent décider du prix optimal de leurs services dans le nouveau contexte européen. Par exemple, certains seraient tentés d'augmenter leur prix pour faire face aux coûts d'adaptation, transférant ainsi lesdits coûts à leurs clients.

L'objectif de l'article est d'apporter des éléments de réponses à ces questions cruciales en modélisant la décision des CSDs de s'adapter à T2S dans le cadre de la théorie des jeux. Deux types de jeux sont utilisés : un jeu fini et un jeu infiniment répété, dans lequel la décision du degré d'adaptation dépend à la fois des coûts fixes d'adaptation, irréversibles, et de la future rentabilité, qui est fonction de la réduction des coûts de traitement des transactions mais aussi du potentiel accroissement des volumes pouvant être obtenu en réduisant les prix, la demande étant supposée partiellement élastique. On suppose, dans la première partie de l'article, un jeu fini avec une décision de s'adapter qui ne peut être prise qu'au début du jeu. Le résultat est que les CSDs ne doivent pas augmenter leurs prix mais au contraire les diminuer en fonction de la réduction des coûts de transaction obtenue. On dérive ensuite une formule analytique close qui donne le degré d'adaptation optimal en fonction des paramètres du modèle, en particulier les coûts de transactions et la quantification de l'effet de substitution entre les services offerts par les différents CSDs. Des simulations sont produites à l'aide de données publiques.

Dans la seconde partie de l'article, on considère un jeu infini et on permet aux CSDs de s'adapter à n'importe quel moment du jeu. Cela introduit un comportement stratégique de la part des CSDs, puisque cela leur permet d'établir des stratégies qui dépendent des stratégies d'autres joueurs. Un important résultat de l'article est la dérivation de conditions suffisantes sous lesquelles un équilibre de Nash en sous-jeux parfait existe dans lequel les CSDs délaient sans cesse leur décision de s'adapter. Cependant cet équilibre semble peu robuste, en particulier si l'on considère la possibilité de nouveaux venus sur le marché. En outre la condition s'avère être aussi nécessaire. Ainsi, on obtient une condition sous laquelle il ne peut y avoir de délai à la décision des CSDs de s'adapter. Les résultats sont robustes vis-à-vis de plusieurs modifications du modèle (délai dans l'observabilité de la décision de s'adapter des autres CSDs, généralisation à un nombre quelconque de CSDs, etc.). On montre également comment une simple translation d'un des paramètres permet d'appliquer notre modèle à la décision simultanée de joindre et de s'adapter à T2S.

Article [4] : Endogenizing the network effect: how Monte-Carlo simulations and Graph Theory could benefit economic modelling

Cet article apporte trois principales contributions à la modélisation des effets de réseau. La première est la dérivation d'une forme de l'utilité de l'effet de réseau à partir de fondements microéconomiques: l'effet de réseau est endogène au modèle, et l'on n'a donc pas besoin d'en spécifier une forme particulière ex-ante. L'effet réseau sur la fonction d'utilité des agents est entièrement déterminé par le nombre de participants, les hypothèses de coûts et de profits, et la structure du marché[15].

La seconde contribution principale de l'article est d'illustrer le problème de coordination des agents pour joindre un nouveau réseau, problème dont la solution optimale tout comme les équilibres de Nash dépendent des coûts opérationnels du nouveau réseau, des coûts individuels de transaction des agents, de leur coût d'adaptation au nouveau réseau, ainsi que de la croyance a priori de chaque agent sur les décisions des autres agents de joindre ou non le réseau. Les conséquences en termes de politique économique sont, tout d'abord, qu'une transparence accrue sur les coûts d'adaptation de chaque agent ainsi que sur les coûts de la plateforme, mais aussi sur les décisions des agents de la joindre, facilite la

[15] Bien sûr, une fois les effets de réseaux dérivés du modèle, ils peuvent être utilisés pour ajouter une autre "couche" de modélisation, et c'est précisément ce qui est fait pour endogénéiser la décision des agents de joindre ou non le réseau à l'aide d'outils de la théorie des jeux, ou encore pour l'étude des équilibres de Nash et de la stabilité de réseaux.

détermination de l'équilibre et son orientation vers un équilibre efficient. En outre, il peut être nécessaire pour un sous-ensemble d'agents de synchroniser leur entrée dans le réseau: cela dépend de la distribution jointe de leurs coûts individuels.

La troisième contribution principale de l'article est plus académique et concerne la détermination de réseaux stables pour un jeu basé sur une règle de partage simple des profits. La modélisation de chaînes d'intermédiaires introduit des considérations d'ordre topologique complexes, et par conséquent déterminer l'ensemble des réseaux stables dans les cas non-triviaux reste un problème ouvert. Par exemple, un agent peut être amené à détruire un lien avec un intermédiaire A qui lui était profitable car il lui permettait d'accéder à un second agent B qui était relié à de nombreux autres agents dans le but de se lier directement à B ou à un autre agent remplissant cette même fonction. La question de savoir si l'agrégation de plusieurs réseaux efficients donne un réseau efficient est aussi étudiée. La réponse est que laisser aux agents le soin de réorganiser le réseau ne permet pas nécessairement la mise en place d'un réseau global efficient, en particulier si les agents cherchent à conserver les liens déjà établis de manière domestique.

Thème 3 : les modèles de défauts

3.1 Contexte et motivation

La crise économique a considérablement augmenté la probabilité de défaut des entreprises tout comme des particuliers. Le risque de crédit est donc à nouveau un facteur essentiel dans l'évaluation des actifs par les acteurs de marché. En outre, le secteur privé n'est pas le seul secteur exposé au risque de crédit : l'Eurosystème, parce qu'il prend des instruments financiers en collatéral des prêts qu'il accorde aux banques commerciales, est aussi exposé a des pertes en cas de double défaut, c'est-à-dire dans le cas ou à la fois une banque et son collatéral ferait défaut. De plus, lors d'achats de titres adossés à des prêts à l'économie réelle, l'Eurosystème serait cette fois directement exposé au risque de crédit. Les modèles de default s'avèrent alors nécessaires pour que les prêteurs puissent mesurer le risque du collatéral reçu, ou pour évaluer le risque associé à des achats directs. De même, du point de vue des régulateurs cette fois, les modèles de défaut sont nécessaires dans le cadre des normes de solvabilité bancaire afin de déterminer les capitaux requis. La réglementation en vigueur de Bâle II, dans son approche avancée, et ses révisions successives ou futures, utilisent en effet un modèle de défaut. On le voit, l'estimation des risques de défaut est une problématique cruciale qui concerne autant le secteur privé, désireux d'optimiser son capital risque et ses décisions d'investissement, que le secteur public, autorités monétaires et régulateurs inclus. Il est donc utile de se demander quels sont les principaux modèles de défaut qui permettent de quantifier le risque de crédit utilisés, et leurs limitations potentielles.

La première difficulté dans la prévision des défauts est souvent le manque de données pour calibrer de tels modèles. Les défauts sont souvent des évènements rares, en tout cas pour les firmes d'une certaine qualité minimale de crédit. Cela explique que la plupart des modèles de défaut couramment employés, que ce soit pour le pricing des Collateralised Debt Securities (CDO), pour le stress-testing des Asset Backed Securities (ABS) par les agences de notations, ou encore pour le calcul des règles de capital des banques, forment en réalité une seule et même famille de modèles dérivés du modèle de Merton (Merton, 1974), qui s'insère dans le cadre théorique du pricing d'option à la Black et Sholes. On y suppose simplement que la firme fait défaut lorsque le processus qui décrit sa valeur intrinsèque tombe en-dessous du niveau de la dette à rembourser.

Une approche différente, et plus macroéconomique, consiste à tenter d'expliquer les séries de défauts par l'ajout de variables macroéconomiques. Deux obstacles doivent être surmontés : la disponibilité de séries suffisamment longues et représentatives de l'économie, d'une part, et le problème du choix des variables macroéconomiques d'autre part.

Récemment Giampeieri et al (Giampieri, Davis, & Crowder, 2005) ont obtenu de Standard and Poor's des séries de défauts concernant les obligations d'entreprises étatsuniennes suffisamment longues pour appliquer une telle approche. Parce que les cycles économiques, tels que mesurés par le PIB par exemple, ne s'avèrent pas correspondre aux cycles de crédits observés, les auteurs ont appliqué une technique empruntée à la reconnaissance de langages (Rabiner, 1989) pour estimer, sans faire aucun choix de variables macroéconomiques, les cycles de crédits : les chaînes de Markov cachées.

Le dernier article de la thèse a pour objectif d'appliquer cette méthodologie dans un contexte Européen. La première difficulté a été d'obtenir des séries de défauts suffisamment longues pour divers secteurs européens. Les agences de notation ne sont pas très coopératives à ce sujet, car leurs séries sont des données confidentielles d'où elles tirent précisément leur avantage stratégique par rapport aux nouveaux entrants potentiels dans leur secteur plutôt fermé. La deuxième étape consiste à implémenter de manière efficace, en évitant les problèmes typiques de sous-flots dans l'estimation des probabilités de défauts conditionnelles, des algorithmes capables d'estimer les cycles cachés de l'économie en termes de risque de crédit. Ces algorithmes nous ont permis d'identifier le cycle de crédit de l'économie à partir de deux types de données observées : la série de taux de défaut des entreprises européennes obtenue de Moody's d'une part et un taux de défaut générique des PME espagnoles calculé à partir de données provenant de Bloomberg d'autre part.

3.2 Résumé détaillé de l'article

Article [5] : The modelling of default rates through hidden Markov chains models

L'article présente les principaux modèles de défaut ainsi que leur utilisation dans le cadre de l'Eurosystème, avec un focus particulier sur les portefeuilles de titres de crédits.

La première partie de l'article présente les modèles communément utilisés, de manière auto-suffisante et avec des notations homogènes, établissant les liens entre les différents modèles. On y montre qu'une grande famille de modèles structurels de défauts provient en réalité d'un même modèle originel proposé par Merton (Merton, 1974), et donc que le risque modèle est concentré : une seule et unique grande famille de modèle est utilisée et l'on manque de modèles alternatifs pour confirmer ou invalider ces modèles. L'article utilise des notations homogènes afin de faciliter la compréhension des liens entre modèles, et pointe leurs principaux atouts et défauts.

La seconde partie de l'article propose un modèle original fondé sur des estimations à base de chaînes de Markov cachées afin de déterminer le cycle de crédit de certains actifs (obligations d'entreprises ou prêts aux PME) en Europe. Des algorithmes efficients sont décrits dans l'article. Implémentés en Matlab, ils sont utilisés afin d'estimer l'état de crédit de l'économie (haut risque / bas risque) indépendamment des variables macroéconomiques utilisées généralement pour prédire les cycles de l'économie (PIB, etc.). Le modèle est d'abord appliqué sur des données générées suivant les hypothèses du modèle à titre d'illustration. Puis, des données observées sont mobilisées, d'une part la série de taux de défaut des entreprises européennes obtenue de Moody's et d'autre part, un taux de défaut générique des PME espagnoles calculé à partir des données de défauts de Bloomberg relatives prêts aux PME espagnoles qui forment le portefeuille sous-jacent des Asset Backed Securities (ABS) éligibles comme collatéral pour les opérations monétaires de l'Eurosystème. L'état du cycle de crédit de l'économie est ainsi identifié.

Conclusion générale

En conclusion, cette thèse offre l'intérêt d'étudier plusieurs problématiques centrales et actuelles de la finance moderne : la rationalité limitée des agents et leurs biais comportementaux vis-à-vis des valeurs nominales, le problème de la juste évaluation du prix des actions, la refonte du paysage de l'industrie post-négociation en Europe suite à l'introduction du projet de l'Eurosystème Target-2 Securities, ainsi que les modèles de défaut et les méthodes d'estimation des cycles de défaut pour un secteur donné. Si les thèmes sont variés, les techniques à la fois théoriques et appliquées le sont aussi, illustrant la capacité sans cesse renouvelée des sciences économiques à mobiliser des méthodes relevant d'une pluralité d'approches et de domaines: enquêtes sur données individuelles, économétrie, théorie des jeux, théorie des graphes, simulations de Monte-Carlo, chaînes de Markov cachées...

Notre enquête sur l'illusion monétaire confirme la robustesse des résultats d'études précédentes tout en dévoilant de nouvelles perspectives de recherche, par exemple tenter d'expliquer la disparité des réponses selon les caractéristiques individuelles des répondants, en particulier leur formation universitaire. Réitérer l'étude sur un échantillon d'étudiants de même niveau, mais provenant de pays dotés d'une culture distincte de l'inflation, pourrait mettre en exergue d'éventuelles différences culturelles liées à la perception de l'inflation. Notre étude du modèle de la Fed montre que la relation de long terme entre taux nominal des obligations d'Etat et rendement des actions n'est ni robuste, ni utile à la prédiction sur des horizons temporels réduits. Néanmoins, tester le modèle en incluant explicitement d'autres variables, de nature macro-économique afin de capturer les tendances et potentiellement les ruptures structurelles, est une voie encore ouverte à la recherche. Enfin, le modèle d'estimation des défauts à partir de chaînes de Markov cachées gagnerait à être appliqué sur des portefeuilles comprenant moins de titres, mais dont les titres sont plus risqués, afin de juger de sa pertinence dans ce cadre plus difficile pour les prédictions.

Travaux cités

Asness, C. (2003). Fight the Fed model. *Journal of Portfolio Management, 30*, 11-24.

Bala, V., & S., G. (2000). A noncooperative model of network formation. *Econometrica 68, no 5, 1181-1229*, 1181-1229.

Bekaert, G., & Engstrom, E. (2010). Inflation and the stock market: understanding the "Fed Model". *Journal of Monetary Economics*, 278-294.

Belleflamme, P., & Bloch, F. (2004). Market sharing agreements and collusive networks. *International Economic Review 45, no 2*, 287-411.

Boucher, C. (2006). Stock prices-inflation puzzle and the predictability of stock market returns. *Economics Letters, 90 52)*, 205-212.

Bourgeois-Gironde, S., & M., G. (2012). Keynes' animal spirits vindicated: An analysis of recent empirical and neural data on money illusion. *Journal of Post-Keynesian Economics, 34(2)*, 331-352.

Calvo-Armengol, A., & Jackson, M. O. (2004). The effect of social networks on employment. *American Economic Review 94*, 426-454.

Camerer, C., Loewenstein, C., & D., P. (2005). Neuroeconomics: How Neuroscience Can Inform Economics. *Journal of Economic Literature*, 9-64.

Campbell, J., & Shiller, R. J. (1989). The dividend-price ratio and expectations of future dividend and discount factors, . *The Review of Financial Studies, Volume 1, no 3*, 195-228.

Campbell, J., & Vuolteenaho, T. (2004). Inflation illusion and stock prices. *American Economic Review, 94(2)*, 19-23.

Clearstream. (2002). *Cross-Border Equity Trading, Clearing and Settlement in Europe.* White Paper.

Dutta, B., Ghosal, S., & Ray, D. (2004). Farsighted network formation. *Journal of Economic Theory 122*, 143-164.

ECB. (2007). *T2S Economic Feasibility Study.* http://www.ecb.europa.eu/press/pr/date/2014/html/index.en.html.

ECB. (2008). *T2S Economic Impact Assessment.* https://www.ecb.int/paym/t2s/about/keydocs/html/index.en.html.

ECB. (2010). *The Payment System. Payments, Securities and Derivatives, and the Role of the Eurosystem.* Tom Kokkola.

Economides, N. (1996). The economics of networks. *International Journal of Industrial Organization 14, no 2*, 675-699.

Economides, N., & C., S. S. (1992). Competition and integration among complements and network market structure. *The Journal of Industrial Economics, 10, no 1*, 105-123.

Fehr, E., & Tyran, J.-R. (2001). Does Money Illusion Matter. *American Economic Review, 91(5)*, 1239-1262.

Fisher, I. (1928). *The Money Illusion.* New York: delphi Company.

Giampieri, G., Davis, M., & Crowder, M. (2005). Analysis of default data using hidden Markov Models. *Quantitative Finance, vol 5, issue 1*.

Giovannini, A. (2001). *Cross-border clearing and settlements arrangement in the European Union.* http://ec.europa.eu/internal_market/%EF%AC%81nancialmarkets/docs/clearing/%EF%AC%81rst_giovannini_report_en.pdf.

Goyal, S., & Joshi, S. (2003). Networks of collaboration in oligopoly. *Games and Economic Behaviour 43*, 57-85.

Jackson, M. (1996). A strategic model of social and economic networks. *Journal of Economic Theory 71*, 44-74.

Jackson, M. (2002). The evolution of social and economic networks. *Journal of Economic Theory 106*, 265-295.

Jackson, M., & Wolinsky, A. (1996). A strategic model of economic and social networks. *Journal of Economic Theory 71 (1)* , 44-74.

Kahneman, D. (2003). Maps of Bounded Rationality: Psychology for Behavioral Economics. *The American Economic Review, 93(5)*, 1449-1475.

Keynes, J. M. (1935). *The General Theory of employment, interest, and money.* First Harvest/Harcourt.

Knieps, G. (2006). Competition in the post-trade markets: A network economic analysis of the securities business. *Journal of Industry, Competition and Trade, vol 6, no 1*, 45-60.

Lucas, R. (1996). obel lecture : Monetary neutrality. *Journal of Political Economy*.

Merton, R. (1974). On the Pricing of Corporate Debt: the Risk Structure of Interest. *Journal of Finance 29*, 449-470.

Modigliani, F., & Cohn, R. (1979). Inflation, rational valuation and the market. *Financial Analysts Journal, 35*, 24-44.

Oxera. (2001). *Methodology for monitoring prices, costs and volumes of trading and post-trading services.* Report prepared for the European Commission DG Internal Market and Services.

Oxera. (2009). *Monitoring prices, costs and volumes of trading and post-trading services.* Report prepared for the European Commission DG Internal Market and Services.

Oxera. (2011). *Monitoring prices, costs and volumes of trading and post-trading services.* http://ec.europa.eu/internal_market/%EF%AC%81nancialmarkets/docs/clearing/2011_oxera_study_en.pdf.: Report prepared for the European Commission DG Internal Market and Services.

Park, Y.-S., & Ahn, B.-H. (1999). Joint ownership and interconnection pricing in network industries. *International Review of Economics and Finance 8*, 183-198.

Rabiner, L. R. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. *Proc. IEEE 77*, 257-286.

Shafir, E., Diamond, P., & Tversky. (1997). Money Illusion. *The Quarterly Journal of Economics, 112(2)*, 341-374.

Shiller, R. (1999). Speculative prices and popular models. *Journal of Economic Perspectives 4(2)*, 55—65.

V., B., & Goyal, S. (2000). A noncooperative model of network formation, 68, no 5. *Econometrica*, 1181-1229.

Van Cayseele, P., & Wuyts, C. (2007). Cost efficiency in the European securities settlement and depository industry. *Journal of Banking & Finance 31, no 10*, 3058-3079.

Weber, B. R. (2009). The medial prefrontal cortex exhibits money illusion. *PNAS*, 106(13), 5025-5028.

Article 1: Money-illusion at individual level: French evidence[1]

1. Introduction

Rational well-informed agents are supposed to be able to appropriately take inflation into account in their decision-making process as well as in exercising their judgment on the economic situations they are faced with. However, the difficulties people experience in distinguishing clearly nominal from real quantities have been recognized since a long time. Irving Fisher was the first to consider that money illusion – "the tendency of people to think of money in nominal, rather than real terms"– was a widespread phenomenon, and an illusion on which people acted upon, with dire economic consequences. He even wrote an entire book (Fisher, 1928) to make the public aware of the difference between nominal and real price. He reported, for example, his talk with a shop tender in Germany who refused to increase the price of the product he was re-selling because he had brought it at a lower price and would consider it unfair to make such a profit, thereby jeopardizing his ability to renew his stock in the future – since the price he himself pays for it had increased. Fisher's main goal was to make both the public and the authorities aware of the great damages that can cause an unstable currency given the prevalence of money illusion. However, money illusion seemed probably so obvious to Fisher that he did not carry out any quantitative study to prove its existence, preferring to prone, with a clear practical goal in mind, more stability in the purchasing power of the currency.

Keynes (1936) considered more generally that money illusion occurs when "individuals do not accurately take inflation into account" and might partly explain the downward stickiness in nominal wages. Commonly used in macroeconomics for many years, money illusion became since the end of the sixties an hypothesis largely abandoned, probably more for normative reasons – the large consensus on the economic rationality model – than for positive ones. If people are assumed to be rational, they should care only about real magnitudes and real changes; hence they are free from money illusion.

The interest in money illusion was renewed from an empirical standpoint and largely grew out of the influence of experimental psychology applied to economics following Tversky and Kahneman,

[1] This paper has been co-written with Marianne Guille (Université Panthéon-Assas, LEMMA (EA 4442), guille@u-paris2.fr) ; the authors would like to thank Anna Bernard for her precious help in organising the experimental sessions and Anne Corcos for her comments on earlier version of this article.

(Tversky & Kahneman, 1981), which focuses on the different biases separating agents from the *homo economicus* model.

The first type of empirical evidence came from large qualitative surveys on wages and prices which show that money illusion can partly explain their nominal stickiness. [2] But, Shafir, Diamond and Tversky (1997) carried out the first extensive survey at the individual level focused on money illusion. Their survey is based on a questionnaire which was completed by several groups of about 100 to 300 respondents in different places: Two New Jersey shopping malls, the Newark Airport and, finally, at Princeton University - the respondents being undergraduate students. The results were presented in an aggregated form since the authors did not find any important difference between the respondent groups. They show that a majority of respondents think in nominal, rather than real, terms, when asked to compare certain situations in various contexts covering wages, transactions and contracts. Moreover, most of them expect that other people would also be subject to money illusion. The authors analyze this strong money illusion effect as a framing effect reflecting agents' preference for the nominal framework given its salience and easiness to process, even though agents may use the real one or a mix of both as well. They consider that the favored cognitive framework may vary depending on the context and the decision-maker's experience, for instance agents are more likely to use the real one during periods of hyperinflation.

The relevance of money illusion was further supported by different type of studies. A first type of studies show that a pure nominal shock as for instance the changeover to the euro in January 2002 triggered significant real effects. [3] Another one support the Modigliani and Cohn (1979) hypothesis that even stock market investors seem to suffer from a particular form of money illusion, discounting real cash flows at nominal discount rates. [4] Finally, money illusion bias at the individual level is supported by experimental studies, including notably Fehr and Tyran (2001, 2007, 2008) price-setting game experiments. They show that most subjects are sensitive to money illusion once they are shown gains expressed in nominal, rather than real, terms; a bias that has powerful effects on equilibrium selection or may cause aggregate nominal inertia. Another sort of evidence is provided by some neuroeconomic studies based on much more simplified experimental-task (see Weber, 2009 and Miyapuram, Tobler, Gregorios-Pippas, & Schultz, 2009). They use advanced techniques of brain imagery to observe the cerebral mechanisms potentially involved in money illusion as expressed by individuals. Their results support the existence of money illusion at the infra-individual level since they identify a specific impact of nominal changes on the reward circuit.

[2] Cf. Kahneman, Knetsch and Thaler (1986), Blinder and Choi (1990) or Blinder (1994).
[3] Cf. Kooreman, Faber and Hofmans (2004), Cannon, E. and Cipriani (2006).
[4] See Campbell and Vuolteenaho (2003), Cohen, Polk and Vuolteenaho (2005) or Boucher (2006).

These findings support the hypothesis of a phenomenon very deeply rooted in human nature, to an extent which was not suspected before (Bourgeois-Gironde & Guille, 2012). However, the precise nature of this bias at the individual level requires further investigation. One of the questions that remain largely unexplored is whether there is an impact of education on money illusion. In particular, is there a learning effect associated to an economics training, since it implies that people have been trained to think in real terms, or to some other particular training? If the answer is affirmative, then a specific program of education could be a means of reducing the prevalence of money illusion in the population, hence increasing the efficiency of markets. If the answer is negative, that is, if there is no or few learning effect linked to a particular training, especially in economics, then a specific program of education could not be sufficient to reduce the degree of money illusion in the population. Governments should also implement measures encouraging people to take into account the real framework, especially in periods of high inflation. Such measures are rare. A notable exception is the creation of an indexed unit of account, the *Unidad de Fomento* (UF), by the Chile in 1980. Defined as the amount of pesos needed to buy the cost of a living bundle, the UF is used to state future payments on forward contracts: They are paid in pesos depending on the UF's value which is published daily. Other measures that could be considered include the creation of markets to hedge inflation risk, such as the European inflation futures market at the Chicago Mercantile Exchange (HICP), as recommended by Shiller, or the adoption of mandatory revaluation measures, such as those that cover rents or child support payments in some countries.

This question was addressed by Cipriani, Lubian and Zago (2008) who carried out a survey among Italian undergraduate students from several backgrounds. Their questionnaire was reduced to a question taken from Kahneman, Knetsch and Thaler (1986) about the fairness of a limited adjustment of nominal wages on prices presented in two versions, in case of inflation and deflation, the two versions being identical in real terms. The proportions of first-year students in economics and mathematics judging the (same) real wage cut unfair were identical in the two versions contrary to the results obtained for other students (law, tourism and foreign languages). Hence, the authors consider that first-year students in economics and mathematics were less likely affected by money illusion than other students. Since these differences were identified from the very beginning of their training and because third-year students in economics were as much affected by money illusion than first year ones, the authors conclude that their findings suggest there is a selection bias – as if the most rational people are drawn to economics or mathematics - rather than a learning effect – a training of two years in economics does not allow to reduce the degree of money illusion.

This article aims to study more precisely the potential impact of different academic trainings on money illusion. To that end, a large questionnaire, close to that of Shafir, Diamond and Tversky (1997), was administered to more advanced students from several backgrounds. All participants are

French graduate students either enrolled in a Master or having a Master degree in economics, in science or in law or humanities. We try to limit some of the biases typically associated to survey-based studies. First, the questionnaire was completed by participants on computer within an experimental framework to control their information and ensure independent answers. Second, to reduce as far as possible the interpretation biases that might introduce errors in subject responses, questions and their context were formulated in clear and explicit terms, in particular the elements useful to give a rational answer were provided to minimize personal views or speculation. Third, individual characteristics were collected to test a possible interaction with money illusion. Hence, participants filled up several questionnaires in addition to that on money illusion, probing usual individual variables as well as their risk aversion, nominal loss aversion and financial literacy skills.

The next section presents the survey and sample characteristics. Section 3 explains the money illusion survey and results obtained. Finally, the last section provides some concluding remarks.

2. Survey design and sample characteristics

The survey was administered to 106 participants (56 females and 50 males) at LEEP (the Experimental Economics Laboratory of Paris) in May 2012. Participants were aged 23 on average. All of them were graduate students either enrolled in a Master or having a Master degree in economics (42), science (42) or in law or humanities (22). We control for previous exposure to economics by asking the non-economics students to mention any economics or finance training.

Type of education:	male	Female	total
Science	17	25	42
Economics	21	21	42
Law and literature	12	10	22
total	50	56	106

The survey was anonymous and all students were told that there were no right or wrong answers and that they will receive a fixed amount of money (5 euros) for their participation.

Before the questionnaire on money illusion, participants filled up three questionnaires aimed to collect individual characteristics that could interact with money illusion. The first one probes usual individual variables such as sex, age and education as well as a brief measure of the Big-Five personality domains. The second questionnaire aims to provide a rapid test of their financial literacy skills. The last one consists in a usual Holt-Laury risk aversion measure procedure (Holt & Laury, 2002). At the end of the questionnaire on money illusion, three questions were added to test their nominal loss aversion.

Financial literacy

Wording

The following two simple questions were asked, testing the understanding of interest compounding, a basic financial concept, with serious implications for usual decisions as savings or retirement planning.

Question 1. Suppose you placed 1000 euros on a saving account which provides a return of 2% per year. After one year, how much will you have on this saving account if you did not made any withdrawal nor added any money to it?

Less than 1020 euros

Exactly 1020 euros

More than 120 euros

Don't know

Question 2. After 5 years, how much will you hold on this saving account if you have not added or withdrawn any money?

Less than 1100 euros

Exactly 1100 euros

More than 1100 euros

Don't know

Categories used in the study

A financial literacy score of 0 was given to all respondent having failed both financial literacy questions (32%) and a score of 1 to those having answered correctly both questions (68%). This allows creating two groups of comparable size to distinguish between financially literate subjects and financially illiterate ones.[5]

Results

Results show that although neither education nor sex has any significant impact on the number of correct answers for the relatively simple first question[6], both education and sex have an impact on the number of correct answers for the second question since students in science and economics gave the best answers, and male underperformed females.

Answer given to Question 2:	Less than 1100 euros	Exactly 1100 euros	More than 1100 euros	Don't know
Science	2%	17%	81%	0%

[5] Only 5% of respondent failed to correctly answer Question 1.

[6] Results can be obtained from the authors on demand.

Economist education	5%	12%	83%	0%
Law and literature	0%	50%	45%	5%
Average	3%	23%	79%	1%

Fischer[7] exact test p-value: 0.002

Fischer exact test indicates a very low p-value of 0.002, hence the differences are highly significant.

Answer given to Question 2:	Less than 1100 euros	Exactly 1100 euros	More than 1100 euros	Don't know
male	0%	30%	68%	2%
female	5%	14%	80%	0%
Average	3%	22%	75%	1%

Fischer exact test p-value: 0.032

Risk aversion

Wording

To measure subjects' risk aversion and study its potential interaction with money illusion, we use the usual lottery-choice method of Holt and Laury (Holt & Laury, 2002), presented in the table below.

	Scenario A		Scenario B	
Question number:	probability to win 2 euros	probability to win 1.60 euro	probability to win 3.85 euro	probability to win 0.10 euro
1	10%	90%	10%	90%
2	20%	80%	20%	80%
3	30%	70%	30%	70%
..	..	..	..	..
9	90%	10%	90%	10%
10	100%	0%	100%	0%

[7] We use Fisher's exact test to test significant differences in answers because it does not require any assumption on the expected frequency of each cell and computes the p-value directly.

Scenarios A as well as Scenario B can have two different realizations with varying probabilities over the 10 choices presented. The later individuals switch from scenario A to scenario B, the higher their risk-aversion.

Categories used in the study

Based on the Holt and Laury (Holt & Laury, 2002) survey it is possible to classify subjects into two groups: the low risk-averse subjects turned to Option B before scenario 4 (50%) and the high risk-averse subjects at scenario 5 or after (50%).

Results

The table below indicates the results split off of the sample.

Question at which subject switch to Scenario B as preferred scenario:	**0**	1	2	3	4	5	6	7	8	9	10
Percentage of subjects	3.77	0.94	4.72	19.81	20.75	17.92	17.92	7.55	4.72	0.94	0.94

Results show that neither education nor sex has any significant impact on the degree of risk aversion which is captured by the scenario in which respondents start to prefer the riskier Scenario, Scenario B, to the other, Scenario A.

Nominal loss aversion

Wording

The questions were as follows.

Question 1. Suppose you are participating in a game in which you start by receiving 100 euros and then can choose between two options:

Option A: You keep 75 euros.

Option B: You have 25% chance to keep 100 euros and 75% chance to lose 100 euros.

Question 2. Suppose you are participating in a game in which you start by receiving 100 euros and then can choose between two options:

Option C: You lose 25 euros.

Option D: You have 75% chance to lose 100 euros and 25% chance to keep 100 euros.

Question 3. Suppose you have the choice between the two following scenarios:

Scenario E: You have 200 euros and have 50% chance to win 100 euros more and 50% chance to lose 100 euros.

Scenario F: You have 100 euros and have 50% chance to win 200 euros more and 50% chance to win nothing more and keep your 100 euros.

Which of the two scenarios would you rather be in?

Categories used in the study

Subjects were considered to exhibit a strong nominal loss aversion whenever they chose the riskier, lower return option to avoid a sure nominal loss in the second question of the nominal loss aversion part of the survey. The score assigned to such subjects (18%) was 2. Question 3 was taken into consideration differently, as it provides evidence of nominal loss aversion but to a much lower degree, the two described situation being identical. The score assigned to subject preferring Scenario F of Question 3, and being not considered as subject to strong loss aversion (56%), was 1. Subjects which were not assigned a score of 1 or of 2 (26%) where considered exhibiting no nominal loss aversion and consequently assigned a score of 0.

Results

Question 1 and Question 2 are phrased such as to have exactly the same payoff for option A and option C, on one hand, and option B and option D, on the other hand. Option B and D have lower expected payoffs than Option A and C, and are more risky. Unsurprisingly, most subjects (95.3%) preferred option A to option B in Question 1. Nevertheless, in Question 2, only 82% of subjects still prefer option C. The only explanation is the framing of Question 2, which explicitly mentions the nominal loss (you *lose* 25 euros), contrary to identical Question 1 which mentions the remaining nominal gain (you *keep* 75 euros). Nominal loss aversion triggered a more risky behavior from some subjects: To avoid the perceived loss, they chose the more risky option despite of its lower expected payoff.

Sex is a completely irrelevant explanatory variable for that behavior (Fisher test p-value of 0.80) contrary to education, which is almost significant at the 10% level. Surprisingly economics and science students exhibit the less rationale behavior since they have a different behavior in both questions and considering that they were no differences between question 1 answers across groups of students.

Answer given to first question:	Option A	Option B
Science	95%	5%
Economics	95%	5%
Law and literature	95%	5%
Average	95%	5%

Fischer exact test p-value: 1.000

Answer given to second question:	Option C	Option D
Science	83%	17%
Economics	74%	26%
Law and literature	95%	5%
Average	82%	18%

Fischer exact test p-value: 0.103

Among the three groups only the literature and law group answers remain unchanged from Question 1 to Question 2. There is no significant difference in the answers to Question 3 regarding education and sex.

Answer given to Question 3:	Option E	Option F
Science	31%	69%
Economics	29%	71%
Law and literature	50%	50%
Average	34%	66%

Fischer exact test p-value: 0.231

3. Money illusion survey and results

The questionnaire on money illusion is based on that of Shafir, Diamond and Tversky (1997) but some questions and/or their context are formulated in a more explicit way to minimize personal asumptions that could biase the answers of respondents who are asked to compare and judge different hypothetical situations i.e. scenarios rather than the situation of different persons. The questionnaire includes 12 questions which have been presented to participants in the following order.

Question 1

This question aims to test if subjects are able to compare in economic terms real vs. nominal wage increase in two situations characterized by different inflation rates.

Wording

Julie has graduated from University. Like all other students, she has neither savings nor debt. Her diploma will allow her to get the job she wants. She does not plan to make any savings nor to take on any debt in the first five years of her working life, meaning she will use, each month, all her monthly pay to buy consumer goods but will not go into debt. All the other students have similar plans. We are concerned exclusively with these five first years. There are two different scenarios that could occur; in each one the inflation rate is perfectly known to Julie and to all other people in the economy, and at the beginning of each year. Also, Julie buys goods whose price increases/decreases are perfectly represented by the inflation rate. After these first five years, there will be no more inflation or pay increase in the economy.

Scenario A: Julie is offered a job that pays EUR 20000 annually, and the economy is under a stable, constant and known in advance annual inflation rate of 4% for all countries during the whole five years period. Julie gets a pay-increase each year, the trajectory of her salary being:

Year	**1**	**2**	**3**	**4**	**5**
Julie salary :	20000 euros	21000 euros	22050 euros	23152 euros	24310 euros

That is, her salary increases of 5% each year.

Scenario B: A: Julie is offered a job that pays EUR 20000 annually, and the economy is under a stable, constant and known in advance inflation rate of 1% for all countries. Julie gets a pay-increase each year, the trajectory of his salary being:

Year	**1**	**2**	**3**	**4**	**5**
Julie salary :	20000 euros	20400 euros	20808 euros	21224 euros	21648 euros

That is, her salary increases of 2% each year.

In your opinion, what do you think would be the best scenario in economic terms?

Scenario A Scenario B Both

Before analyzing the results, it is useful to review the four main considerations that led to the precise wording of Question 1.

First, it seems useful to **state explicitly Julie possible prior savings or debt**, since a higher inflation clearly benefits debt holders. In Shafir et al. previous study (Shafir, Diamond, & Tversky, 1997), the respondents could guess or imagine this part of the problem, which can lead, particularly in a country like the United States where many students contract student loans, to an overstatement of the real extend of money illusion in the respondents' minds.

Second, it seems also useful **to be precise with Julie's plans for savings or debt**. It is very possible that savings are made more difficult in an inflationary environment than in a stable price environment due to the lack or inaccessibility of proper hedging instruments. Debt could potentially be less expensive in a higher inflation environment if the lender failed to properly predict the future inflation rate to adjust the required nominal rate of interest. In the question Julie's plans for the five year period ahead are made as simple as possible: Julie will not save nor contract debt. Hence no freedom of interpretation that could interfere with our measure of money illusion is left to the respondents.

Third, the phrasing of the question avoids comparing the situations of two different people, as in Shafir et al, in order not to favor some **interregional arbitrage** in respondents' mind. Indeed, if someone has a better pay, in nominal terms, in some region, he or she might take advantage of the lower inflation rate in another region to buy goods or services from there. Hence an inflationary environment can seem more attractive to respondents, whatever their money illusion. Another detrimental effect of framing the problem with two different people instead of using different scenarios is that it prompts interpersonal comparisons in a way that may capture some other part of human psychology independently of money illusion, for instance **nominal comparisons between salaries may flatter one's ego relative to others**. Indeed, empirically it is true that the most prestigious places to live are also the most expensive. Although the real salary may be the same, the person having the higher nominal salary thus may feel superior (even if not geographically mobile). Hence, the wording is such as to avoid any social consideration based on nominal earnings, and uses the concept of "Scenario". This implies different possible universes for the same person instead of referring to comparisons between individuals living in the same universe, which could provoke unwanted inter-individual comparisons. To avoid individual comparisons the question also mentions that other students do not have any savings or debt. We also stated other students have similar plans than Julie, such that no comparative advantage exists between Julie and the others[8].

Fourth, the question assumes that the inflation rate is both representative and an accurate measure of the cost of living. The relative high pay of Julie might otherwise lead respondents to believe he buys goods which differ, in their price increases, from those used in the basket of goods used to compute inflation. This is relevant in the context of our current economic situation where food prices are more volatile than technological products while making a larger relative share of modest households expenditures compared to above-the-average-earnings households. Similarly it is important that respondent knows that the inflation rate is known to Julie and the other students, such that they do not imagine they could have been fooled in a way or another by different than expected inflation when choosing the goods they buy and consume[9].

Results

Since both scenarios are approximately equivalent in real terms (+ 1% real wage growth per year) while a precise calculation indicates a small advantage in the scenario B (+1.0099 vs. +1.0096), the results suggest that only about one third of the subjects, those who have chosen the high inflation scenario (A), can be considered as exhibiting money illusion. Therefore, almost half of the subjects

[8] If not, it could have been argued, for example, that Julie, who is more inclined to consume than to save could *comparatively* benefit from a higher inflation environment and of the non-availability to saving-prone students of appropriate financial instruments to hedge against inflation, although this argument seems a little far-fetched.

[9] Moreover, in Shafir et al article (Shafir, Diamond, & Tversky, 1997) the adjustment of the salary to the inflation rate is made with a full year lag. It could thus come as no surprise that Barbara, who endured a 4% inflation rate before getting, at the end of the year, her 5% salary increase, was deemed, at that moment, happier than Ann, who had no inflation during the first year and then received a 2% salary increase – although the response to the third question, that it is Ann who is more likely to quit for another job, is consistent with money illusion.

correctly chose the scenario B – either because they managed to do the precise calculation[10] or because they had a preference for a low inflation – and a minority considered that both scenarios were identical in economic terms (19%). Answers differ depending on the subjects' type of education, and depending on their level of financial literacy, with financially literate subjects able to see that the two situations are roughly equal much more often than financially illiterate ones (24% compared to 9%).

Question 1: What is the best scenario in economic terms?	Scenario A (higher inflation)	Scenario B (lower inflation)	Both scenario are the same
Results:	35%	46%	19%

It is likely that the careful wording of Question 1, as detailed above, explains why our results differ significantly from Shafir's survey, finding fewer subjects exhibiting money illusion (35%). But the difference could also be the consequence of different samples (nationality, level of education).

Education type

The term "economically" may have prompted students to adopt a more economically oriented frame of thought, and possibly contribute to some extent to the lesser degree of money illusion observed from the answers. To try to assess in what extend, one can compare the non-economics students' answers to the economics students' one. The expectation would be that the economics students would give more often the correct answer than non-economics students, in particular when reminded to think in economics terms. The results confirm this expectation.

Question 1: What is the best scenario in economic terms?	Scenario A (higher inflation)	Scenario B (lower inflation)	Both scenario are the same
Science	45%	36%	19%
Economics	17%	60%	23%
Law and literature	50%	41%	9%

Fischer exact test p-value: 0.02; Pearson chi2 p-value: 0.025

Less than 2 economics students in 10 chose the high inflation scenario (A) vs. around half of the other students. Moreover, a large majority of them (60%) correctly chose the low inflation scenario (B) vs. only 3 or 4 other students in 10. These results mean that a serious training in economics, during which students are used to think in real terms, may be useful to reduce money illusion bias when subjects are asked to compare different scenarios in economic terms. Moreover, this training seems especially

[10] The similar percentage of financially literate subjects and financially illiterate subjects preferring the low inflation environment does not sustain this particular interpretation, which is understandable given that scenario B is only 0.11% better than scenario A.

useful to manage to compare exactly both scenarios, even if a preference for a low inflation environment cannot be excluded since they have learned that a low rate of inflation is better. Surprisingly, the only difference observed between the other students is related to the third answer: students in science are more likely to consider that both scenarios are the same than students in law and literature. It could be that financial literacy has an impact on subject's preferences. This hypothesis is supported by the results obtained from the test of subjects' financial skills.

Financial literacy

Financial literacy has a strong impact on the type of subject's preferences. Financial literate subjects are less likely to choose the high inflation scenario (A) and more able to consider both scenarios are the same, thus exhibiting less money illusion. This result supports the hypothesis that money illusion may be partly explained by a lack of financial skills.

Question 1: What is the best scenario in economic terms?	Scenario A (higher inflation)	Scenario B (lower inflation)	Both scenario are the same
Not financially literate (score = 0)	50%	41%	9%
Financially literate (score = 1)	28%	49%	24%

Fischer exact test p-value: 0.046

<u>In conclusion:</u>

1. These results show that the average prevalence of money illusion (about one third) conceals significant differences between students: Only 17% of the students in economics exhibit money illusion vs. about 50% of the other students.

2. Students in economics are especially more likely to prefer the low inflation scenario (60%) than other students (about 40%).

3. Both other groups of students prefer the high inflation scenario (A).

4. Financially literate subjects are more able to see that the two situations are the same (24%) compared to financially illiterate subjects (9%). Financially illiterate students tend also to prefer the high inflation scenario (50%) while financially literate students preferred the low inflation scenario (49%).

Question 2

This question, close to Problem 2 from Shafir et al, refers to real vs. nominal house price change but makes explicit the use of the funds obtained from the selling of a house and eliminates the risk of interregional arbitrage based decisions.

Wording

You have been living in a house you brought 200 000 euros one year ago, but decided this year to sell it to immediately re-purchase another house in the same region with all the proceedings of the sale. Consider the three alternative scenarios:

Scenario A: The prices of all services and goods, including houses and flats, in the economy have decreased by 25%. You are able to sell the house at 154 000 euros (23% less than what you paid).

Scenario B: There has been neither inflation nor deflation, and you sold the house for 198 000 euros (1% less than what you paid).

Scenario C: The prices of all services and goods, including houses and flats, in the economy have increased by 25%. You are able to sell the house for 246 000 euros (23% more than what you paid).

Please indicate:

According to you which Scenario describes the most profitable double house transaction (selling your house to buy another one at the market price)?

Which Scenario describes the least profitable one?

Which Scenario would bring you the most satisfaction?

Which Scenario would bring you the least satisfaction?

The most profitable double transaction is of course that of Scenario A, with a real gain of about +2% in terms of house values. Here again, the phrasing in terms of Scenario is aimed to prevent cross-region arbitrage.

Results

We find that the subjects seem adverse to the nominal losses entailed by Scenario A, and thus prefer the scenarios without deflation (Scenarios B and C) even if they entail real losses. Answers are almost completely unchanged across each different type of education and respondents' nominal loss aversion scores. By contrast, sex and especially financial literacy have an impact.

Interestingly, there are many subjects among those choosing the correct answer about the most profitable scenario (Scenario A) which still considered it is not the scenario which will bring them the most satisfaction. More precisely, 24% of the subjects correctly choosing Scenario A as the most profitable scenario also answered it is not the Scenario that would bring them the most satisfaction, and 21% answered it would bring them the least satisfaction.

Question 2	Scenario A Deflation, large nominal loss but real gain	Scenario B No inflation, small real loss	Scenario C Inflation, large nominal gain but real loss
Scenario that describes the most	39%	27%	33%

profitable double house transaction			
Scenario that describes the least profitable double house transaction	39%	17%	44%
Scenario that would bring you the most satisfaction	30%	19%	51%
Scenario that would bring you the least satisfaction	54%	10%	36%

Contrary to Question 1, students in economics do not provide the right answer more often than other student groups. It could be that they are more trained to think in real terms about wages than real estate or that a nominal loss aversion or anchoring bias is at play even if subjects' nominal loss aversion score has no impact on these answers .

Interestingly compared to other studies, the presence of an intermediate case has a significant impact since it was selected by a considerable number of participants: 26% recognized that Scenario B, with its **explicit** 1% loss – explicit since there is no inflation – is <u>*better*</u> than the Scenario C and its apparent 23% gains, while at the same time surprisingly failing to understand that it is not better than Scenario A, which translates into a 2% real gain, despite the 23% nominal loss[11]. These inconsistent answers at the individual level support the money illusion hypothesis.

Sex

Women preferred the no inflation scenario in larger number than men (25% instead of 8%) making less often the wrong choice of high inflation scenario (Scenario C) but also less often the right choice of deflationary scenario (Scenario A).

Question 2 bis: Scenario that describes the least profitable double house transaction	Scenario A Deflation, large nominal loss but real gain	Scenario B No inflation, small real loss	Scenario C Inflation, large nominal gain but real loss
Male	44%	8%	48%
Female	34%	25%	41%

Fischer exact test p-value: 0.063

[11] Indeed, because all houses have fallen by more (25%), the 2% gains compared to an average transaction will result either in a net positive cash flow from the two transactions (selling then buying a house of same standing and size) or by the ability to obtain, with the 154 000 euros of proceedings from the selling of the first house, a better house than the one previously owned ("2% better"). This is neglecting the (huge) transaction costs in real estate, but including transaction costs would make Scenario A even more attractive compared to other scenarios as the nominal value of houses is lower in Scenario A and real estate fees are often a percentage of the nominal rather than real value. Hence, transaction costs could not explain the results obtained.

Financial literacy

Again, financial literacy has a strong impact on subject's preferences: Financial literacy students were more likely to choose the scenario A, exhibiting then less money illusion than others students in Question 2 (even if they were surprisingly more likely to choose the scenario C) and Question 2 ter.

Question 2: Scenario that describes the most profitable double house transaction	Scenario A Deflation, large nominal loss but real gain	Scenario B No inflation, small real loss	Scenario C Inflation, large nominal gain but real loss
Not financially literate (score = 0)	26%	44%	29%
Financially literate (score = 1)	46%	19%	35%

Fischer exact test p-value: 0.027

Question 2 ter: Scenario that would bring you the least satisfaction	Scenario A Deflation, large nominal loss but real gain	Scenario B No inflation, small real loss	Scenario C Inflation, large nominal gain but real loss
Not financially literate (score = 0)	15%	21%	65%
Financially literate (score = 1)	38%	18%	44%

Fischer exact test p-value: 0.047

<u>In conclusion:</u>

1. In this question focused on real estate transaction, participants gave similar results whatever their education, in striking contract to Question 1 which focused on wages. A training in economics is here not efficient to reduce money illusion exhibited in subjects answers, contrary to financial literacy skills.

2. Despite the double-transaction phrasing, there is still a large majority of respondents who manifest money illusion (about 60% of them do not select the Scenario A, which is the best in real terms).

3. There is an intra-individual inconsistency which would need further exploration. Indeed the respondents who answered scenario B manage to see that in real terms B is superior to C, while failing to see it is inferior to A. There seem to be a nominal anchoring at play there, where a transaction realized at a much lower level than the initial buying is deemed a poor transaction independently of the level of utility it brings (as in Scenario A, it brings the maximum utility for the double-transaction envisioned). Hence one could see interacting

anchoring bias with money illusion, which would result in the intra-individual inconsistent answers observed for this group.

4. 21% of subjects choosing correctly answer A as the most profitable scenario also consider it as the least satisfying scenario. This result shows that even subjects able to indicate the best scenario in real terms from an economic viewpoint would not be satisfied with that scenario. The perception of satisfaction is here clearly distinct from economic assessment, and more sensitive to money illusion.

Question 3

The aim of this question is to test subjects' expectations about wage increase in case of inflation.

Wording

During higher inflation times, is it more or less probable that salary increases?

More probable Less probable The same

Results

Answers are reported below:

Question 3	More probable	Less probable	The same
During higher inflation times, is it more or less probable that salary increases?	58%	21%	21%

The answers indicate that the assumed close relationship of salaries moving one-to-one with inflation is being questioned by respondents, including students in economics. Among them 17% believe it is less likely that salary increase in high inflation times.

Sex

A significantly higher proportion of women than men (rightly) believe salary increases are more likely during high inflation times.

Question 3	More probable	Less probable	The same
male	48%	20%	32%
female	68%	23%	9%

Fischer exact test p-value: 0.011; Pearson chi2 p-value: 0.011

Financial literacy

Financial literacy is a very significant explanatory variable of the subject's belief on the probability of salary increases during inflationary times. More precisely, subject who were able to correctly discount future cash-flows in the financial literacy survey gave twice more often the answer that salary increases are more likely in higher inflationary environment than those who did not.

Question 3	More probable	Less probable	The same
Not financially literate (score = 0)	38%	29%	32%
Financially literate (score = 1)	68%	18%	14%

Fischer exact test p-value: 0.012

Risk aversion

Low risk-averse subjects are more likely to believe salary increases are more probable in inflationary times than do the high risk adverse subjects but the difference is low (62% vs. 55%).

Question 3	More probable	Less probable	The same
Low risk adverse	62%	26%	11%
High risk adverse	55%	17%	28%

Fischer exact test p-value: 0.078

Question 4

This question aims to collect subjects' judgments of fairness and luck about adjustment of wage in case of inflation. Tom wage was chosen such that it is not a round number to avoid involving psychological levels in numbers. The question was phrased as follows.

Wording

The economy this year has been experiencing an annual rate of 5% inflation. The firm which employs Tom has seen its profits increase by 5% compared to last year, and decided to increase all of its employee salaries by 5%. Tom's salary is thus increased from 1120 euros to 1176 euros.

Do you think Tom will consider himself:

lucky unlucky neither

In your opinion this situation is …….. (choose below) to Tom and the other employees.

fair unfair depends

Results

The majority of respondents –about 58% - answer correctly "none of those", exhibiting no money illusion. Nevertheless 38% of subjects do answer "lucky", potentially supporting the hypothesis that because it could have been the case that the salary of Tom would not have been increased, Tom may give a higher positive value to the present situation. Also, a significant number of subjects (56%) correctly believe that the situation of Tom is fair compared to other employees. There is no significant difference in answers across education group or sex type.

Question 4	lucky	unlucky	none of those
Do you think Tom will consider himself	38%	4%	58%
Question 4 bis	fair	unfair	depends
In your opinion this situation is …….. (choose) to Tom and the other employees	56%	7%	38%

Sex

Similar to Question 5, women seems to exhibit slightly less money illusion than men as they answer that Tom will consider himself neither lucky nor unlucky in a higher proportion (64%) than men (52%). However, more of them also consider the situation of Tom as unfair compared to other employees (13% vs. 0% for men), which is difficult to explain given than other employees have the same salary increase than Tom.

Question 4: do you think Tom will consider himself	lucky	unlucky	none of those
Male	48%	0%	52%
Female	29%	7%	64%
Question 4 bis: in your opinion this situation is …….. (choose) to Tom and the other employees	fair	Unfair	depends
Male	68%	0%	32%
Female	45%	13%	43%

Fischer exact test p-value first question: 0.025; second question: 0.060

Nominal loss aversion

Non nominally adverse subjects are better at providing the correct answer to Question 4 bis that the situation is fair (68%). Nominally adverse subjects are less prone to propose this correct answer but surprisingly the relation between that dimension and money illusion exhibited is non-monotonous since strongly nominally adverse subject are more likely to give the correct answer than slightly adverse ones (58% vs. 49%).

Question 4 bis: In your opinion this situation is …….. (choose) to Tom and the other employees	fair	Unfair	depends
Not nominal loss adverse	68%	7%	25%
Slightly nominal loss adverse	49%	2%	49%
Very strongly nominal loss adverse	58%	21%	21%

Fischer exact test p-value: 0.008

Question 5

The same two questions were asked, but within a deflationary context.

Wording

All prices in the economy have decreased by 5%. The firm which employs Tom has seen its profits decreased by 5% this year compared to last year. It decided to cut employees salary by 5%. Tom's salary is thus decreased from 1120 euros to 1064 euros.

Results

There is an important asymmetry between the two contexts while they are identical in real terms. Only 7% of subjects think that Tom will consider himself as lucky in deflation and 60% as unlucky vs. 38% and 4% in inflation. Moreover, about 5 times more subjects consider his situation unfair when his salary is adjusted downward in a global deflationary economy than when it is adjusted upwards in a similar inflationary environment, indicating huge differences of perception between the two contexts. There is no significant difference in answers across sex type for both questions as across education type for the first question, but there are significant differences across education type for the second question.

Question 5	Lucky	Unlucky	none of those
Do you think Tom will consider himself	7%	60%	34%
Question 5 bis	Fair	Unfair	depends
In your opinion this situation is …….. (choose) to Tom and the other employees	33%	32%	35%

In matters of fairness, respondents judge asymmetrically nominal adjustments of wages in case of inflation and deflation: It is considered fair to increase salary in case of inflation, but not fair to decrease it in deflation. This may be the root of downward wages rigidity. Possible reasons to explain that asymmetry include a nominal loss aversion or a nominal anchoring. Nevertheless, in the present case there is no significant differences between subjects which show nominal loss aversion and those who do not: The p-value of Fisher exact test is high, about 65%, for the usual categories defined earlier. [12] Hence the explanation could be a differentiated perception of wage adjustment according to price change that may be reinforced by a form of insurance implicit contract between employers and employee. Indeed, adjustment of wages in case of inflation is in favor of employees, who would otherwise loose some purchasing power, while in case of deflation, the absence of adjustment is an advantage. If, inflation is more prevalent than deflation, and wage increases often delayed and lower than price changes, agents may have an unfairness feeling about inflation: They consider they are often affected by inflation, then deflation may be an opportunity to regain purchasing power. Hence they could consider as fair not to decrease wages in case of deflation. Moreover, times of deflation are associated with more difficult economic conditions, and the employee might expect some protection from the firm, with no salary decrease; in return he may implicitly accept lower salary increases in times of inflation or growth to compensate.[13]

Education type

Students in economics and in law and literature provided more often the “it depends” answer when asked to assess Tom situation than students in science. Apart from this result, each of the groups is equally split between “fair” and “unfair” answers in the deflationary scenario of Question 5, indicating a similar confusion about fairness in case of a deflationary scenario. Hence the only difference seems to be that students in science are less likely to express a conditional assessment on the fairness of the situation.

Question 5 bis: in your opinion this situation is …….. (choose) to Tom and the other employees	Fair	Unfair	depends
Science	43%	43%	14%
Economics	24%	24%	52%
Law and literature	32%	27%	41%

Fischer exact test p-value: 0.006; Pearson chi2 p-value: 0.008

Contrary to Question 4, its symmetric in inflationary times, there is no effect of nominal loss aversion on answers. Ficher exact test p-values is close to 40 %, indicating no significant difference between groups.

Question 5 bis: In your opinion this situation is …….. (choose) to Tom	fair	Unfair	Depends

[12] The p-value of Fisher exact test is 0.648, that is, about 65%, for the usual categories of loss aversion defined earlier.

[13] Of course, the situation of the firm and market power of unions and employers may have an impact on wage adjustments and their perception.

and the other employees			
Not nominal loss adverse	4%	57%	39%
Slightly nominal loss adverse	7%	56%	37%
Very strongly nominal loss adverse	5%	79%	16%

Fischer exact test p-value: 0.396

Question 6

This question is designed so as to ask individuals to compare both situations i.e. adjustment of wages in case of inflation and deflation.

Wording

In which situation of the one described in Question 4 or in Question 5 do you think Tom would be luckier?

Question 4 Question 5 The same

In which situation of the one described in Question 4 or in Question 5 do you think Tom would be better off?

Question 4 Question 5 The same

In which situation of the one described in Question 4 or in Question 5 do you think the firm's decision is the fairest?

Question 4 Question 5 The same

Results

69% of the subjects believe Tom would be luckier in the inflationary scenario of Question 4 whereas only 40% believe it is the situation in which Tom is better off. Hence, a substantial proportion of the subjects who correctly chose the answer that none of the two scenarios is better for Tom than the other, and hence have no money illusion, consider that he would be luckier either because they expect him to have some money illusion or think that his wage could not have been raised.

About fairness, the huge asymmetry previously noticed applies again: Increasing the salary in inflation is deemed fair while decreasing it in deflation is deemed unfair since only 10% of subjects believe the firm decision is fairest in Question 5 as opposed to 57% who believe it is fairest in Question 4. There is no significant difference in answers across education group or sex type.

Question 6	Question 4 (inflationary environment)	Question 5 (deflationary environment)	none of those

In which situation of the one described do you think Tom would be luckier?	69%	8%	24%
In which situation of the one described do you think Tom would be better off?	40%	13%	47%
In which situation of the one described do you think the firm's decision is the fairest?	57%	10%	33%

Sex

More women indicated the deflationary scenario as being the fairest than men.

Question 6 ter: In which situation of the one described in Question 6 or in Question 7 do you think the firm's decision is the fairest?	Question 4 (inflationary environment)	Question 5 (deflationary environment)	none of those
Male	66%	4%	30%
Female	48%	16%	36%

Fischer exact test p-value: 0.066

Question 7

This question is aimed to check if an explicit indexation wage scheme has an impact on subjects' judgment.

Wording

The economy this year has been experiencing an annual rate of 5% inflation. The firm which employs Tom has seen its profits increase by 5% compared to last year. The firm had signed a contract indexed to inflation with all its employees, meaning the salaries it pays automatically adjust to inflation. Hence this year its employee will have a 5% salary rise. Tom's salary is thus increased from 1120 euros to 1176 euros.

Do you think Tom will consider himself:

lucky unlucky neutral

In your opinion this situation is …….. (choose below) to Tom and the other employees.

fair unfair depends

Results

Results are very consistent with those of Question 4. A slightly higher number of subjects consider the situation to be fair compared to Question 4, with less “it depends” answers collected. Clearly the fact to have signed and to respect an inflation-indexed contract reduces uncertainty about the situation fairness (less “it depends” answers) and honoring the contract is just “fair”. Moreover as many respondents consider Tom to be lucky in both situations, contrary to our expectations since an automatic adjustment might be more neutral than a decision of the firm. Hence we expected to collect more answers “neither lucky nor unlucky” to this question. It is possible that respondents judge Tom to have been lucky to sign an indexed contract in the first place, or simply that this particular discretionary item does not influence the respondents opinion concerning Tom luck, but only their opinion about fairness.

Question 7	lucky	Unlucky	none of those
Do you think Tom will consider himself	38%	6	56%
Question 7 bis	fair	Unfair	depends
In your opinion this situation is …….. (choose) to Tom and the other employees	70%	8%	22%

Nominal loss aversion

Again, the results are similar to those of Question 4 bis.

Question 7 bis: In your opinion this situation is …….. (choose) to Tom and the other employees	fair	Unfair	depends
Not nominal loss adverse	78%	11%	11%
Slightly nominal loss adverse	75%	3%	22%
Very strongly nominal loss adverse	47%	16%	37%

Fischer exact test p-value: 0.044

Question 8

The same two questions were asked, but within the following deflationary context.

Wording

All prices in the economy have decreased by 5%. The firm which employs Tom has seen its profits decreased by 5% this year compared to last year. The firm had signed a contract indexed to inflation with all its employees, meaning the salaries it pays automatically adjust to inflation. Hence this year its employee will have a 5% salary cut. Tom's salary is thus decreased from 1120 euros to 1064 euros.

Results

Again, results are consistent with those of Question 5, except the significant higher proportion of respondents finding the situation as fair: 52% believe it is fair, given the indexed contract, for the firm to decrease Tom salary, compared to only 33% when the firm decides (Question 5). Also, 49% believes Tom is unlucky in this question compared to the higher 60% of Question 5. Indexation clearly has an impact on fairness and luck perception, in the exact same environment and according to the same respondents. This result supports the previous hypothesis of a differentiated perception of wage adjustments: In case of an indexed-contract, the feeling of unfairness in case of inflation disappears so the asymmetry between the results obtained in both contexts reduces. Again, there is no significant difference in answers across education group or sex type.

Question 8	lucky	Unlucky	none of those
Do you think Tom will consider himself	8%	49%	43%
Question 8 bis	fair	Unfair	Depends
In your opinion this situation is …….. (choose) to Tom and the other employees	52%	21%	27%

Education type

Education type has an impact on answers to Question 8 bis. Students in science provided less often the answer "it depends" compared to students in economics (12 vs. 40%) and in law and literature (32%) when asked if Tom situation is fair, a result consistent with that obtained for Question 5 bis.

Question 8 bis: in your opinion this situation is …….. (choose) to Tom and the other employees	fair	Unfair	Depends
Science	62%	26%	12%
Economics	45%	14%	40%
Law and literature	45%	23%	32%

Fischer exact test p-value: 0.044; Pearson chi2 p-value: 0.052

Question 9

Like question 6, this question aims to ask individuals to compare both situations, the only difference being the indexation of the contract.

Wording

In which situation of the one described in Question 7 or in Question 8 do you think Tom would be luckier?

Question 7 Question 8 The same

In which situation of the one described in Question 9or in Question 10 do you think Tom would be better off?

Question 7 Question 8 The same

In which situation of the one described in Question 8 or in Question 9 do you think the firm's decision is the fairest?

Question 7 Question 8 The same

Results

Results are consistent with previous questions, but also show that changing the wording to imply passivity from the firm (the contract is now indexed) does not modify the desirability of one of the two scenario compared to the other in terms of luck: It only affects fairness, as the percentage of subjects believing the inflationary scenario is the fairest falls from 57% to 39% when indexation is added to the contract. Again, there is no significant difference in answers across education groups or sex type.

Question 9	Question 7 (inflationary environment)	Question 8 (deflationary environment)	none of those
In which situation of the one described do you think Tom would be luckier?	65%	9%	24%
In which situation of the one described do you think Tom would be better off?	44%	9%	46%
In which situation of the one described do you think the firm's decision is the fairest?	39%	10%	51%

Financial literacy

Financial literacy is a very significant explanatory variable of the type of subject's preferences. A majority of the financial literate subjects consider both situations as equivalent while a majority of the others consider Tom would be better in the inflationary one.

Question 9 bis: In which situation of the one described do you think Tom would be better off?	Question 7 (inflationary environment)	Question 8 (deflationary environment)	none of those
Not financially literate (score = 0)	56%	15%	29%
Financially literate (score = 1)	39%	7%	54%

Fischer exact test p-value: 0.048

Question 10

This question aims to assess if the mention of no price increase, because it triggers no alternative scenario and gives directly the real increase of wage, has an impact on respondents answers.

Wording

Assume you have neither debt nor savings, and that you spend each month all your 1600 euros monthly salary, but not more, that is you do not contract debt. Which situations would you rather be in (assuming all other things are equal in the two scenarios)?

Scenario A: Prices stay the same in the economy and you got salary increase of 2%. Your salary thus increases from 1600 euros to 1632 euros.

Scenario B: There is 4% inflation in the economy and you got a salary increase of 5%. Your salary thus increases from 1600 euros to 1680 euros.

Scenario A   Scenario B   Both

Results

Subjects exhibit a low sensitivity to money illusion in this question since 82% of them rightly prefer the higher real pay increase of Scenario A to the higher nominal pay increase of Scenario B. It is very likely that, as for Question 1 which was also about wages, providing a clear context allows to reduce the money illusion exhibited in answers, even more here, since the question includes a simpler alternative, in which the real increase of wage was obvious. There is no significant difference in answers across sex type or across education groups, contrary to question 1, which may due to the fact the comparison between both scenarios being easier, all groups of students were more able to give the correct answer.

Question 10	Scenario A (low inflation environment)	Scenario B (high inflation environment)	Both
Which situations would you rather be in (assuming all other things are equal in the two scenarios)?	82%	15%	3%

Question 11

The framing of this question enable us to check if changing the "no event" level of inflation of the previous question, that is, the "no price increase" assumption, while keeping the real situations unchanged, provokes different answers.

Wording

Assume you have neither debt nor savings, and that each month you spend all your 1600 euros monthly salary (but not more, that is, you do not contract debt). Which situations would you rather be in (assuming all other things are equal in the two scenarios)?

Scenario A: There is 1% inflation in the economy and you got a salary increase of 3%. Your salary thus increases from 1600 euros to 1648 euros.

Scenario B: There is 4% inflation in the economy and you got a salary increase of 5%. Your salary thus increases from 1600 euros to 1680 euros.

Scenario A Scenario B Both

Results

Subjects provided answers consistent with previous question. The question provides a robustness check to the previous answers to Question 10. Again, there is no significant difference in answers across education group or sex type.

Question 11	Scenario A (low inflation environment)	Scenario B (high inflation environment)	Both
Which situations would you rather be in (assuming all other things are equal in the two scenarios)?	79%	12%	9%

Nominal loss aversion

The higher degree of nominal loss aversion respondents have, the higher their preference for the high inflation environment, even if the real returns are twice lower in the high inflation environment of Scenario B than in the low inflation environment of Scenario A. This result supports the hypothesis of a relation between both biases of money illusion and nominal loss aversion.

Question 11	Scenario A (low inflation environment	Scenario B (high inflation environment)	Both
Not nominal loss averse	92%	7%	0%
Slightly nominal loss averse	83%	9%	8%
Very strongly nominal loss averse	47%	32%	21%

Fischer exact test p-value: 0.044

Risk aversion

High risk-averse subjects tend to prefer the low inflation scenario more often (87%) than low risk-averse respondents (72%). The difference between both groups, higher than that observed in question 3, cannot be easily interpreted since there is no reason for risk-aversion and money illusion to be inversely related. Further exploration is needed to test this relation as well as other hypothesis like a preference of high-risk averse subjects for low inflation contexts that could bias the results.

Question 11	Scenario A (low inflation environment	Scenario B (high inflation environment)	Both
Low risk adverse	72%	21%	8%
High risk adverse	87%	4%	10%

Fischer exact test p-value: 0.030

Question 12

This question aims to test if subjects are able to compare two different real vs. nominal rates of return in case of inflation from various viewpoints.

Wording

A fund, by buying and selling securities, makes some return for investors.

Scenario A: There is a 1% inflation rate and the fund manages to get a 6% return.

Scenario B: There is a 10% inflation rate and the fund manages to get a 15% return.

Which of the two scenarios do you believe is the best in economic terms?

Scenario A Scenario B Same

Which scenario do you think is the most beneficial for the fund manager?

Scenario A Scenario B Same

Which scenario do you think is the most beneficial for the investors?

Scenario A Scenario B Same

Results

As in question 1, while both scenarios are approximately equivalent in real terms, a precise calculation indicates that the low inflation scenario (A) is slightly better. However, students were more able to compare both scenarios in economic terms since only 13% of them considered that the high inflation scenario (B) was better and can then be considered as exhibiting money illusion. A vast majority correctly chose the scenario A (63%) and the others considered both scenarios as equivalent in economic terms. These results may be due to the fact the comparison deals only with two rates, hence subjects were not influenced by nominal values, in contrast to other questions, like question 1 which refers to nominal wages. A closer look at respondents answers indicate they are similar for all education type.

Nonetheless, a significant proportion of students have changed their answers when they were asked to give an opinion on the scenario in which the fund manager, or alternatively the investors will be happier. About 40% of them considered that the fund manager or the investors will be happier with the high inflation scenario (B). Given that most of these respondents were able to see that scenario was not the better in economic terms, these results suggest that even if individuals are able to compare different situations in real terms, they could prefer (or consider that others could prefer) the one which is better in nominal terms. Again, there is no significant difference in answers across education groups or sex type, except the third question with one significant difference across education type.

Question 12	Scenario A (low inflation environment)	Scenario B (high inflation environment)	the same
Which of the two scenarios do you believe is the best in economic terms?	63%	13%	24%
In which scenario do you think the fund manager will be happier?	40%	42%	18%
In which scenario do you think the investors will be happier?	41%	38%	22%

Education type

Students in economics chose less often the high inflation scenario (B) in the third version of the question than other students, exhibiting then less money illusion. The rest of the results barely differ across education groups.

Question 12 ter: in which scenario do you think the investors will be happier?	Scenario A (low inflation environment)	Scenario B (high inflation environment)	the same
Science	43%	43%	14%
Economics	40%	24%	36%
Law and literature	36%	55%	9%

Fischer exact test p-value: 0.036; Pearson chi2 p-value: 0.0030

Financial literacy

Financial literacy does not have on impact on the answers to the first two questions (Question 12 and Question 12 bis). But Question 12 ter shows that a significant higher proportion of non-financially literate subjects believe that investors will be happier in the high inflation scenario (B), more than half of them can then be considered as exhibiting money illusion vs. about 30% of the financial literate subjects. Of course, the former can have considered that investors would not have the hedge against inflation provided by the fund return but in that case, financially literate students might also have been sensitive to that argument, and clearly they didn't, as indicated by their different answers.

Question 12 ter: in which scenario do you think the investors will be happier?	Scenario A (low inflation environment)	Scenario B (high inflation environment)	the same
Non-financially literate (score = 0)	24%	56%	21%
Financially literate (score = 1)	49%	29%	22%

Fischer exact test p-value: 0.048

Risk aversion

Risk-aversion has an impact on answers. High risk-averse subjects are more likely to prefer the low inflation scenario (74%) than low risk-averse ones (53%). The results are consistent with question 11.

Question 12	Scenario A (low inflation environment	Scenario B (high inflation	the same

		environment)	
Low risk adverse	53%	19%	28%
High risk adverse	74%	8%	18%

Fischer exact test p-value: 0.030

4. **Conclusion**

The results obtained support the hypothesis of a significant prevalence of money illusion bias. Even students enrolled or having a Master degree expressed answers biased by money illusion to most of the questions of the survey. Moreover, any of the individual characteristics collected seem to be an important factor to explain that bias. However, students in economics seem less affected by money illusion than other students in some questions and financial literacy skills seem to have a true negative impact on that bias. Finally, the money illusion exhibited depends on the context and the way questions are formulated. Individuals are less likely to exhibit money illusion when alternatives are easy to compare in real terms and when asked to express a judgment on economic terms than on fairness, luck or happiness. We observe significant asymmetries between inflationary and deflationary contexts, particularly in that last type of judgments.

Education type

Differences observed between education groups were mentioned when appropriate in the previous section. Education type especially matters for money illusion in Question 1, where students in economics manifest clearly less money illusion than others. The question was phrased in economic terms. It also matters for Question 5, Question 8 bis and Question 12 ter, but to a lesser extent. Because the survey contains many questions similar to Question 1, but often without the mention "in economic terms", it is likely that the phrasing of the question had an impact on the answers collected from economists.

For all other questions in the survey, students in economics exhibit the same prevalence of money illusion than other students.

Hence, education type doesn't have an important impact on money illusion bias subjects manifested in their answers to that survey. Even an important economics training (4 or 5 years of higher education in economics) is not sufficient to reduce significantly that bias. This result supports the hypothesis of a strongly rooted bias.

Sex

Differences related to sex were also mentioned when appropriate in the previous section. Except for Question 2, where women preferred the no inflation scenario in larger number than men (25% instead of 8%) making less often the wrong choice of high inflation scenario but also less often the right choice of deflationary scenario, differences between men and women were more pronounced when respondents were asked to give judgment of fairness or luck about wages adjustments on prices and probability of wage increases during inflationary times (Question 3, Question 4, Question 4 bis and Question 6 ter). Even if the answers to these questions might be influenced by the degree of money illusion respondents have, additional factors might be at play.

Hence sex does not a major influence on respondents' degree of money illusion from respondents, although it seems to have a non-negligible impact on some particular types of judgments, invoking fairness or luck.

Financial literacy

Financial literacy, as captured by the ability to correctly discount future payments of the two questions included in the survey, seems strongly inversely correlated with money illusion.

Indeed, the two groups of respondents have statistically significant different answers to Question 1, Question 2, Question 9 bis, Question 12 ter. **Financially illiterate students tend to be more subject to money illusion, which is an understandable result in view of the importance of discounting in assessing real values.**

Answers to Question 3, the probability of wage increase in inflationary times, is also influenced by financial literacy. **Financially literate students provided the correct answer, that is, that there is a correlation between wages increases and inflation, twice more often than non-financially literate subjects.**

Finally, only one question out of the four related to a judgment on satisfaction (Question 2 ter) is influenced by financial literacy. Hence **it seems financial literacy has little or no impact on the perception of happiness linked to higher nominal wages.**

According to these results, the hypothesis of an inverse relationship between financially literacy skills and money illusion bias cannot be excluded. Programs of education devoted to specific financial learning could then help people to think in real, rather than nominal, terms, i.e. to reduce money illusion prevalence in the population.

Nominal loss aversion

Nominal loss aversion matters for money illusion in Question 11 with **a clear positive correlation between the degree of nominal loss aversion and the preference for higher inflation scenarios**. Although the result to Question 11 are clear cut, this correlation between nominal loss aversion and money illusion is mitigated by the fact **there is no statistically significant differences between the answers of subjects with different degree of nominal loss aversion in the other questions, except two of these.**

Indeed, nominal loss aversion is correlated to the answers related to two questions about fairness (Question 4 bis, Question 7 bis). Nominally loss adverse subjects more often fail to assess properly situation of fairness in inflation time, but here again the results are mitigated by the fact no statistically different answers were obtained from nominally loss adverse subjects to the other questions relating to fairness.

Hence, there is insufficient evidence to conclude on a clear positive relation between nominal loss aversion and money illusion. Further exploration is clearly needed to conclude on the nature of the relation between both biases.

Risk aversion

High risk adverse subjects exhibited less money illusion in two questions than low risk averse ones but their answers may be due to a preference for low inflation (Questions 11 and 12). Indeed, these results are mitigated by the fact they less often believe that salary increases are more probable in inflationary times than low risk adverse subjects (Question 3) and moreover by the similar answers of both groups of subjects to other questions related to money illusion.

References

Boes, S., Lipp, M., & Winkelmann, R. (2007). Money illusion under test. *Economics Letters 94*, 332-337.

Bourgeois-Gironde, S., & Guille, M. (2012). Keynes' animal spirits vindicated: An analysis of recent empirical and neural data on money illusion. *Journal of Post-Keynesian Economics, 34(2)*, 331-352.

Cipriani, G. P., Lubian, D., & Zago, A. (2008). Money illusion: are economists different? *Economics Bulletin*, Vol 1, no 3, p 1-9.

Fehr, E., & Tyran, J.-R. (2007). Money Illusion and Coordination Failure. *Economic Behavior*, 58(2): 246-68.

Fischbacher, U. (2007). Z-tree. Zurich Toolbox for Readymade Economic Experiments. *Experimental Economics*.

Fisher. (1928). *Money Illusion.* New York: Adelphi Company.

Holt, C., & Laury, S. K. (2002). Risk Aversion and Incentive Effects. *American Economic Review, 92(5)*, 1644-1655.

Koskela, E., & Sullström, R. (n.d.). Some tests for money illusion and distribution effects in the aggregate consumption function. *Empirical Economics* , Volume 4, Number 4, 221-234.

Lian, P., & Plott, C. R. (1998). General Equilibrium, Markets, Macroeconomics and Money in a Laboratory Experimental Environment. *Economic Theory*, 12(1): 21-75.

Mishra, H., Mishra, A., & Nayakankuppam, D. (2006). Money: A Bias for the Whole. *Journal of Consumer Research: An Interdisciplinary Quarterly, University of Chicago Press*, vol. 32(4), pages 541-549, 03.

Miyapuram, K. P., Tbler, P. N., Gregorios-Pippas, L., & Schultz, W. (2009). Money Illusion in the human brain. *Proceedings of the 31st annual meeting of the Cognitive science society, Amsterdam*, 1657-1662.

Noussair, C. R. (2001). Price Bubbles in Laboratory Asset Markets with Constant Fundamental Values. *Experimental Economics*, 4(1): 87-105.

Odean, T. (1998). Are Investors Reluctant to Realize Their Losses? *Journal of Finance*, 53(5): 1775-98.

Oechssler, J., Schmidt, K., & Schnedler, W. (2007). Asset Bubbles without Dividends. *University of Heidelberg*, Working Paper no. 439.

Shafir, E., Diamond, P., & Tversky, A. (1997). Money Illusion. *The Quaterly Journal of Economics, vol 112, no. 2, In Memory of Amos Tversky*, 341-374.

Tversky, A., & Kahneman, D. (1981). The Framing of Decisions and the Psychology of Choice. *Science*, New Series, Vol 211, No. 4481, pp 453-458.

Weber, B. R. (2009). The medial prefrontal cortex exhibits money illusion. *PNAS*, 106(13), 5025-5028.

Wong, W.-K. (2005). Does money illusion still matter after some economics education?

Summary statistic tables of exact Fisher test p-values by theme

		Effect of :				
Theme :	Questions of the Survey	education type	sex	financial literacy	nominal loss aversion	risk aversion
Money illusion in salary increase, economic	Question 1	**0.020**	0.255	**0.046**	*0.601*	*0.819*
	Question 6 bis	0.280	*0.436*	0.146	*0.406*	*0.780*
	Question 9 bis	*0.373*	*0.883*	**0.048**	0.136	0.136
	Question 10	*0.581*	*0.761*	*0.425*	0.262	*0.695*
	Question 11	*0.802*	*0.496*	0.128	**0.003**	**0.030**
Money illusion in house or financial asset returns increase, economic	Question 2	*0.990*	*0.325*	**0.027**	*0.798*	*0.780*
	Question 2 bis	*0.944*	**0.063**	0.187	*0.875*	*0.662*
	Question 12	*0.724*	0.367	0.223	*0.735*	**0.078**
	Question 12 bis	*0.819*	0.265	*0.735*	0.249	*0.361*
	Question 12 ter	*0.036*	*0.823*	**0.020**	*0.516*	*0.773*
Satisfaction	Question 2 ter	*0.979*	*0.616*	**0.047**	*0.907*	*0.966*
	Question 2 quar	*0.596*	0.285	*0.331*	*0.417*	*0.473*
	Question 6	*0.805*	*0.949*	*0.943*	*0.438*	0.177
	Question 9	*0.843*	*0.656*	*0.522*	*0.806*	*0.624*
Probability salary increase	Question 3	*0.873*	**0.011**	**0.012**	*0.513*	**0.078**

in inflation time						
Luck in salary increase	Question 4	*0.423*	**0.025**	*0.396*	*0.338*	*0.657*
	Question 5	*0.627*	*0.851*	*0.650*	*0.648*	*1.000*
	Question 7	0.255	**0.075**	*0.401*	*0.581*	*1.000*
	Question 8	*0.370*	*0.706*	*0.331*	*0.878*	*0.595*
Fairness	Question 4 bis	0.276	**0.006**	*0.607*	**0.008**	*0.610*
	Question 5 bis	**0.006**	*0.940*	*0.810*	*0.831*	*1.000*
	Question 6 ter	*0.731*	**0.066**	*0.826*	*0.547*	*0.361*
	Question 7 bis	0.114	*1.000*	*0.831*	**0.044**	*0.580*
	Question 8 bis	**0.044**	*0.575*	*0.729*	*0.622*	*0.580*
	Question 9 ter	*0.599*	*0.817*	*0.574*	0.277	*1.000*

NB: Fisher exact test p-values of *more than 30% are in italics*, those **less than 10% are in bold**, indicating statistically different answers for the question at a 10% significance level.

Article 2: The validity and time-horizon of the Fed model:

a co-integration approach[1].

1. Introduction

The so-called Fed model for valuing stocks, which became widely popular after the release of the Humprey-Hawkins report to the Board of Governors of the Federal Reserve System (1997), has been and remains a widely controversial model among both scholars and practitioners. This simple model stipulates that comparing the ratio of the earnings over the price of the stock market to the yield of long-term government bonds indicates whether the stock market is overvalued or undervalued. The main idea is that bonds and stocks being competitive investments, their expected return – at least when adjusted for the perceived risk of each asset class – should be set equal by profit-maximizing investors. We will, throughout this article, denote by Y the yield on long-term government bonds and by E/P the earning yield of stocks, that is, the ratio of earnings E divided by the price level of the stock market P. Let us make the bold assumption, for simplicity of illustration, that both asset classes are perceived as incurring the same risk by market participants[2]. Then the Fed model argues that the two quantities Y and E/P will be set equal by rational, return-oriented investors. For example, if it happens that Y<E/P, it would mean that stocks are cheap (relative to bonds), and investors will sell bonds and buy stocks, pushing the price P of the stocks higher until the E/P ratio is again set to the level of Y. We will refer to the Fed model that asserts a strict equality of Y and E/P as to the "strict" Fed model. Adjusting for perceived risks or other factors can yield other Fed models: the common idea of those "general" Fed models is that the position of E/P *relative to Y* is a pertinent indicator for valuing stocks. Such models can include adjustments for perceived risks, as the one proposed by Asness (2003) which uses prior realized volatility on stocks and bonds as a proxy for the risk perceived by investors.

From a theoretical perspective, it can be objected that E/P is not the return for investing in stocks, nor is the yield to maturity the return for investing in bonds (except if the bond is held to maturity and if all coupons can be reinvested at the same rate). But the main criticism made to the Fed model is that it compares a real quantity, E/P, to a nominal quantity, the government bond yield Y, and thus assumes that economic agents are victims of money-illusion, not being able to properly distinguish between real and nominal values. Indeed, stocks are claims on real assets, which are expected to appreciate with inflation, whereas the government bond yield is clearly a nominal quantity. Hence not only the Fed model would lack any firm theoretical grounding, but it is in flagrant contradiction with the

[1] This paper has been accepted by the *Journal of Empirical Finance*, September 2014, code: EMPFIN-S-14-00287.

[2] This hypothesis is clearly unrealistic as bonds, and in particular government bonds, are usually assumed less risky than stocks.

hypothesis of efficient markets[3]. One of the most vivid criticisms of the Fed model is certainly the one made by Asness (2003). Asness asserts the Fed model is correct as a *descriptive* tool – documenting how investors set the prices of stocks - but performs poorly as a *predictive* tool of subsequent returns. According to him investors are consistently making the same error, and the Fed model documents this error. This seems a paradox: how can a model which correctly describes the level to which investors set the price P of the stock market be denied of any merit as a trading strategy? One element of answer could be the time horizon. It would be possible that the Fed model allows predicting correctly the future level of prices P of the stock, but without specifying the time horizon correctly. For example, and to sustain Asness claim, it could be that the time horizon is excessively short compared to the longer term horizon of the buy and hold strategy which Asness assumes when he shows that investing according to the signal offered by the Fed model actually incurs huge losses compared to, for example, a passive "buy and hold" investment strategy in bonds. Also consistent with this explanation are the findings of Owain ap et al. (2001). These authors study the historical in-sample performance of the Fed model as an tactical asset allocation tool for six different countries (the United States, the United Kingdom, France, Germany, Switzerland and Japan) and conclude that for short-term tactical allocation, the Fed model has some merit as a relative valuation tool, whereas over a 5 year horizon it underperforms the more traditional indicators, like the E/P ratio, in each country. Together with Asness' own long-term study of the returns this suggests that the time horizon of the Fed model should be less than 5 years. Interestingly, *no study trying to estimate which minimum time period is required for applying the Fed model with success has been carried out to our knowledge*. In this paper we propose a methodology based on in and out of sample comparisons of different model predictions to do so. The existence and predictability of a horizon for the Fed model would have huge practical and theoretical consequences. We use daily, end of the day, weekly and monthly, either end of the period or average over the sampling frequency period, data to assess the possible time horizon of the Fed model over a longer period and for different countries.

Because the Fed model posits the existence of a linear relationship between the earning yield and the nominal bond yield, linear regressions would be the natural candidate for evaluating the Fed model, as used by Asness (2003). Nevertheless, unit-root tests indicate the series are non-stationary (independently of the frequency). Hence, a full co-integration approach would be more methodologically robust. More precisely, in Section 2 we explain why non-stationarity of what we call "Fed model indicators" implies the Fed model cannot hold. This is because, when these indicators are not stationary, any deviations to their assumed fair level do not decrease over time. Because the stationarity of these indicators depends on the time-horizon and on the time-period, stationarity tests

[3] For example, in times of higher inflation, we would expect E/P to be roughly stable, since both the earnings E and the price P will co-move with inflation, whereas Y will move close to one-to-one with inflation, as accredited by numerous studies concerning the so-called "Fisher effect".

on the Fed indicators are carried out at different frequencies and for different sub-samples. Section 3 studies further the predictive ability of the model for the United States and the Standard & Poor's 500 index (S&P 500). It uses both the Engle-Granger and the Johansen co-integration framework to build Vector Error Correcting (VEC) models. To compare the predictive value of the linear, long-term part of the model, it is compared in terms of performance to a simple Vector Auto-Regressive (VAR) model of similar size in terms of explanatory variables. Because the only difference between the two models is the long-term relationship (the "error-correcting term") when the VAR outperforms the VEC according to different out-of-sample measures, it shows that this long-term linear relationship does not have any ability to predict future yields. It is also quite possible that, for example, the VAR would outperform the VEC for only one of the two variables forecast. For example it could be that the VAR outperform the VEC for the nominal government yield but not for the earning yield. In that case we will conclude that the Fed model only has predictive power for the earning yield, and such results depend on the frequency at which the analysis is performed. Section 4 tries to determine the time horizon of the Fed model, also for the United States and the Standard and Poor's, and uses Engle-Granger methodology and the same methodology as Section 3, while Section 5 and concludes.

2. International evidence for the strict Fed models

The data used in this section consists of the ten year generic government bond yields of different countries, as reported by Bloomberg, and of the earnings ratio E/P where E is defined as the trailing 12 months earnings of the enterprises quoted in some stock index of the country under study, and P is the price level of this index. List 1 in the Annex indicates the stock exchanges earning-ratios studied. Admittedly, the choice of the trailing earnings for E is rather arbitrary[4]. The source of our data is Bloomberg, and our dataset extends from the first of January 1963 to the 24 of March 2011 included. The frequencies of the data used here are daily, end of the day, monthly, end of the month and monthly, daily average of the month.

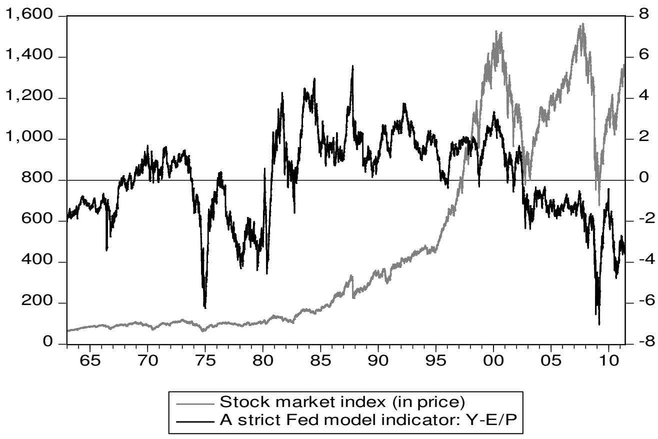
Evaluating the strict Fed relationship

What we defined in the introduction as the *strict Fed model* asserts equality between the earning yield and the government yield. If the strict Fed model is right and this is indeed the long-term relationship between the two quantities, then taking the difference or the log-difference of the two series should result in a stationary series. Indeed if the resulting series has a unit root, then the equilibrium is meaningless, since any disequilibrium to the relation Y=E/P does not tend to correct over time, but

[4] Expected forward earnings would be a most appropriate candidate for E. But because market expectations of forward earnings are rather imprecise, and in any case much influenced by recent past earnings, and also because of the non-availability of such series for most countries, we decided to take as a proxy for the future earnings the past, realized earnings, thus following Asness (2003) among others. Asness actually considered the ten-year trailing earnings, but we believe taking ten year trailing earnings is definitely too backward-looking, in particular considering the sharp volatility in earnings than the stock exchange has experienced in our sample periods. It would be, of course, interesting to use other proxies for the expected future earnings, like the aggregate market consensus, when such data is available, and compare results.

exhibits the highest degree of persistency. We can thus define different strict *Fed model indicators* and assess their stationarity. The most obvious indicator is the simple difference indicator Y-E/P. When it has a large negative sign, it coins that stocks are cheap relative to bonds, hence rational investors seeking to maximize returns will bid up stock prices and the indicator will return towards 0. If the indicator is positive, then stocks are too expensive relative to bonds and the inverse will happen. Figure 1 below draws this indicator together with the stock index level[5].

Figure 1: simple strict Fed indicator for the S&P 500



For the United States and the S&P 500, the indicator did give a strong buying signal in the 1973 to 1980 period, which preceded the formidable increase in stock prices of the 1980 to 2000 period, but it also gave a strong selling signal from 1980 – or at least, from 1985, to 1995. An investor making investment decisions accordingly would have missed the run-up in equities of that period. Since 2002 the indicator has consistently given a buying signal, thus being unable to predict the stock market crash of 2008. Still, visual inspection also reveals it has correctly started giving a buying signal in 2002, precisely when the stock market had bottom from its previous 2000 crash. It also gave a strong correct buying signal in 2009, but failed to predict the 2007 crash. We leave to the reader to try to draw eventual long-term trading rules from the data presented in Figure 1, but our overall sentiment is that visual inspection does not seem to indicate a strong and timely predictive power for this simple strict Fed indicator. Other possible strict Fed model indicators include the log difference log(Y)-log(E/P)= log(Y/(E/P)), and the simple ratio Y/(E/P). Similar observations can be drawn for these other strict Fed indicators. Of course, one could always try to design trading rules, and see if it incurs any more gains compared to some other strategies. Since any Fed indicator assumes the relevance of looking at both the government bond yield and the earning yield while assessing the level of stock

[5] The graph of this indicator for each other pair of country and stock market studied can be found in the document "Additional supporting evidence related to the article: 'the validity and time-horizon of the Fed model: a co-integration approach'", available from the author on request. The least we can say is that evidence is not overwhelming.

prices, such a trading strategy should, in any study, be compared to the one using only the earning yield E/P as an indicator, since it is well known that the earning yield does have some predicting power for future stock returns. The problem is that there are an infinite number of possible trading strategies, depending on what we define to be a trading signal from the indicator, and the time-horizon to move in and out of the stock market[6]. Hence we need a formal test that would confirm or invalidate the relevance of strict Fed model indicators. As mentioned earlier, a necessary condition is certainly the *stationarity* of the indicator. An indicator whose realization points towards a unit-root process is certainly not relevant: it means there are no forces that push the indicator back to equilibrium, and thus past shocks have a permanent effect on its level: not very pleasant since our trading signal involves comparing the indicator to the 0 level (or to any bands around 0, depending on how one defines a trading signal). If past shocks have a permanent effect on the indicator, then comparing it to any fixed band is irrelevant.

For each of the three different strict Fed model indicators mentioned we perform Augmented Dickey-Fuller (ADF) tests and Philipps-Perron tests with daily and monthly data, for each of our 24 samples. When the time-series is long enough, we also perform ADF tests on sub-samples. Table 1, 2 and 3 in Annex recapitulate the results, reporting the MacKinnon p-values corresponding to the t-statistics[7].

The results are that the null of a unit root is rejected, at a 10% significance level and for all three indicators, for only four out of the 24 samples: the United States (S&P 500), Mexico, Italy and Australia. In the most recent period, from 1993 to 2011, none of the three indicators pass the test for the United States and the S&P 500: the tests indicate, for this subsample, very large p-values of more than 0.40. Not surprisingly, when looking at monthly data it is precisely the same countries which pass the tests. With only one exception (Canada), the *relative* strict Fed model indicator, that is, the second and third one, which involves the *quotient* of Y by E/P in log form or in simple form[8], passes much more often the stationarity tests than the first absolute indicator, which assumes stationarity of the absolute change level in yields Y-E/P.[9]

---

[6] This is a non-trivial problem: because of this infinite number of possible trading rules, one will always end up finding a rule which will work in-sample (that is, ex-post). To conclude that the Fed model is valuable on the sole basis of such a finding would thus be methodologically incorrect.

[7] The lag for all the ADF tests was chosen as to maximize the Schwarz Information Criteria, with a maximum of up to 25 lags. Philipps-Perron tests were performed using the Newey-West Bandwith to select the bandwith. Table 1 presents the result of the tests using daily, end of the day data, whereas Table 2 and Table 3 present the monthly, end of the month, and average of the month results. Notice that looking at both daily and monthly frequency adds really a new dimension to the problem, since the maximum lags used in ADF tests at daily frequency being 25, these daily stationarity tests necessarily have a shorter-time horizon than the monthly ones, for which the time lag is at minimum one month, and at maximum 25 months.

[8] Of course, since (Y-E/P)/(E/P)=Y/(E/P) – 1 and that adding a constant does not change the results of the stationarity tests, this indicates that changes relatively to the current level of earning yieds are usually more stationary than changes in the absolute levels.

[9] Indeed, at both monthly and daily frequency the third Fed indicator is stationary for the United Kingdom, Germany, and Japan, whereas the first indicator fails to satisfy stationarity in these three countries. All in all, it is now 7 out of 24 samples which pass the basic test of stationarity at both the daily and monthly frequency if we look at the third indicator instead of the first. This still leaves 17 samples in which the null of a unit-root process is not rejected, that is, the indicator seems to have a stochastic trend, hence behaving more like a random walk, with no economic forces pushing it back to 0 when shocks makes it deviate from this assumed “fair” value.

To conclude, for most markets, including the United States (S&P 500) from 1993 to March 2011, the Dow Jones Industrial and the NASDAQ, the strict Fed model fails to pass the most basic test: the so-called Fed indicator has probably a stochastic trend, hence its current level, partly due to the accumulation of *past* previous shocks, is irrelevant to assess *current* the over or under-evaluation of the stock market relative to bonds. Our results are thus in line with those of Gwilym et al. (2001) and Asness (2003). The (strict) Fed model does not travel well, nor does it fit the United States (S&P 500) for most recent periods, nor does it seem relevant for the other US stock indexes.

Now the strict Fed model indicator seems stationary for the United State and S&P 500 for the whole sample period, for Mexico, Italy and Australia. We can assess the stationarity of the strict Fed indicator Y-E/P more precisely on each sub-sample by looking at rolling ADF tests. The statistical results are detailed in the supporting documentation downloadable from the author website[10].

They indicate that for some five-year periods, in particular the 2002 to 2008 period, the simple strict Fed indicator Y-E/P is not stationary, suggesting it cannot have been a good indicator during that period. The test with the relative strict Fed indicator gives similar results.

3. Estimated Linear Fed Models: performance evidence

Outline of the methodology

As coined in the introduction, a general Fed model long-term relation could have any other form than a strict equality, in so far it relies on comparing the earnings-to-price ratio E/P to the earning yields Y. A particular, yet more general case than a *strict* equality between the two ratios would be *linear* general Fed models, in which the assumed long-term relationship is:

$$Y = b * E/P + b2, \text{ with } b>0 \text{ and } b2 \text{ two constants}$$

This linear relation could be interpreted in a way similar to those of general Fed model coined in the introduction which takes into account perceived relative risks between bonds and stocks: for example, the above relation with b2=0 would mean that stocks are deemed riskier than bonds if 0<b<1, and less risky than bonds if b>1. A negative b would be contradictory to the Fed model as it implied opposite dynamics for the earning yields and the government yields.

We study in this section the relevance of the linear general Fed model for the United States and the Standard and Poor's, using data from different frequencies, and within a co-integration theory framework as defined by Engle and Granger (1987) and using maximum likelihood methods as

[10] See the document "Additional supporting evidence related to the article: 'the validity and time-horizon of the Fed model: a co-integration approach'", downloadable from the author's website: http://fabienecb.fr/menurapport2.html.

defined by Johansen, Stock and Watson (1995). This allows to build various Vector Error Correcting (VEC) models, whose merits as a quantitative version of the Fed model we assess by comparing the results of their predictions to those of similar models without error-correcting terms, like those of a simple Vector Autoregressive Models, in levels or in first difference, or to the naïve predictions of the explained variable lagged one period. The idea is to disentangle how much the Fed indicator actually improves forecasts when compared to these different alternative models which do not contain any long-term relationship between yields, earnings and stock market prices.

Visual inspection: a changing relation

The graphs below depicts the very strong (visual) relation between the government yield and the earning yield of the S&P 500 over the whole sample period (graph on the left), and the price level of the S&P 500 (graph on the right).

Figure 2: time series of Y, E/P and P for the US and the S&P 500.

Fig 2A: US government yields and S&P 500

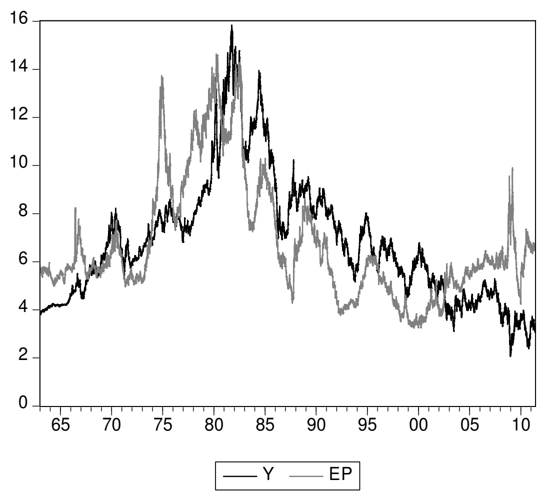
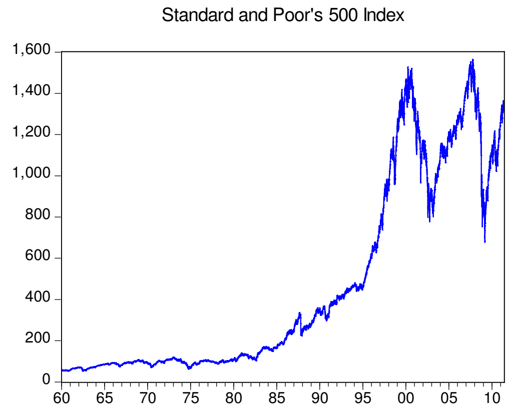


Fig 2B: S&P 500 in level

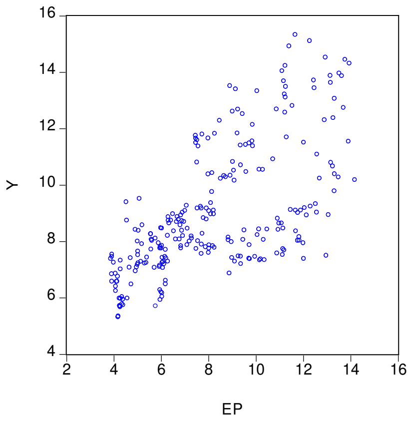


A first remark is that by looking at shorter sub-sample, the relationship between the two series is far from obvious. The graphs below represent scattered plots of the 10 year government bond yield against the earnings ratio of the S&P 500, for six different periods. They were obtained by using monthly data corresponding to the average of the daily observations of each month. For some period, as for the 1985 to 1995 period, the two series seem to move together in the same direction, whereas for other periods, for example the 1975 to 1985 period, there does not seem to be any clear pattern. For the 2000 to 2011 period, the two series even seem inversely correlated, in plain contradiction with what the Fed model would predict, but in accordance with the “flight to safety” phenomenon, very strong in that period: afraid of stock returns, investors would flee the stock market and bid government bond prices up.

Figure 3: Scatter plots of Y and E/P for the US and the S&P 500 on different sub-samples

Fig 3A : 1963-1975

Fig 3B : 1975-1985



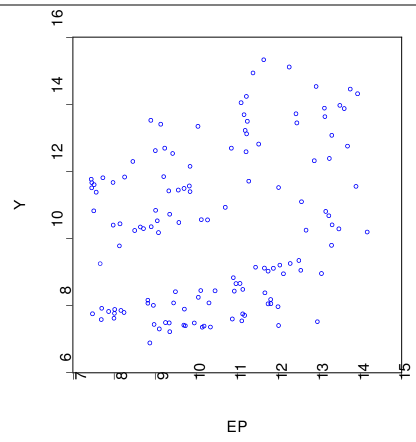
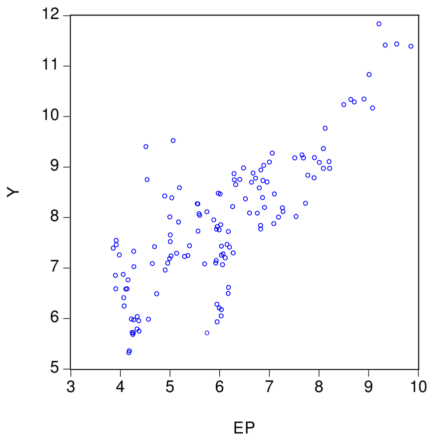


Fig 3C : 1985-1995

Fig 3D : 1995-2005



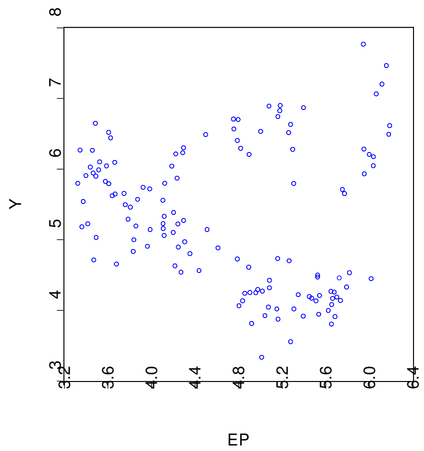
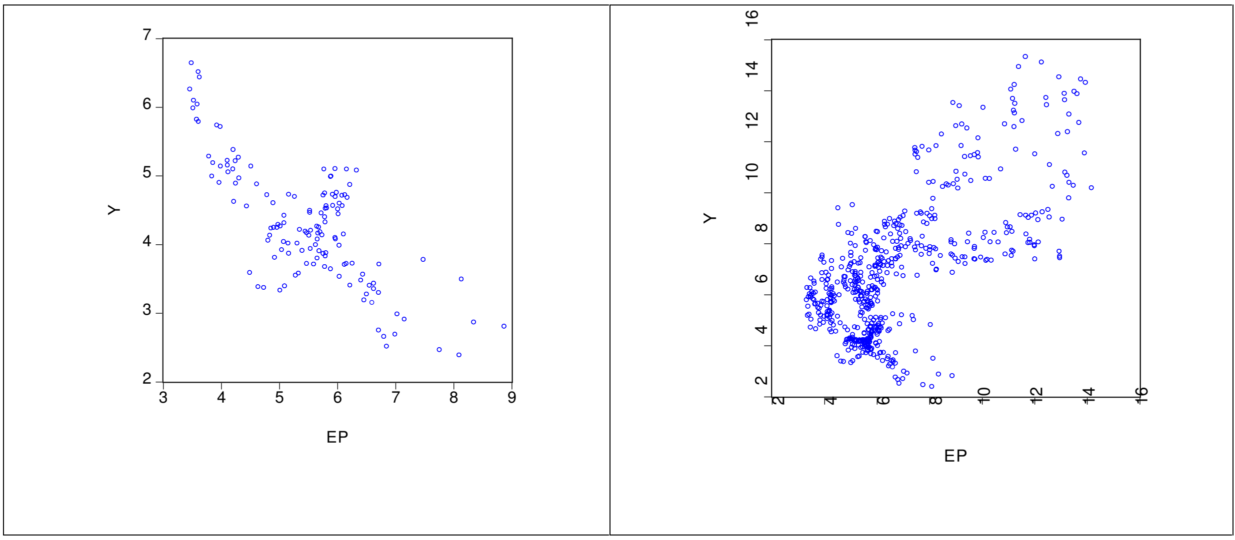


Fig 3E : 2000-2011

Fig 3F : 1963-2011



Unit root tests for the government yield and the earning yield series

Johansen co-integration tests assume that the involved series are integrated of order one, hence we have to make sure this is the case, otherwise the tests are meaningless. We thus apply unit root tests to each of the two series Y and E/P. The detailed statistics results of these routine tests can be found on the author's website[11].

Tests at different frequencies, and for different sub-periods

For the sample from the first of January 1963 to the 24 of May 2011, Johansen tests performed with either daily, weekly or monthly data do not reject the null of no co-integration at the 10% significance level, for any lag length selected between 1 and 25.[12] The table below reports the Trace statistics and the Maximum Eigenvalue statistics of this test for the whole sample, together with the corresponding MacKinnon-Haug-Michelis (1999) p-values:

Table 1A: Johansen Co-integration Rank Test: Trace statistics

Number of co-integration relationships assumed in the null hypothesis of the test	Trace statistics	0.05 critical value	p-value
None	11.12764	15.49471	0.2038
At most 1	1.807713	3.841466	0.1788

Table 1B: Johansen Co-integration Rank Test: Maximum eigenvalue statistics

Test of the null of r co-integrating relationships against the alternative of r+1 co-integrating relationships	Trace statistics	0.05 critical value	p-value
None	9.319930	14.26460	0.2605
At most 1	1.807713	3.841466	0.1788

[11] These results are reported in the Annex 4 of the document quoted in footnote 5, page 6.

[12] For example, for daily data and allowing 4 lags, an intercept but no deterministic trend in the co-integrating relation and a trend in the VAR gives a Trace statistics of 11.127 corresponding to a p-value of 0.2039 for the null hypothesis of no co-integrating relationship.

Dismissing the results of the tests and using co-integration techniques (Engle-Granger two steps method and Johansen maximum likelihood method) gives, unsurprisingly, inconsistent and unstable results[13].

Now looking at the trends for the two series (see first chart of Figure 2) it is clear that we can separate our sample into two sub-samples so as to allow the Johansen test to capture the two deterministic trends better: the first linear deterministic trend is positively sloping from 1963 to 1980 whereas the second, from 1980 to 2011, has a negative slope.

We get, for the same tests as previously – in particular, using daily data – the following results for the two periods:

Table 2A: Johansen Co-integration Rank Test for the 1963 to 1980 period: Trace statistics

Number of co-integration relationships assumed in the null hypothesis of the test	Trace statistics	0.05 critical value	p-value
None	8.220717	15.49471	0.4421
At most 1	0.104091	3.841466	0.7470

Table 2B: Johansen Co-integration Rank Test from 1963 to 1980 period: Maximum eigenvalue statistics

Test of the null of r co-integrating relationships against the alternative of r+1 co-integrating relationships	Trace statistics	0.05 critical value	p-value
None	8.116626	14.26460	0.3670
At most 1	0.104091	3.841466	0.7470

Table 3A: Johansen Co-integration Rank Test for the 1980 to 2011 period: Trace statistics

Number of co-integration relationships assumed in the null hypothesis of the test	Trace statistics	0.05 critical value	p-value
None	15.99375	15.49471	0.0420
At most 1	1.100870	3.841466	0.2941

Table 3B: Johansen Co-integration Rank Test for the 1980 to 2011 period: Maximum eigenvalue statistics

Test of the null of r co-integrating relationships against the alternative of r+1 co-integrating relationships	Trace statistics	0.05 critical value	p-value
None	14.89288	14.26460	0.0397
At most 1	1.100870	3.841466	0.2941

Results for both periods are robust relative to the frequency used. The difference between the two sub-samples is striking: in the first sub-sample period the null of no co-integration is accepted with a high

[13] See the document quoted in footnote 5, page 6.

p-value of 0.44, whereas in the second the null of no co-integration is rejected at the 5% significance level, for daily data, and at a 10% significance level for weekly and monthly data.
The tables below detail the results for the 1980 to 2011 period, presenting the Trace statistics and the Maximum Eigenvalue statistics at each different frequency[14].

Table 4A: Johansen Co-integration Rank Test for the 1980 to 2011 period: Trace statistics

Number of co-integration relationships assumed in the null hypothesis of the test	Frequency of data used	Trace statistics	0.05 critical value	p-value
None	daily end of the day	15.99375	15.49471	0.0420
	weekly end of the week	14.29764	15.49471	0.0751
	weekly, daily average	14.01039	15.49471	0.0827
	monthly end of the month	15.46071	15.49471	0.0506
	monthly, daily average	15.83884	15.49471	0.0444
At most 1	daily end of the day	1.100870	3.841466	0.2941
	weekly end of the week	1.915946	3.841466	0.1663
	weekly, daily average	2.079140	3.841466	0.1493
	monthly end of the month	1.910101	3.841466	0.1670
	monthly, daily average	1.947226	3.841466	0.1629

Table 4B: Johansen Co-integration Rank Test for the 1980 to 2011 period: Maximum eigenvalue statistics

Number of co-integration relationships assumed in the null hypothesis of the test	Frequency of data used	Trace statistics	0.05 critical value	p-value
None	daily end of the day	14.89288	14.26460	0.0397
	weekly end of the week	14.29764	15.49471	0.0751
	weekly, daily average	11.93125	14.26460	0.1133
	monthly end of the month	13.55061	14.26460	0.0646
	monthly, daily average	13.89161	14.26460	0.0572
At most 1	daily end of the day	1.100870	3.841466	0.2941
	weekly end of the week	1.915946	3.841466	0.1663
	weekly, daily average	2.079140	3.841466	0.1493
	monthly end of the month	1.910101	3.841466	0.1670
	monthly, daily average	1.947226	3.841466	0.1629

So far, we made certain that our two series, the government bond yield and the earnings ratio, were integrated of order one, using both augmented Dickey-Fuller tests and Phillips and Perron tests. We now see there is also strong evidence of co-integration for the 1980 to 2011 period, in particular at the daily and at the monthly frequency. We are thus in a position to follow the Johansen method for writing down a Vector Error Correction model for the two variables Y and E/P. We chose to build a VEC with 4 lags[15], in order to get both a parsimonious model yet uncorrelated residuals.
The Johansen maximum likelihood methodology gives the following co-integration relationship:

[14] The p-values suggest there is more evidence of co-integration at the daily level than at lower frequency levels. This is in line with the results of the study of the strict Fed indicators of the fourth section, where the p-value for stationarity of the strict Fed indicators were, most of the time, higher for daily frequencies than for monthly.

[15] The co-integration relation seems rather stable with respect to the number of lags selected. For example, writing a VEC with 25 lags give the long term relationship: Y= 1.84161717598*E/P + 4.70640290327

$$Y = 1.8519020924*E/P + 4.77205658219$$

Interestingly – and surprisingly – following Engle-Granger method for the same period gives a different co-integrating relationship:

$$Y = 0.900439413701*E/P + 1.30149920728$$

Hence the two usual methods for estimating a VEC do not, in our case, give comparable results: the maximum likelihood method to estimate the long-term relationship used in the Johansen approach gives radically different results than the ordinary least square method used by Engle and Granger.

Now, the historical stability of the relation is another problem, and so is its economic interpretation. Concerning the Johansen estimation of the long-term relationship, it is rather dubious that in the long term an increase of 1 unit of the earning to price ratio would be followed by an increase of 1.84 unit of the ten-year government yield. The Engle-Granger method[16] makes more sense economically: an increase of one unit of the earnings ratio would translate, at equilibrium, into an increase of 0.9 unit for the yield Y. If we restrict our sample from 01/05/2011 to 25/05/2011, estimating the VEC using the Johansen approach gives the following relation:

$$Y = - 0.775192725401*E/P - 8.54556842774$$

Whereas the Engle-Granger approach gives the following:

$$Y = -0.601155146767*E/P + 7.59160572346$$

The change of sign shows how much the "long term relationship" depends on the sample. This result confirms our visual inspection of the scattered plots in the beginning of this section. Of course, there could be many possible structural factors explaining this change of the relationship. But maintaining that there is a stable relationship between earnings ratio and yield, as does the Fed model, is contrary to evidence. Moreover, the sign is the opposite of what the Fed model suggests. This could be the effect of the so-called "flight to safety" where investors sell stocks to buy bonds, and hence, if earnings are more or less constant, E/P would go up while Y would go down. But if this explanation could hold for relatively short periods of stock market distress, it is hard to see why it would be the case for the whole first January 2000 to 24 March 2011 period.

Historical stability of the long-term relationship

As noted earlier through visual inspection of the scattered plots, and in the previous paragraph, the long-term relationship does not seem particularly stable from an historical perspective. Table 5 below indicates the different long-term relationship of VEC models with 9 lags estimated during the mentioned period (the results are quite stable with respect to the number of lags):

Table 5: co-integration relationship estimations, with associated speed of adjustment

[16] Recall the first step of Engle-Granger is a simple an ordinary least square regression, as opposed to Johansen maximum likelihood approach.

Series	Co-integrating relation Y=-b.E/P-b2		Speed of adjustment[17]	
	b	b2	of Y	of E/P
1/01/1980 to 1/01/1985	-0.273985	-9.400664	-0.002821	0.007262
1/01/1985 to 1/01/1990	-1.427755	1.260989	0.001691	0.004768
1/01/1990 to 1/01/1995	-1.129606	-1.494971	-0.003522	0.001464
1/01/1995 to 1/01/2000	0.029319	-6.182875	-0.004146	0.001351
1/01/2000 to 1/01/2005	1.087241	-9.775998	-0.008214	-0.004656
1/01/2005 to 1/01/2010	1.078404	-10.68921	-0.004732	-0.002102
1/01/2005 to 24/05/2011	1.521183	-13.30656	-0.002583	-0.003555
1/01/1980 to 24/05/2011	-1.922247	5.221106	0.000127	0.000876

The "long-term" relationship is very dependant of the sample selected, and in some cases the series involved does not even error-correct in the estimated VEC. A possible way to visually examine the stability (or the instability) of such a changing relationship is to plot the coefficients of successive (rolling) estimations of the long-term relationship[18].

We conclude that contrary to what asserts the Fed model, the long-term relationship is not stable. In particular the sign of b, for the most recent periods, is the opposite of what theory underlying the Fed model would suggest.

Hence there is no evidence of a stable relationship between earnings yield and bond yields. We now turn to the evaluation of the general linear Fed model as a predictive tool. Could the Fed model, despite the evidence of instability in the relationship, be used as a predictive tool for timing the stock market? We will, in particular, look if allowing the long term relationship to change continually gives better forecasts than those of other, simpler models.

<u>In sample evaluation of the historical predictive power of the model</u>

We evaluate, on historical data, the predictive power of the above VEC model for the earning price ratio and the government yield. To do so we compare our VEC model with other models, in particular with the simple model obtained by lagging the explained variable itself, which corresponds to naïve predictions one step ahead.

There is no unanimity relatively to knowing whether or not VAR models can be written in level for non-stationary models. Sims (1980) and Sims, Stocks and Watson (1990) recommended against differencing even if the variables contain a unit root. We compare our VEC model with such a VAR in level, with the same number of lags. We use Root Mean Square Errors (RMSE) and the number of correctly predicted directions of change to compare these models.

---

[17] The speed of adjustment is the coefficient in front of the long term relationship in the VEC obtained following Johansen methodology. The higher it is, the more impact the long-term relationship.

[18] See the document quoted in footnote 5, page 6.

Now there are two possibilities for comparing these models on different sub-samples: one can either estimate the models a single time for the whole sample, from the first of January 1980 to the 24 of March 2011 and then look at RMSE and correctly predicted changes for different sub-samples (Table 6), or re-estimate the coefficients of the models for each sub-sample under study (Table 8).

Table 6: models were estimated using the whole 01/01/1980 to 24/05/2011 period

Series: Y	VEC model	VAR model in level	Lagged variable	VEC model	VAR model in level	Lagged variable
Indicator:	RMSE	RMSE	RMSE	Correct changes	Correct changes	Correct changes
1/01/1980 to 1/01/1985	**0.198081**	0.198116	0.198154	623	629	**786**
1/01/1985 to 1/01/1990	0.197379	**0.197364**	0.197453	602	605	**759**
1/01/1990 to 1/01/1995	0.203514	**0.203502**	0.203507	619	641	**772**
1/01/1995 to 1/01/2000	**0.205790**	0.205798	0.205822	632	630	**768**
1/01/2000 to 1/01/2005	**0.204308**	0.204312	0.204301	633	624	**755**
1/01/2005 to 1/01/2010	0.182645	**0.182641**	0.182695	636	632	**786**
1/01/2005 to 24/05/2011	0.160218	**0.160216**	0.160260	801	788	**984**
1/01/1980 to 24/05/2011	**0.067971**	0.067979	0.068068	3910	3917	**4822**
Series: E/P	VEC model	VAR model in level	Lagged variable	VEC model	VAR model in level	Lagged variable
Indicator:	RMSE	RMSE	RMSE	Correct changes	Correct changes	Correct changes
1/01/1980 to 1/01/1985	**0.203602**	0.203670	0.203885	648	641	**786**
1/01/1985 to 1/01/1990	0.203703	**0.203656**	0.203876	596	595	**739**
1/01/1990 to 1/01/1995	0.203752	0.203743	**0.203729**	575	578	**748**
1/01/1995 to 1/01/2000	0.203760	0.203748	**0.203739**	641	645	**789**
1/01/2000 to 1/01/2005	0.203672	**0.203634**	0.203675	608	609	**747**
1/01/2005 to 1/01/2010	**0.203720**	0.203868	0.204408	571	573	**735**
1/01/2005 to 24/05/2011	**0.180216**	0.180354	0.180758	742	750	**961**
1/01/1980 to 24/05/2011	**0.081274**	0.081321	0.081684	3810	3818	**4768**

We can see that the Fed model performs poorly compared to the lagged variable or a simple VAR in level. We would like to be more precise and assess the usefulness of *the co-integrating relationship* in helping to improve the forecast. A simple idea would be to compare our VEC model with a VAR *in difference* with the same number of lags to assess if the error correcting term really brings any additional information to the forecast. Hence, the only explanatory variable that the VEC contains and that the VAR does not contain is the error-correcting term, that is, the estimated linear long term Fed relationship. The following table, for each sample period, presents the obtained results:

Table 7: models were estimated using the whole 01/01/1980 to 24/05/2011 period

Series	Y	E/P

	VEC	VAR in difference	VEC	VAR in difference
1/01/1980 to 1/01/1985	**0.198081**	0.198084	**0.203602**	0.203689
1/01/1985 to 1/01/1990	**0.197379**	0.197382	**0.203703**	0.203744
1/01/1990 to 1/01/1995	**0.203514**	**0.203514**	0.203752	**0.203744**
1/01/1995 to 1/01/2000	**0.205790**	0.205791	0.203760	**0.203744**
1/01/2000 to 1/01/2005	0.204308	**0.204305**	**0.203672**	0.203689
1/01/2005 to 1/01/2010	0.182645	**0.182644**	**0.203720**	0.203744
1/01/2005 to 24/05/2011	0.160218	**0.160217**	**0.180216**	0.180224
1/01/1980 to 24/05/2011	**0.067971**	0.067972	**0.081274**	0.081326

If we use the RMSE criteria to assess performance, then the VEC model outperforms both the VAR in level and the VAR in difference for the whole sample period (see last column of Table 6 and 7). Nevertheless we note that on sub-samples it only outperforms the VAR in about half of the sub-samples. Moreover, the amount by which it outperforms the VAR on the whole period does not seem very large.

Overall, the Fed model performs poorly compared to similar models. The changing nature of the VEC long term relationship could explain such poor performance. If we re-estimate the models for each of the sub-sample period we find the results presented in Table 8 and 9:

Table 8: models were estimated on each sub-sample period

Series: Y	VEC model	VAR model in level	Lagged variable	VEC model	VAR model in level	Lagged variable
Indicator:	RMSE	RMSE	RMSE	Correct changes	Correct changes	Correct changes
1/01/1980 to 1/01/1985	**0.197088**	0.197144	0.197368	610	613	**786**
1/01/1985 to 1/01/1990	0.196417	**0.196394**	0.196677	598	598	**759**
1/01/1990 to 1/01/1995	**0.202589**	0.202594	0. 202636	666	656	**772**
1/01/1995 to 1/01/2000	**0.204806**	0.204822	0.204891	620	625	**768**
1/01/2000 to 1/01/2005	**0.203380**	**0.203380**	0.203437	632	628	**755**
1/01/2005 to 1/01/2010	**0.181964**	0.182077	0.182209	642	636	**786**
1/01/2005 to 24/05/2011	**0.159562**	0.159652	0.159784	802	800	**984**
1/01/1980 to 24/05/2011	0.067971	0.067979	0.068068	3910	3917	4822
Series: E/P	VEC model	VAR model in level	Lagged variable	VEC model	VAR model in level	Lagged variable
Indicator:	RMSE	RMSE	RMSE	Correct changes	Correct changes	Correct changes
1/01/1980 to 1/01/1985	**0.202566**	0.202639	0.203120	644	640	**786**
1/01/1985 to 1/01/1990	**0.202722**	0.202725	0.203124	588	594	**739**
1/01/1990 to 1/01/1995	**0.202787**	0.202794	0.202859	592	590	**748**
1/01/1995 to	0.202772	**0.202770**	0.202798	650	641	**789**

1/01/2000						
1/01/2000 to 1/01/2005	**0.202737**	0.202746	0.202808	605	604	**747**
1/01/2005 to 1/01/2010	**0.202977**	0.203157	0.203974	583	559	**735**
1/01/2005 to 24/05/2011	**0.179558**	0.179729	0.180335	746	743	**961**
1/01/1980 to 24/05/2011	**0.081274**	0.081321	0.081684	3810	3818	**4768**

Our Vector Error Correcting model, with a different, changing "long term" relationship for each subsample, now performs better than the VAR in level for most of the time. Yet, it is important to note that the lagged variable better predicts the *direction changes* of the explained variable than both models. Comparing our VEC to a VAR in first difference gives the following results:

Table 9: models were estimated on each sub-sample period

Series	Y		E/P	
	VEC	VAR in difference	VEC	VAR in difference
1/01/1980 to 1/01/1985	**0.197088**	0.197107	**0.202566**	0.202783
1/01/1985 to 1/01/1990	**0.196417**	0.196427	**0.202722**	0.202842
1/01/1990 to 1/01/1995	**0.202589**	0.202602	**0.202787**	0.202789
1/01/1995 to 1/01/2000	**0.204806**	0.204825	**0.202772**	0.202775
1/01/2000 to 1/01/2005	**0.203380**	0.203404	**0.202737**	0.202745
1/01/2005 to 1/01/2010	**0.181964**	0.181985	**0.202977**	0.203013
1/01/2005 to 24/05/2011	**0.159562**	0.159576	**0.179558**	0.179617
1/01/1980 to 24/05/2011	**0.067971**	0.067972	**0.081274**	0.081326

Nevertheless, the amount by which the VEC, and hence the Fed model, outperforms a simple VAR in difference seems small and *possibly only due to the fact that in the OLS used to estimate the VEC we have one more explanatory variable than in the corresponding VAR (the error correcting term).* One should try to factor out this purely mechanical effect of more explanatory variables before asserting the estimated long term relationship improves the forecasts.

To detangle the real effect of the error correcting term from the purely mechanical effect of improving the forecast by having one explanatory variable more in the VEC than in the VAR, it could be useful to allow one more lag to the VAR. Unfortunately this solution increases by *two* the number of explanatory variables of the VAR, and not by one. Hence the results would be now biased toward finding that the VAR gives better prediction. A way to diminish that bias would be to allow more lags for both models, which will make the one more explanatory variable of the VAR less important. Hence we allow 10 lags in the VAR and only 9 in the VEC, which implies the OLS for the VAR has one more explanatory variable than for the VEC. Table 10 below reports the results:

Table 10: comparison between VEC and VAR

Series	Y		E/P	
	VEC	VAR in difference	VEC	VAR in difference
1/01/1980 to 1/01/1985	**0.197166**	0.197184	**0.202755**	0.202957
1/01/1985 to 1/01/1990	**0.196685**	0.196692	**0.202942**	0.203016
1/01/1990 to 1/01/1995	**0.202686**	0.202699	0.202800	**0.202794**
1/01/1995 to 1/01/2000	**0.203944**	0.203963	**0.201959**	**0.201959**
1/01/2000 to 1/01/2005	**0.203228**	0.203250	0.202602	**0.202598**
1/01/2005 to 1/01/2010	**0.181109**	0.181127	0.202043	**0.201959**
1/01/2005 to 24/05/2011	**0.158848**	0.158854	0.178792	**0.178775**
1/01/1980 to 24/05/2011	0.067889	**0.067887**	**0.081196**	0.081244

For the each subsample period, the VEC model is still the best for predicting the earning to price ratio E/P, but the evidence is much more mixed than in the previous table. After all, the Fed model pretends the long-term relationship (which changes on each of the 7 sub-samples indicated here) is relevant for predicting the E/P ratio, and hence the prices of stocks, since the earnings we are using here are backward looking (trailing 12 months earning are already known to the market). What we actually observe is that the VAR in first difference forecasts as often as the VEC the E/P ratio. And the Fed model seems better to forecasts the yield Y on each sub-sample (because the long-term relationship is re-estimated on each sample period here), but not on the whole sample[19] (suggesting, again, an instable relationship over longer horizons).

We conclude that, at a daily frequency, the VAR in first difference or the VEC are about equivalent models. At the daily frequency the error correcting term does not seem to add much to the prevision of the forecast. Moreover, to predict the direction of the changes in the variables (increase or decrease), both models do worse than the lagged variable itself.

Out-of-sample evaluation of the historical predictive power of the model

Now all the previous historical assessments of the models were done *in-sample*: although no information available at time t or after is given to the model to predict values at time t, the coefficients of the co-integrating relationship and of the VEC model itself are *estimated* using the *whole* sample period (that we occasionally also changed to get the results for Table 8 and 9). Hence these are *in-sample* predictions. How does the model perform for *out-of-sample* predictions?

[19] In other words, a single estimate of the long term relationship completely deprive the Fed model from its predictive power: the relationship needs to be estimated more often, on smaller sub-sample.

To get an idea of the predictive power, out-of-sample, of a VEC model between the Y and E/P series we obtained Table E and F.

To get the first line of our tables, we first estimate the two models from the date in the first column, which is 01/01/1980, to the 1/1/1990. Then we use the model to predict out-of-sample estimates for the next date, 2/1/1990. We then re-estimate the model from 01/01/1995 to 2/1/1990 to get next forecast, for 3/1/2000. We are thus assured not to "peer" into the future, and that, at the same time, to use all the information available up to and including time t-1 to make our predictions for time t. These can thus translate into practical, concrete models for predicting future movements of the explained variables. We proceed similarly with the other dates, starting always our estimation on the first date for a 10 year period before starting to predict for the next (daily) period.

Here too, we can define the residuals as the difference between the realised value and the predicted value and compare them with the residuals obtained from a "naïve" forecast (the variable itself lagged one period), or to those of a VAR, either in level or in first difference. The results are presented in the Table 11 and 12 below:

Table 11: out-of-sample results with re-estimation of all models at each step

Series: Y	VEC model	VAR model In level	Lagged variable	VEC model	VAR model In level	Lagged variable
Indicator:	RMSE	RMSE	RMSE	Correct changes	Correct changes	Correct changes
Since 1/01/1980	0.052259	**0.052204**	0.052249	2700	2696	**3274** / 7817
Since 1/01/1990	**0.069018**	0.069090	0.069254	1454	1434	**1735** / 4164
Since 1/01/2000	**0.254042**	0.254046	0.254086	161	149	**194** / 512
Series: E/P	VEC model	VAR model In level	Lagged variable	VEC model	VAR model In level	Lagged variable
Indicator:	RMSE	RMSE	RMSE	Correct changes	Correct changes	Correct changes
Since 1/01/1980	**0.067341**	0.067411	0.067686	2555	2571	**3241**
Since 1/01/1990	**0.091924**	0.092067	0.092802	1343	1363	**1705**
Since 1/01/2000	**0.264602**	0.264565	0.264299	171	174	**223**

Table 12: out-of-sample results with re-estimation of all models at each step

Series	Y		E/P	
	VEC	VAR in difference	VEC	VAR in difference
from 1/01/1980	0.052259	**0.052258**	0.067341	**0.067247**
from 1/01/1990	**0.069018**	0.069021	**0.091924**	0.091987

from 1/01/2000	0.254042	**0.254041**	**0.264602**	0.264618

Of course, this could also be done for other sub-samples[20].

Estimating the long term relationship over the lagged 5 year period seems to give better results for predicting the earning yield at daily frequency, but does worse for the bond yield. Overall, at daily frequencies, the Fed model does not perform better than a simple VAR in first difference, in particular when the mechanical effect of a better fit due to one more explanatory variable is factored out by the addition of one more lag for the VAR than for the Fed model.

4. The time-horizon of linear Fed model: how short is the long-run?

In this section we first determine the period where a linear "Fed relationship" seems stable enough, and then try to determine the frequency at which the Fed model is relevant using daily, weekly and monthly data. The same methodology, which consists in comparing results obtained from a VEC model to a similar VAR model in first difference, is used. Because the methodology is similar to the one already described in Section 3 we do not fully describe the different statistics and steps again. The interested reader can find a full description on the author's webpage[21], Annex 6. Hence this section only recapitulates the main findings and the references to the additional documents of the detailed study.

Estimating a linear relationship between earning yields and bond yields for the US indicates that the relationship is highly unstable over the 1980 to present day period. More precisely, it seems to be highly unstable since 2001 (part A of Annex 6).

First, a log term linear relationship between earning yield and bond yield does not exist besides the 1980 to 2000 period. For example, in the most recent period (from 2001 onward) there is no evidence of co-integration between the two variables. Nevertheless, at monthly frequency the short-term interactions are significant and help improve the goodness-of-fit as evaluated by RMSE, MSAV or the number of correct changes, compared to the lagged variable (part D of Annex 6).

Restricting to the 1980 to 2000 period, the Fed model cannot be observed at daily frequency as a "long-term relationship", although short term dynamics between the earning yield and the bond yield exists (part B of Annex 6). At the weekly frequency, the Fed model is only useful for explaining the earnings ratio, and not to the bond yield (Part E of Annex 6). At the monthly frequency, and for the 1980 to 2000 period, there is indeed evidence that assuming and estimating a long-term relationship between the earning yield and the bond yield does improve the forecasts (Part C of Annex 6).

[20] See the Annex 7 of document quoted in footnote 5, page 6.

[21] See the document quoted in footnote 5, page 6.

5. Conclusion and direction for further research

In accordance with other studies, we found that the Fed model is based on rather limited empirical evidence. Concerning its **geographical scope**, it does not “travel” well. Even restricted to the US, it does not work for other stock indexes than the S&P 500 and the relevant **historical scope** is confined within the 1980 to 2000 period. In particular, although co-movements in the earning yields and the bond yield existin the 2001 to 2011 period, these movements are short-term in nature, in the sense that they are captured by immediate feedback when a VEC model is estimated, and do not contribute to the (long-term) error correcting term of the VEC in any meaningful way. In particular, building a VEC gives less precise results for forecasting than a VAR in first difference: the long-term relationship is a hindrance more than a useful tool.

Concerning its **frequencies**, we found that the Fed model, during the 1980 to 2000 period, does not have any predictive power at the daily frequency. It starts having some predictive power for the earning yield (but not for the bond yield) at the weekly frequency, and for both the earning yield and the bond yield at monthly frequency. Although we use different sub-samples as a robustness-check for checking the better goodness-of-fit of the Fed model compared to simple similar models which do not contain a long-term relationship, a word of caution is warranted by the very sample-dependence observed of the Fed model in general.

Bibliographie

Asness, C. (2003). Fight the Fed model. *Journal of Portfolio Management, 30*, 11-24.

Board of Governors of the Federal Reserve System (1997). *Monetary policy report to the congress pursuant to the full employment and balanced growth act of 1978.* Technical report, 23-24 of www.federalreserve.gov/boarddocs/hh/1997/july/fullreport.htm.

Campbell, J., & Shiller, R. (1988). Stock prices, earnings and expected dividends. *Journal of Finance 43*, 661–676.

Dickey, D., & Fuller, W. (1981). Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica 49*, 1057–1072.

Durré, A., & Giot, P. (2004). Endorse or fight the Fed model? An international analysis of earnings, stock prices and bond yields.

Engle, R., & Granger, C. (1987). Co-integration and error correction: Representation, estimation, and testing. *Econometrica 55*, 251-276.

Gordon, M. (1962). The Investment, Financing, and Valuation of the Corporation. *Irwin, Homewood, IL.*

Gwilym, O. a., Seaton, J., Suddason, K., & Thomas, S. (2001). Does the Fed Model Travel Well?

Johansen, S. (1995). Likelihood-based Inference in Cointegrated Vector Autoregressive Models. *Oxford University Press*.

Thomas, J., & Zhang, F. (2003). Don't fight the Fed Model.

Williams, J. B. (1938). Evaluation by the rule of the present worth. *The Theory of Investment Value*, 55-75.

ANNEXES

LIST 1: Stock exchange indices description

INDU Index (US):	the Dow Jones Industrial Average is a price-weighted average of 30 blue-chip stocks that are generally the leaders in their industry.	01/01/1963
SPX Index (US):	the Standard and Poor's 500 Index is a capitalization-weighted index of 500 stocks. The index is designed to measure performance of the broad domestic economy through changes in the aggregate market value of 500 stocks representing all major industries.	02/05/1993
CCMP Index (US):	The NASDAQ Composite Index is a broad-based capitalization-weighted index of stocks in all three NASDAQ tiers: Global Select, Global Market and Capital Market.	01/01/1995
SPTSX Index (Canada):	The S&P/Toronto Stock Exchange Composite Index is a capitalization-weighted index designed to measure market activity of stocks listed on the TSX.	01/06/1994
MEXBOL Index (Mexico):	The Mexican IPC index (Indice de Precios y Cotizaciones) is a capitalization weighted index of the leading stocks traded on the Mexican Stock Exchange.	01/08/2001
IBOV Index (Brazil):	The Bovespa Index is a total return index weighted by traded volume and is comprised of the most liquid stocks traded on the Sao Paulo Stock Exchange.	01/01/2007
SXXE Index (Euro Area):	The EURO STOXX (Price) Index is a capitalization-weighted index which includes countries that are participating in the EMU.	01/01/2002
SXXP Index (Euro Area):	The STOXX Europe 600 (Price) Index is a broad based capitalization-weighted index of European stocks designed to provide a broad yet liquid representation of companies in the European region.	01/01/2002
SX5E Index (Euro Area):	The EURO STOXX 50 (Price) Index is a free-float market capitalization-weighted index of 50 European blue-chip stocks from those countries participating in the EMU. Each component's weight is capped at 10% of the index's total free float market capitalization.	01/05/2001
UKX Index (UK):	The FTSE 100 Index is a capitalization-weighted index of the 100 most highly capitalized companies traded on the London Stock Exchange. The equities use an investibility weighting in the index calculation.	01/01/2002
CAC Index (France):	The CAC-40 Index is a narrow-based, modified capitalization-weighted index of the 40 highest market caps on the Paris Bourse.	01/06/2001
DAX Index (Germany):	The German Stock Index is a total return index of 30 selected German blue chip stocks traded on the Frankfurt Stock Exchange. The equities use free float shares in the index calculation.	01/04/1997
IBEX Index (Spain):	The IBEX 35 is the official index of the Spanish Continuous Market. The index is comprised of the 35 most liquid stocks traded on the Continuous market. It is calculated, supervised and published by the Sociedad de Bolsas.	01/04/1993
FTSEMIB (Italy):	The Index will consist of the 40 most liquid and capitalised stocks listed on the Borsa	01/01/2004

	Italiana. In the FTSE MIB Index foreign shares will be eligible for inclusion.	
AEX Index (Sweeden):	The AEX-Index is a free-float adjusted market capitalization weighted index of the leading Dutch stocks traded on the Amsterdam Exchange. The index was adjusted to the Dutch Guilder fixing rate.	01/01/2004
OMX (Sweeden):	The OMX Stockholm 30 Index is a capitalization-weighted index of the 30 stocks that have the largest volume of the trading on the Stockholm Stock Exchange. The equities use free float shares in the index calculation.	01/01/1994
SMI Index (Switzerland):	The Swiss Market Index is a capitalization-weighted index of the 20 largest and most liquid stocks of the SPI universe. It represents about 85% of the free- float market capitalization of the Swiss equity market.	01/01/2002
NKY Index (Japan):	The Nikkei-225 Stock Average is a price-weighted average of 225 top-rated Japanese companies listed in the First Section of the Tokyo Stock Exchange. The Nikkei Stock Average was first published on May 16, 1949, where the average price was ¥176.21 with a divisor of 225.	01/02/2000
TPX Index (Japan):	The TOPIX, also known as the Tokyo Stock Price Index, is a capitalization weighted index of all companies listed on the First Section of the Tokyo Stock Exchange. The index is supplemented by the sub-indices of the 33 industry sectors.	01/04/1993
SHASHR Index (China):	The Shanghai A-Share Stock Price Index is a capitalization-weighted index. The index tracks the daily price performance of all A-shares listed on the Shanghai Stock Exchange that are restricted to local investors and qualified institutional foreign investors.	01/06/2005
AS51 Index (Australia):	The S&P/ASX 200 measures the performance of the 200 largest index-eligible stocks listed on the ASX by float-adjusted market capitalization. Representative, liquid and tradable, it is widely considered Australia's preeminent benchmark index. The index is float-adjusted. The index was launched in April 2000.	01/01/2000

NB: The unavailability of data in the Bloomberg series often came from the earning yield. The availability of the dividend yield, or of the dividend payout ratio, is even worse.

TABLE 1: ADF and Philipps-Perron test[22]s with daily data for Fed indicators

		Y-E/P		Log(Y)-log(E/P)		Y/(E/P)	
	Sample starts in	ADF	P-P	ADF	P-P	ADF	P-P
US and SPX	01/01/1963	-2.96 0.03	-2.87 0.04	-2.44 0.13	-2.42 0.13	-2.68 0.07	-2.65 0.08
US and SPX	01/01/1993	-1.51 0.53	-1.50 0.53	-1.41 0.57	-1.43 0.56	-1.64 0.45	-1.55 0.51
US and SPX	01/01/2001	-2.08 0.25	-2.09 0.24	-1.99 0.29	-2.03 0.27	-2.27 0.18	-2.21 0.20
US and INDU	02/05/1993	-2.38 0.15	-1.93 0.31	-2.21 0.20	-1.95 0.30	-2.07 0.25	-1.92 0.32
US and CCMP	01/01/1995	-2.03 0.27	-2.04 0.26	-2.11 0.24	-3.28 0.01	-11.77 0.00	-1.87 0.00
CA and SPTSX	01/06/1994	-2.52 0.11	-2.42 0.13	-2.05 0.26	-1.97 0.29	-2.37 0.14	-2.28 0.17
MX and MEXBOL	01/08/2001	-4.35 0.00	-4.00 0.00	-4.38 0.00	-4.14 0.00	-4.44 0.00	-4.29 0.00
BR and IBOV	01/01/2007	-2.06 0.26	-1.95 0.31	-1.78 0.38	-1.74 0.40	-1.61 0.47	-1.57 0.49
EU and SXXE	01/01/2002	-2.01 0.28	-1.91 0.32	-1.98 0.29	-1.92 0.31	-2.62 0.08	-2.57 0.09
EU and SXXP	01/01/2002	-2.08 0.24	-1.98 0.29	-1.97 0.29	-1.91 0.32	-2.39 0.14	-2.28 0.17
EU and	01/05/2001	-1.79	-1.58	-1.69	-1.67	-2.96	-2.96

[22] The first line of each cell is the t-statistics (either from an ADF tests whose lag is selected by optimizing the Schwarz Information Criterion, with a maximum lag of 25, or a Phillips-Perron test, using the Newest-West Bandwith test). Data used is daily, end-of-the day data (closing prices).

SX5P		0.38	0.48	0.43	0.44	0.03	0.03
UK and UKX	01/01/2002	-2.54 0.10	-2.29 0.17	-2.31 0.16	-2.13 0.23	-2.85 0.05	-2.61 0.08
FR and CAC	01/06/2001	-2.70 0.07	-2.44 0.13	-2.44 0.12	-2.26 0.18	-2.53 0.10	-2.39 0.14
DE and DAX	01/02/1997	-2.30 0.16	-2.30 0.16	-2.22 0.19	-2.24 0.19	-3.42 0.01	-3.52 0.00
DE and DAX	01/01/2001	-2.15 0.22	-2.16 0.22	-2.14 0.22	-2.15 0.22	-3.17 0.02	-3.26 0.01
ES and IBEX	01/04/1993	-2.87 0.04	-2.95 0.03	-2.06 0.26	-1.91 0.32	-2.65 0.08	-2.38 0.14
ES and IBEX	01/01/2001	-1.98 0.29	-1.71 0.42	-2.15 0.22	-1.93 0.31	-2.59 0.09	-2.49 0.11
IT and FTSEMIB	01/01/2004	-2.91 0.043	-2.88 0.047	-3.32 0.0139	-3.28 0.0158	-4.06 0.0011	-4.05 0.0012
SE and OMX	01/01/1994	-2.95 0.039	-2.91 0.04	-1.92 0.32	-1.92 0.32	-2.18 0.21	-2.17 0.21
CH and SMI	01/01/2002	-1.76 0.39	-1.75 0.40	-2.21 0.20	-2.04 0.26	-3.22 0.018	-3.27 0.016
JP and NKY	01/02/2000	-2.51 0.11	-2.38 0.14	-2.78 0.06	-2.82 0.055	-3.41 0.0104	-3.46 0.0089
JP and TPX	01/04/1993	-1.67 0.44	-1.57 0.49	-2.92 0.04	-2.94 0.04	-7.21 0.00	-7.99 0.00
CN and SHASHR	01/06/2005	-1.2 0.67	-1.30 0.62	-1.13 0.70	-1.25 0.65	-1.13 0.70	-1.19 0.67
AU and AS51	01/01/2000	-3.26 0.016	-3.19 0.02	-3.07 0.02	-2.98 0.03	-3.27 0.016	-3.15 0.02

TABLE 2: ADF and Philipps-Perron tests[23] with monthly (end of the month) data for Fed indicators

		Y-E/P		Log(Y)-log(E/P)		Y/(E/P)	
	Sample starts in	ADF	P-P	ADF	P-P	ADF	P-P
US and SPX	01/01/1963	-2.90 0.04	-2.81 0.05	-2.51 0.11	-2.35 0.15	-2.73 0.06	-2.12 0.23
US and SPX	01/01/1993	-1.65 0.45	-1.62 0.46	-1.59 0.48	-1.55 0.50	-1.61 0.47	-1.59 0.48
US and SPX	01/01/2001	-2.21 0.20	-2.22 0.20	-2.18 0.21	-2.16 0.22	-2.22 0.19	-2.05 0.26
US and INDU	02/05/1993	-1.53 0.51	-1.85 0.35	-1.52 0.52	-1.84 0.35	-2.02 0.27	-1.79 0.38
US and INDU	01/01/2001	-1.87 0.34	-1.72 0.41	-1.73 0.41	-1.73 0.40	-2.06 0.26	-1.80 0.37
US and CCMP	01/01/1995	-2.16 0.21	-2.15 0.22	-1.40 0.57	-1.40 0.57	-0.94 0.77	-1.68 0.43
US and CCMP	01/01/2001	-1.48 0.54	-1.48 0.54	-1.44 0.55	-1.44 0.56	-1.90 0.33	-1.92 0.32
CA and SPTSX	01/06/1994	-2.83 0.05	-2.65 0.08	-1.81 0.37	-2.15 0.22	-2.18 0.21	-2.38 0.14
MX and MEXBOL	01/08/2001	-4.04 0.0017	-4.01 0.0017	-4.10 0.0014	-3.97 0.0021	-4.14 0.0012	-4.10 0.0014
BR and IBOV	01/01/2007	-1.81 0.37	-2.00 0.28	-1.62 0.46	-1.88 0.38	-1.51 0.51	-1.90 0.32
EU and SXXE	01/01/2002	-1.85 0.35	-1.82 0.36	-1.90 0.32	-1.80 0.37	-1.76 0.39	-2.32 0.16
EU and SXXP	01/01/2002	-1.92 0.31	-1.90 0.33	-1.89 0.33	-1.90 0.32	-2.67 0.08	-2.19 0.20
EU and SX5P	01/05/2001	-1.08 0.72	-1.39 0.58	-1.13 0.70	-1.56 0.49	-1.88 0.33	-2.84 0.054
UK and UKX	01/01/2002	-2.15 0.22	-2.02 0.27	-1.94 0.30	-2.30 0.17	-3.59 0.0069	-2.97 0.03
FR and CAC	01/06/2001	-2.23 0.19	-2.37 0.15	-2.08 0.25	-2.24 0.19	-2.21 0.20	-2.30 0.17

[23] The first line of each cell is the t-statistics (either from an ADF tests whose lag is selected by optimizing the Schwarz Information Criterion, with a maximum lag of 25, or a Phillips-Perron test, using the Newest-West Bandwith test). Data used is daily, end-of-the day data (closing prices).

DE and DAX	01/02/1997	-2.03 0.27	-2.28 0.17	-2.14 0.22	-2.48 0.12	-3.27 0.0178	-3.64 0.0059
DE and DAX	01/01/2001	-1.87 0.34	-2.03 0.27	-2.05 0.26	-2.05 0.26	-3.04 0.0337	-3.33 0.015
ES and IBEX	01/04/1993	-1.53 0.51	-1.62 0.46	-1.72 0.41	-1.71 0.42	-1.97 0.29	-1.86 0.34
ES and IBEX	01/01/2001	-1.61 0.47	-1.68 0.43	-1.93 0.31	-1.89 0.33	-2.04 0.26	-2.11 0.23
IT and FTSEMIB	01/01/2004	-2.56 0.104	-2.65 0.08	-2.65 0.08	-2.77 0.65	-2.85 0.054	-2.92 0.046
SE and OMX	01/01/1994	-2.31 0.16	-3.00 0.03	-1.80 0.37	-1.97 0.29	-2.11 0.23	-2.27 0.18
CH and SMI	01/01/2002	-1.55 0.50	-1.84 0.35	-1.85 0.35	-1.86 0.34	-1.84 0.35	-2.47 0.12
JP and NKY	01/02/2000	-2.34 0.15	-2.41 0.14	-2.98 0.039	-3.05 0.03	-5.29 0.00	-3.58 0.007
JP and TPX	01/04/1993	-1.50 0.52	-1.53 0.51	-1.85 0.35	-2.48 0.12	-6.10 0.00	-7.92 0.00
CN and SHASHR	01/06/2005	-1.91 0.32	-1.79 0.37	-1.44 0.55	-1.87 0.34	-1.43 0.55	-1.84 0.35
AU and AS51	01/01/2000	-3.22 0.020	-3.29 0.017	-3.51 0.009	-3.35 0.014	-4.18 0.001	-3.57 0.007

TABLE 3: ADF and Philipps-Perron tests[24] with monthly (average of the month) data for Fed indicators

		Y-E/P		Log(Y)-log(E/P)		Y/(E/P)	
	Sample starts in	ADF	P-P	ADF	P-P	ADF	P-P
US and SPX	01/01/1963	-2.90 0.04	-2.81 0.05	-2.51 0.11	-2.35 0.15	-2.73 0.06	-2.12 0.23
US and SPX	01/01/1993	-1.58 0.48	-1.60 0.47	-1.55 0.50	-1.54 0.50	-1.52 0.52	-1.56 0.50
US and SPX	01/01/2001	-2.26 0.18	-2.23 0.19	-2.23 0.19	-2.17 0.21	-2.29 0.17	-2.04 0.26
US and INDU	02/05/1993	-1.53 0.51	-1.85 0.35	-1.52 0.52	-1.84 0.35	-2.02 0.27	-1.79 0.38
US and INDU	01/01/2001	-1.84 0.35	-1.79 0.38	-1.70 0.42	-1.70 0.42	-1.67 0.44	-1.76 0.39
US and CCMP	01/01/1995	-2.18 0.21	-1.94 0.30	-1.24 0.65	-1.29 0.62	-1.22 0.66	-1.87 0.34
US and CCMP	01/01/2001	-1.34 0.60	-1.43 0.56	-1.40 0.57	-1.40 0.57	-1.81 0.37	-1.80 0.37
CA and SPTSX	01/06/1994	-2.89 0.04	-2.57 0.10	-2.43 0.13	-2.12 0.23	-2.46 0.12	-2.00 0.28
MX and MEXBOL	01/08/2001	-3.42 0.0118	-3.42 0.0118	-3.50 0.0094	-3.56 0.0080	-3.61 0.0068	-3.67 0.0057
BR and IBOV	01/01/2007	-2.16 0.22	-1.97 0.29	-2.03 0.27	-1.85 0.35	-2.73 0.07	-1.85 0.34
EU and SXXE	01/01/2002	-1.77 0.39	-1.84 0.35	-1.77 0.39	-1.77 0.39	-1.76 0.39	-2.12 0.23
EU and SXXP	01/01/2002	-1.85 0.35	-1.92 0.32	-1.79 0.38	-1.88 0.33	-2.10 0.24	-2.08 0.25
EU and SX5P	01/05/2001	-1.31 0.62	-1.41 0.57	-1.28 0.63	-1.47 0.54	-2.20 0.20	-2.66 0.08
UK and UKX	01/01/2002	-2.12 0.23	-2.18 0.21	-2.89 0.04	-2.36 0.15	-3.95 0.0020	-2.87 0.05
FR and CAC	01/06/2001	-2.14 0.23	-2.37 0.15	-1.99 0.28	-2.22 0.19	-2.45 0.12	-2.29 0.17
DE and DAX	01/02/1997	-1.55 0.50	-1.73 0.41	-1.78 0.38	-2.06 0.26	-2.65 0.08	-3.14 0.025
DE and DAX	01/01/2001	-1.55 0.50	-1.73 0.40	-1.78 0.38	-2.06 0.26	-2.65 0.08	-3.14 0.02
ES and IBEX	01/04/1993	-2.40 0.14	-2.58 0.09	-1.72 0.41	-1.86 0.34	-2.11 0.23	-2.17 0.21
ES and IBEX	01/01/2001	-1.53 0.51	-1.62 0.46	-1.72 0.41	-1.71 0.42	-1.97 0.29	-1.86 0.34

[24] The first line of each cell is the t-statistics (either from an ADF tests whose lag is selected by optimizing the Schwarz Information Criterion, with a maximum lag of 25, or a Phillips-Perron test, using the Newest-West Bandwith test). Data used is daily, end-of-the day data (closing prices).

IT and FTSEMIB	01/01/2004	-2.56 0.10	-2.65 0.08	-2.65 0.08	-2.77 0.065	-2.85 0.054	-2.92 0.046
SE and OMX	01/01/1994	-1.9 0.33	-2.61 0.09	-2.25 0.18	-1.92 0.31	-2.46 0.12	-2.18 0.21
CH and SMI	01/01/2002	-1.48 0.53	-1.78 0.38	-1.74 0.40	-1.77 0.39	-2.55 0.105	-2.51 0.11
JP and NKY	01/02/2000	-2.17 0.21	-2.30 0.17	-2.74 0.06	-2.91 0.04	-5.08 0.00	-3.70 0.005
JP and TPX	01/04/1993	-1.45 0.55	-1.50 0.52	-1.83 0.36	-2.24 0.19	-10.74 0.00	-10.93 0.00
CN and SHASHR	01/06/2005	-2.31 0.16	-1.85 0.34	-2.21 0.20	-1.84 0.35	-2.18 0.21	-1.81 0.36
AU and AS51	01/01/2000	-2.98 0.039	-3.14 0.02	-3.37 0.013	-3.19 0.02	-3.92 0.00	-3.39 0.013

Article 3: Optimal CSD reshaping towards T2S[1]

Non technical summary

TARGET2-Securities (T2S) is a project for creating a single European IT-platform for settling securities transactions. It will offer the various national Central Securities Depositories (CSDs) harmonised securities settlement services in central bank money for a single European market. CSDs deciding to join T2S have to adapt their own IT infrastructure to T2S, that is, to "reshape" towards T2S. This adaptation towards T2S requires some investments by CSDs. In practice anything between two extreme cases is possible: in the first extreme scenario, a CSD could apply a "T2S-fees-on-top approach" in which the CSD simply keeps its existing IT infrastructure and puts T2S on top of it. In this case, the CSD will have to replicate all the data from T2S to its own legacy system before it can further operate and provide its services. Such a strategy probably implies no cost savings compared to the status quo and puts the operational costs of T2S on top. This strategy may, however, limit the immediate CSD adaptation costs to T2S. In the other extreme scenario, a CSD could apply a "Greenfield approach" in which it sets up a completely new infrastructure that is entirely shaped to incorporate T2S in the most cost-efficient way. In this case, this CSD could minimise the part of its own costs that come on top of the fees to be paid for the use of T2S, but at the same time it would probably face significant upfront investment costs. Hence, each CSD is faced with an individual strategic decision, and the overall gain in efficiency brought by T2S to the settlement industry depends on each CSD's individual decision as to which degree it will reshape.

Since reshaping involves a significant lump sum investment, CSDs could be tempted to delay the moment when they reshape. Because higher operating costs in the industry justify higher prices and possibly bigger CSD profit margins, there could be an incentive for CSDs to tacitly collude to avoid reshaping. Hence, besides the problem of knowing what is a CSD's individual optimal degree of reshaping is the problem of studying the possibility (and plausibility) of tacit collusion not to reshape.

This study makes use of a simple analytical model, cast into a dynamic game theoretic setting, to answer both issues. A closed-form expression for the optimal degree of reshaping is derived from the model, and simulations are provided as an illustration. These simulations involve ranges of plausible parameters derived from public data: For example, the cost of settling in T2S is taken from the T2S price list published by the ECB and the adaptation costs for a large CSD are assumed within the very broad range of EUR 5 million to EUR 50 million. Clearstream and Euroclear, the two largest European CSD groups, have announced that they intend to charge their users a maximum of EUR 30 million and EUR 25 million, respectively, as adaptation costs to T2S. In particular, these simulations allow to experimentally visualize the sense of variation of the optimal degree of reshaping as a function of the various parameters of the model, such as the market size, the costs per transaction of a given CSD, the price elasticities of the demand of settlement services, etc. A simple condition under which tacit collusion not to reshape is an equilibrium is provided. Interestingly, when this condition fails, and when price competition is assumed in each price-setting stage of the model, there can be no such tacit collusion. The stability of the various equilibria is then discussed. Finally, it is shown analytically that two natural generalisations of the model, i.e. allowing for an arbitrary number of CSDs and introducing a delay in the observability of the reshaping decision of other CSDs, do not affect the results of this paper.

The focus of the paper is on the optimal decision of CSDs for their reshaping towards T2S. The results provide interesting insights about an optimal, non-reversible investment decision that can be applied more broadly in the Industrial Organisation literature.

[1] This paper has been co-written with Stephan Sauer (stephan.sauer@ecb.europa.eu), and published in the ECB Working Paper Series no 1549, May 2013, and in the Journal of Financial Market Infrastructures, vol. 2, no. 2, pages 3-51, 2014.

1 Introduction

1.1 What is T2S?

TARGET2-Securities (T2S) is a project for creating a single European IT-platform for settling securities transactions. Developed and – as of 2015 – operated by the central banks of the euro area (the *Eurosystem*), T2S will offer the various national Central Securities Depositories (CSDs) harmonised securities settlement services in central bank money for a single European market.

The introduction of T2S is a significant change to the European post-trading industry. Today, around 40 CSDs in Europe operate in a domestically-oriented, fragmented and thus largely monopolistic market structure. An analogy often used with T2S is a railway system in which national monopolists own and operate the tracks and the trains. In this analogy, both tracks and trains have a number of national specificities, which makes travelling across borders a costly, lengthy and sometimes risky process because of the frictions associated in particular with non-harmonised tracks.

Today's European settlement industry is in many ways similar to such a fragmented railway system. Already in 2001, a group chaired by Alberto Giovannini published a report that asserted that "inefficiencies in clearing and settlement represent the most primitive and thus the most important barrier to integrated financial markets in Europe"(see Giovannini [13]). Differences for example in technical interfaces, message formats, intraday settlement finality rules, opening days and daily timetables became known as the 15 Giovannini barriers and they have largely remained in place since 2001.

T2S aims at overcoming these fragmentations in European securities settlement. In the image used above, it will provide a single set of tracks on which the railway companies may operate their trains. T2S will thus allow economies of scale, more cost-efficient processes and new business opportunities[2]. The Eurosystem has invited all CSDs in Europe to outsource their settlement services to T2S. By 2012, 22 CSDs have signed a legal agreement ("Framework Agreement") with the Eurosystem, including almost all CSDs of the euro area. While CSDs will keep the legal relationship with their customers and can provide various services to them, the settlement of securities will be technically processed in T2S.

1.2 What is CSD "reshaping" towards T2S?

CSDs need to "reshape" their existing infrastructures to access T2S. This adaptation towards T2S requires some investments by CSDs. In practice anything between two extreme cases is possible: in the first extreme scenario, a CSD could apply a "T2S-fees-on-top approach" in which the CSD simply keeps its existing IT infrastructure and puts T2S on top of it. In this case, the CSD will have to replicate all the data from T2S to its own legacy system before it can further operate and provide its services. Such a strategy probably implies no cost savings compared to the *status quo* and puts the operational costs of T2S on top. This strategy may, however, limit the immediate CSD adaptation costs to T2S.

In the other extreme scenario, a CSD could apply a "Greenfield approach" in which it sets up a completely new infrastructure that is entirely shaped to incorporate T2S in the most cost-efficient way. In this case, this CSD could minimise the part of its own costs that come on top of the fees to be paid for the use of T2S, but at the same time it would probably face significant upfront investment costs.

Hence, European CSDs face crucial strategic decisions: as the participation in T2S is voluntary, should CSDs jump with their trains on the new tracks, i.e. should they join T2S? And, for those who decide to join, to which degree should they adapt – or "reshape" – their current trains, i.e. their own IT platform, their human resources and even their business model to T2S? In particular, CSDs are faced with the problem of deciding what is the optimal pricing scheme in the new post-trade environment. For example, CSDs might be tempted to increase their fees in order to recover their immediate investment costs from adapting to T2S, thus passing these to their customers.

[2] The European Central Bank (ECB) provides information about T2S via the website www.t2s.eu. Non-technical general information on T2S is available in particular in the T2S brochures (`http://www.ecb.europa.eu/paym/t2s/about/brochures/html/index.en.html`) and in some videos in the multimedia room (`http://www.ecb.europa.eu/paym/t2s/about/multimedia/html/index.en.html`).

1.3 Outline of the model and main results

In this article we address these questions using a finite and then infinite dynamic game, where the decision to reshape incurs immediate fixed costs but future potential benefits, stemming both from cost reductions and from increased market shares. The focus of the paper is on the optimal decision of CSD for their reshaping towards T2S. The results provide interesting insights about an optimal, non-reversible investment decision than can be applied more broadly in the Industrial Organisation literature.

In our first (finite-time) model of Section 2, we give all CSDs the choice to reshape only at the beginning of the game, and then determine the unique subgame perfect Nash equilibrium of this game. In particular, we prove that in the absence of capital constraints, CSDs should not increase the price of their transaction services after reshaping (Proposition 1), and should even decrease them correspondingly to the costs-reduction obtained from their reshaping. We derive closed-form formulas giving the optimal, profit-maximizing pricing at each period of the model as well as the degree of optimal reshaping as a function of the main parameters of the model (Theorem 1), and in particular as a function of the costs per transactions of each CSD as well as of the different substitution effect among CSD settlement services.

We carry out simulations with different parameters derived from publicly available data. For example, the cost of settling in T2S is taken from the T2S price list published by the ECB. The costs for fully reshaping a large CSD towards T2S are assumed within the very broad range of EUR 5 million to EUR 50 million. At the end of 2012, Clearstream and Euroclear announced that they intend to charge their users a maximum of EUR 30 million and EUR 25 million, respectively, as adaptation costs to T2S (see Clearstream [4] and Euroclear [10]). Since greater adaptation costs reduce the benefits of reshaping, our parameters are chosen in a conservative way which is detrimental to reshaping. Also the other parameters are based on rough approximations from publicly available data, primarily from the two largest European CSD groups, Clearstream and Euroclear.

In the second model, presented in Section 3.1.2, we modify the rules of the game to allow participating CSDs to reshape at any time. This induces more strategic behaviour on the part of CSDs, because it allows them to design reshaping strategies that are dependent on others' reshaping timing. An important result is an explicit sufficient condition on costs per transactions and price-elasticities under which tacit collusion to perpetually delay the reshaping can occur (Theorem 2). That is, we provide a condition under which there also exists a particular subgame perfect Nash equilibrium in which no existing CSD reshapes towards T2S and all CSDs just puts T2S on top of its current infrastructure, because each CSD expects all other CSDs to behave like this. However, this equilibrium seems rather fragile as it likely to break down as soon as one of the CSDs deviates from it, or as soon as only marginal changes to the basic model are introduced: possibility of new entrants in the settlement industry, observability of reshaping behaviour, or enough homogeneity in cost-efficiency and price-elasticities of the main CSDs. Since all of these effects are likely to be fulfilled in the post-trading market with T2S, it seems unlikely that it will be the optimal strategy for a CSD not to reshape its current infrastructure to T2S. Moreover, this condition we prove to be also necessary (Theorem 3). Hence, we also get a condition under which there can be no collusion to delay the reshaping. The main results are further discussed in Section 3.2.3. They are robust with respect to a delay in the observability of the reshaping of other CSDs as discussed in Section 4.

We also briefly discuss – since it is not the main focus of the article – how to apply this model and hence derive the same type of results with respect to the combined decision to join and reshape. In particular, while we note that tacit collusion for not joining T2S could be one of several possible equilibria, the publicly stated intentions of some CSDs to join T2S imply that this equilibrium will not apply in reality. In particular, Theorem 7 gives a sufficient condition for a CSD to join and to completely reshape no matter what the other CSDs decide.

1.4 Literature review

There has been, until fairly recently, a relative scarcity of research articles on the post-trading industry, at least when compared to the trading industry or to other well-studied network industries such as telecommunications. For example, the 2007 survey by post-trading by Milne [21] includes only papers published after 2001. Milne's conclusion that more research in this field is needed is as valid as ever, not least because of a significant amount of regulatory (such as the Dodd-Frank act in the United States or the European Market Infrastructure Regulation (EMIR) in the European Union) and technical initiatives (such as T2S).

On the empirical side, the papers by Cayseele and Wuyts [2], and by Schmiedel, Malkamaki and Tarkka [24] have received most attention in the literature. They show that the post-trading industry exhibits significant economies of scale and scope.

On the theoretical side, the relevant literature combines elements of the classic theory of industrial organization, as described by, for example, Tirole [27] with more distinctive features of the post-trading industry. It focuses on network effects, two-sided markets and vertical as well as horizontal integration as key features of post-trading infrastructures.

A description of general network effects in an industrial organization setting is given for example by Economides [8]. The participation of additional market participants in a settlement system increases the benefits of participation for other market participants. For example, if market participants can do business with more counterparties, liquidity increases and as a consequence, capital costs tend to decrease (see for example [7], chapter 5). An introductory general discussion of these network effects in the setting of the post-trade industry, with some comparisons with other network industries, can be found in Knieps [16] and an application of the network effect in settlement in Holthausen and Tapking [14].

Post-trading market infrastructures and CSDs in particular are necessary to bring together issuers and investors in the primary market and to allow the exchange of securities between investors in the secondary market. The cost of capital for issuers and the returns of issuers after transaction costs can create a virtuous circle with the liquidity in the market. Hence, post-trading market infrastructures exhibit features of two-sided markets where for example pricing decisions should take into account both sides of the market. For example Kauko [15] is an attempt to capture this feature by designing ad-hoc models to illustrate this dual aspect specific to the post-trading industry.

The majority of the theoretical literature in this field looks at vertical and horizontal integration in trading and post-trading. In some national markets such as Germany or Italy, a single "vertical silo" operates the trading, clearing and settlement infrastructure. Other markets such as France use separate firms for the different elements of the value chain. Some of the firms active in these markets have horizontally integrated their activities across different national markets. A prominent example is Euroclear's ESES-platform (Euroclear Settlement of Euronext-zone Securities). A key difference to the standard industrial organisation literature as in [29], for example, is that market infrastructures with economies of scale and scope and network effects operate at each layer of the value chain (trading, clearing, settlement). Many financial market participants use each infrastructure directly so that the standard perspective of upstream and downstream firms producing goods or services for the final user is not necessarily directly applicable.

Tapking and Yang [26] use a theoretical model for cross-border trading to systematically compare the welfare of different industry structures for securities trading and settlement. In their model, which captures competition and transaction cost effects, welfare is the highest for horizontal integration, while it is higher for vertical integration than for complete separation. Pirrong [22] focuses on the industry structure in a single market and emphasises transaction cost reductions from vertical integration. He argues that the transaction cost benefits can be (partly) offset by increasing differences between the scope economies at the different layers of the value chain. Cherbonnier and Rochet [3] look at a different possibility of vertical integration, namely between a single CSD as provider of settlement and custody and banking services[3] and one of two competing custodian banks. They build on the model developed in Holthausen and Tapking[14] to analyse the welfare effects and optimal regulation of such vertical integration. Cherbonnier and Rochet conclude that the possibility of vertical integration would require the regulation of access pricing which would introduce other inefficiencies. They suggest that competition between several CSDs, which is one of the visions of T2S promoted by the Eurosystem, could limit the need for regulation, but this idea is not included in their formal model.[4]

By looking at the optimal reshaping decision of CSDs, we add another aspect to the literature on post-

[3] The scope of services offered by CSDs differs and can include custody and banking services as some CSDs have a banking licence.

[4] Koeppl and Monnet [17] and Rochet [23] analyse further interesting theoretical models about vertical integration in the post-trading industry.

trading that is closely linked to the industrial organisation literature. In fact, the basic mechanisms of the model remain fairly general and thus contribute to literature on investments, which has also examined the trade-off between paying a lump sum today to reduce costs or improve quality tomorrow. There has also been considerable research in a general setting about investments in productive capacity, as well as in modelling competing firms' strategies (with or without investment). For example, Fine and Freund [11] examine the trade-off faced by firms in investing into a product-specific or a more expensive flexible production capacity. Nevertheless investment in our setting would be dedicated to reducing the average cost-per-transaction, not to aim at a more flexible business. Eliashberg and Steinberg [9] examine the possible strategies of two competing firms with different cost structures. Most of these models were designed for the manufacturing industry. While capacity constraints and thus optimal capacity planning is also relevant for the processing of securities settlement transactions, a key element in the previous literature is the management of the inventory. This does not apply to the production of settlement services. The reader interested in the specific problems of investment in capacity, in particular in high IT-industry, can refer to Wu's survey [31].

One of the closest article to ours in the stream of the literature dealing with investment seems to be Spence [25]. It tackles the problem of the trade-off of paying a higher lump-sum of investment in order to reduce costs per unit of production in the subsequent periods of the game. Although Spence's article allows for very general forms of profit and demand functions and could in theory be applied to many investment decision problems, the investment decision in the article is presented as an investment in research and development, more than in infrastructure. Hence, Spence's main concerns are tied to the potential spillover effects of research on the optimal degree of investment by firms as well as the optimal level of subsidies from the authorities required to maximize aggregate welfare. Another article close to ours is Valletti and Cambini [28] where investment is designed to improve the quality of the services provided by a network, rather than to reduce costs as in our article. This investment can also a positive external effect on the quality of the service offered by other networks. The paper is similar to our article in two different aspects: first, the higher quality attracts more demand, as in our model lower costs allows lower equilibrium prices and ultimately attract more demand. Second, the authors demonstrate that high termination charges, as well as an increased interplay between the quality level of both networks, can lead to collusion to under-invest. A difference with our collusion theorem is that collusion here stems from the direct interplay between the two competitors, either through the termination fees they have to pay to each other's or through the quality spillover of their investment decision, while our tacit collusion theorem comes from the links of competitors as providing partially substitutable services of settlement. For a survey indicating how access prices can be used as an instrument of tacit collusion in a communication network setting, please refer to Vogelsang [30]. For an review of the price-competition in a dynamic setting, please refer to Maskin and Tirole [19].

Our paper obviously also contributes to the literature on T2S. The Eurosystem published an economic feasibility study about T2S in 2007 [5], followed by an economic impact assessment in 2008 [6]. These studies describe and quantify the cost savings and economic benefits associated with T2S for CSDs, their users and final investors and issuers. In particular the later study ([6]) was developed in close cooperation with market participants and the quantitative results are based on data provided by market participants.

There is, however, a significant lack of studies that use analytical models to investigate the competition, network and welfare effects of T2S more specifically. Cales et al [1] is the only paper to our knowledge that is concerned explicitly with T2S. It is based on Matutes and Padilla [20], who use the Salop model of space competition, which is itself a generalization of the Hoteling model used in many articles of the post-trading literature. An advantage of using the Salop model instead of the more classical Hoteling model is to allow for competition between three CSDs providing depository and settlement services to banks. An important assumption is that securities must be deposited in their domestic CSD, whereas banks (the users of CSDs) can choose the CSD for settlement if the CSD is in T2S. Hence, three alternative situations for competition in settlement services can be modelled: the so-called "compatibility case", where all the three CSDs join T2S for settlement, the "incompatibility" case, which corresponds to the current pre-T2S situation and no CSD joining T2S, and the "partial compatibility" case, where precisely two CSDs join T2S while the other stays out. The authors capture the benefits of T2S for cross-border / cross-CSD settlement transactions by a network benefit for banks that increases in the number of CSDs joining T2S. This is an important difference to our model in which we focus on cost-reductions for CSDs that depend on the reshaping of CSDs. They arise from the availability of a single settlement platform allowing for economies of scale.

Cales et al [1] conclude that competition will lead to price decreases if all CSDs join T2S. If only a subset of CSDs joins T2S, the effect on CSD prices depends on the relative importance of the competition and network effects as CSDs in T2S can gain part of the network benefit from banks.[5] Importantly, banks' welfare increases also in this case, as for them the network benefit always dominates the potential price increase of CSDs. Similar to our paper, Cales et al [1] mention a possible coordination problem for CSDs joining T2S and the possibility of tacit collusion among CSDs. However, we look at the possibility and sustainability of tacit collusion in much more detail.

2 A dynamic N-period model for the degree of optimal reshaping

2.1 Motivation

We consider the subset of CSDs having joined T2S and ask the question of the optimal degree of reshaping that they should target. Indeed, reshaping incurs immediate costs, but long term potential benefits, hence a balance should be found between a complete reshaping and a pure juxtaposition of the T2S IT-infrastructure and business model on top of the CSD infrastructure and its own business model, with no costs reduction, neither technological or in the servicing staff, from the participating CSD. Also, the settlement industry is a somewhat competitive environment - and even if it were not, it is expected to become much more competitive with the introduction of T2S. Hence the optimal pricing and reshaping decisions do not only depend on a given CSD's own characteristics but on the similar decisions from others.

2.2 The dynamic game

To try to capture how the degree of optimal reshaping - interpreted in a rational setting as the degree of reshaping in a Nash equilibrium of the game - depends on these various parameters and participants' action, we define the following dynamic game:

1) At the beginning of the game all CSDs having joined T2S simultaneously decide to which extent they will reshape[6]. This is modelled by the choice, for each CSD i, of a pair (a_i, b_i) of real numbers between 0 and 1, each pair being associated with a given adaptation cost $C_{i,adapt}(a_i, b_i)$. The choice of a greater reshaping to reduce fixed costs is modelled by a higher a_i, while the choice of a greater reshaping to reduce transaction costs is modelled by a greater b_i. The adaptation costs function $C_{i,adapt}$ is thus logically increasing in both its parameters, and the maximum possible reshaping corresponds to $(a_i, b_i) = (1, 1)$. Such a reshaping is not necessarily optimal, since although it would produce the

[5] The reason is that in Cales et al [1] demand is perfectly inelastic, the greater utility of being a client of compatible CSDs allows them, in some case, to price their services higher than in the incompatibility case. This result should be taken with care: first, it relies on perfect inelasticity of demand: the banks are not given the opportunity to trade and settle less often in front of higher prices in the model: they have to be some CSD clients no matter the overall prices. Second, Cales et al distinguish in total four ranges of parameters, each with its own implications in terms of profitability. In particular, when the added utility of the network effect (and which relies exclusively on the inter-substitutability of CSDs having joined T2S in terms of their settlement services) is high enough compared to transportation costs (which represents product differentiation and banks 'preferences for their closest CSD), Cales et al find that the CSDs which had opted out would be better off inside T2S. This suggests CSDs which settle products that have very close substitutes in the other CSDs' market would make a very poor bet by staying out of T2S to preserve their natural monopoly on the settlement services (service "bundled" with depository services), in particular in a context of globalization and harmonisation in Europe where many issuers could easily issue in other European countries and thus completely avoid a CSD outside of T2S.

[6] The legal contract governing the relationship between CSDs and the Eurosystem as the provider of T2S services, the Framework Agreement, requires that "The Contracting CSD shall use reasonable efforts to adapt its operational, internal guidelines as well as its processes and related technical systems in order to foster the development of the European post-trading infrastructure, make efficient use of the T2S Services and maintain the Multilateral Character of T2S" (Article 4, Chapter 2). This formulation leaves a very large degree of freedom to CSDs concerning their reshaping.

maximum cost-reduction in the future, it entails a the highest cost at the beginning. We assume CSDs can finance their adaptation costs at the cost of capital r.

2) Each CSD observes the reshaping decision of others, then play repeatedly, for a certain number $N \geq 2$ of times, the price-setting game $G(\widetilde{C}_{i,fixed}, \widetilde{c_i})$ described in the section below, which consists in a simultaneous choice by each CSD i of the price p_i it sets for its settlement service, given the associated costs $(\widetilde{C}_{i,fixed}, \widetilde{c_i})$ assumed in the following table:

period	fixed costs $\widetilde{C}_{i,fixed}$ of CSD i	cost per transaction $\widetilde{c_i}$ of CSD i
1	$(1-a_i)C_{i,fixed} + C_{i,adapt}(a_i, b_i)$	$(1-b_i)c_i + c_{T2S}$
2	$(1-a_i)C_{i,fixed}$	$(1-b_i)c_i + c_{T2S}$
...	...	...
N	$(1-a_i)C_{i,fixed}$	$(1-b_i)c_i + c_{T2S}$

Hence these costs depend on each CSD reshaping decision at the first stage of the game: if CSD i has chosen a (a_i, b_i)-reshaping at the beginning of the game, then it pays immediately $C_{adapt}(a_i, b_i)$, the corresponding adaptation costs, which is assumed increasing in both its variable a_i and its variable b_i. The cost per transaction $\widetilde{c_i}$ consists of the individual CSD's cost c_i after reshaping and a constant fee per transaction c_{T2S} for the use of T2S.

2.3 The price-setting game $G(\widetilde{c_i})$

The model chosen for capturing inter-CSD competition and substitution effect among the settlement services provided by CSDs is the following: at each period of the game the demand for settlement transactions q_i in CSD i is given by

$$q_i = \alpha_i - \gamma_{ii} p_i + \gamma_{ij} p_j$$

where p_i is the average price charged by CSD i for its settlement services, and γ_{ii} and γ_{ij} are nonnegative constants. Note that γ_{ii} is the elasticity of the volume of transaction q_i to the price per settlement transaction p_i, and γ_{ij} is its elasticity to the price per settlement transaction set by CSD j; hence greater competition and substitution effect translates into a greater γ_{ij}.

Remark 1: The demand equation could be normalised in p_i, but keeping the parameter γ_{ii} makes interpretations of the simulation part easier to understand.

Remark 2: We consider in this Remark the many possibilities presented by this very simple stage game model. First, for two CSDs i and j the model consists in assuming:

$$\begin{cases} q_i = \alpha_i - \gamma_{ii} p_i + \gamma_{ij} p_j \\ q_j = \alpha_j - \gamma_{jj} p_j + \gamma_{ji} p_i \end{cases}$$

Hence total demand for settlement services is

$$q = q_i + q_j = \alpha_i + \alpha_j + (\gamma_{ji} - \gamma_{ii}) p_i + (\gamma_{ij} - \gamma_{jj}) p_j$$

The assumption $\gamma_{ji} \leq \gamma_{ii}$ as well as $\gamma_{ij} \leq \gamma_{jj}$ ensures that there is no possible way to increase the price level (p_i, p_j) that would result in a higher *aggregate* demand (because we assume γ_{ii} and γ_{jj} nonnegative the local demand will necessarily be lower). We will thus work under this assumption. Nevertheless, our formal derivations only require the weaker assumptions $\gamma_{ji} \leq \gamma_{ii}\sqrt{2}$ and $\gamma_{ij} \leq \gamma_{jj}\sqrt{2}$.

In terms of competition through price-setting, the model can be interpreted as follows. Consider, for example, a price increase of one unit for the settlement services provided by CSD i, while the price set by the other CSD stays constant. It translates into a loss of demand of γ_{ii} for settlement services offered by CSD i, of which γ_{ji} is re-captured by CSD j: market participants which could do so have reacted by switching to the settlement services provided by j, still the higher prices result in a decrease of $\gamma_{ji} - \gamma_{ii}$ of aggregate settlement demand.

Demand is indeed responsive to price; a lower local price translate in higher settlement demand since transactions that were not profitable before now become profitable; still, we can ensure (if we ever want to) an inelastic *aggregate* demand by assuming the extreme case where $\gamma_{ji} = \gamma_{ii}$ and $\gamma_{ij} = \gamma_{jj}$. Then $q = \alpha_i + \alpha_j$ is constant, whatever the price level, while demand for settlement in a given CSD still responds negatively to prices set by this CSD. The case where settlement services are completely non-substitutable – and where, as a consequence, there can be no competition between CSD i and CSD j – is obtained by setting $\gamma_{ij} = \gamma_{ji} = 0$.

In order not to mix the incentives for the investment / reshaping decision given price competition with the effects of increases in price competition, the main part of the paper does not consider changes in price competition over time. However, Annex 7.8 shows that the set of parameters for which tacit collusion is sustainable in a high competition / substitution environment is larger than in a low competition / substitution environment.

Interpreting CSD j as consisting of all other agents providing settlement services except CSD i, instead of a particular given CSD, is a convenient simplification to generalizing the model to n CSDs and multiplying the number of parameters given the complexity of analytical solutions. The focus is then set on CSD i, while the parameters defining CSD j reflect an average of the parameters for the rest of the settlement industry.[7] In particular, new entrants in the industry will modify the value of these parameters, for example entrance of new low-cost CSDs will certainly affect c_j negatively, since they decrease the aggregate costs in the settlement industry. The price p_j would, following the same idea, be the average price for settlement services in the market, excluding the price set by CSD i.

2.4 Analytical Resolution

For dynamic games, the concept of subgame-perfect Nash equilibrium is particularly relevant, because contrary to the simple Nash Equilibrium (Nash equilibrium) concept, it does not include strategies involving threats that are not credible. This is the reason why it is widely used in economics today and will be used extensively in this article. Note that although it is common to distinguish between strict and weak Nash equilibria, we will often not do so as the limiting conditions that translate into a weak Nash equilibrium are most of the time of very little interest. In fact, we will always talk of Nash equilibrium (resp. subgame perfect Nash equilibrium) when we actually mean strict Nash equilibrium (resp. subgame perfect Nash equilibrium). Solving the game by backward induction as we proceed to do allows to find *all* subgame perfect Nash equilibria. We will see that, because each stage game has only one Nash equilibrium, there exists in fact only one subgame perfect Nash equilibrium: the one where the Nash equilibrium of each subgame is played repeatedly. Only the main results are reported here: the detailed computations are provided in Annex 7.1.

2.4.1 A Cost Structure Assumption (CSA)

We will assume for all our derivations that the costs of settlements are always smaller than the price set by the CSD for the settlement services. This allows to get rid of unwanted corner solutions, where a given CSD would in practice stop all settlement activities (which amount to withdrawing from the market), because its optimal price, that is, the price at which maximizes its profits, still does not yield positive profits. Formally, the CSA states that for any costs $(\widetilde{c}_i, \widetilde{c}_j)$ involved in the model, we have $p_i^* - \widetilde{c}_i \geq 0$ and $p_i^* - \widetilde{c}_i \geq 0$, so that no CSD makes a loss by settling an additional transaction. As we will see, at equilibrium of the price-setting stage we have $p_i^* - \widetilde{c}_i = \frac{1}{\gamma_{ii}} q_i^*$, hence the CSA also translates into stating that the quantities produced at equilibrium are positive, i.e. $q_i^* \geq 0$ and $q_j^* \geq 0$. Note the assumption of CSA has to be checked whenever simulations using the equilibrium formulas provided here are performed.

[7] Of course, such simplification does not allow to analyse all potential strategic aspects for such an additional CSD. However, the focus below (see Section 3.2.3) is on the impact of an additional CSD on CSD i by lowering c_j and p_j. Annex 7.5 includes a more formal generalisation of the model to n CSDs under the simplifying assumption of symmetry between the CSDs.

2.4.2 Nash equilibrium of the price-setting game $G(\widetilde{c_i})$

Profits for a given CSD i in a period where the costs per transaction are $\widetilde{c_i}$ are:

$$\pi_i = q_i(p_i - \widetilde{c_i})$$

The best-response $p_i^*(p_j)$ of CSD i to a price p_j from CSD j is thus

$$p_i^*(p_j) = \frac{1}{2\gamma_{ii}}(\alpha_i + \gamma_{ii}\widetilde{c_i} + \gamma_{ij}p_j)$$

and similarly

$$p_j^*(p_i) = \frac{1}{2\gamma_{jj}}(\alpha_j + \gamma_{jj}\widetilde{c_j} + \gamma_{ji}p_i)$$

Corresponding profits are

$$\pi_i(p_i^*(p_j), p_j) = \frac{1}{\gamma_{ii}}\left(\frac{1}{2}(\alpha_i - \gamma_{ii}\widetilde{c_i} + \gamma_{ij}p_j)\right)^2$$

and an analogous formula holds for the maximum profits derived by CSD j for any choice of price p_i by CSD i.

At equilibrium of the stage-game G, the level of prices (p_i, p_j) satisfies:

$$\begin{cases} p_i = p_i^*(p_j) \\ p_j = p_j^*(p_i) \end{cases}$$

Let us denote by p_i^* the price set by CSD i in the equilibrium of the stage-game. Then we get, by solving the above system:

$$\begin{cases} p_i^* = \frac{1}{4\gamma_{ii}\gamma_{jj} - \gamma_{ij}\gamma_{ji}}(2\gamma_{jj}\alpha_i + 2\gamma_{ii}\gamma_{jj}\widetilde{c_i} + \gamma_{ij}\alpha_j + \gamma_{ij}\gamma_{jj}\widetilde{c_j}) \\ p_j^* = \frac{1}{4\gamma_{ii}\gamma_{jj} - \gamma_{ij}\gamma_{ji}}(2\gamma_{ii}\alpha_j + 2\gamma_{ii}\gamma_{jj}\widetilde{c_j} + \gamma_{ji}\alpha_i + \gamma_{ji}\gamma_{ii}\widetilde{c_i}) \end{cases}$$

Note in passing that the fixed costs do not influence the equilibrium prices of the price-setting game. That is, the model predicts that CSDs should not try to recover the adaptation costs by increasing their price per transaction compared to the situation with no adaptation costs. CSDs will recover their adaptation costs over time from cost reductions and increased demand for settlement services. Indeed, if CSD i deviates from the Nash equilibrium prices by trying to set p_i higher to compensate for its *adaptation* costs, then, assuming CSD j does maximize its profits and play $p_j^*(p_i)$, CSD i does not maximize its profits (This is because it is not playing its best response $p_i = p_i^*(p_j)$ - if it was, then since CSD j plays $p_j = p_j^*(p_i)$ they would indeed both be playing the Nash equilibrium (p_i^*, p_j^*), a contradiction with CSD i deviating from it).

By applying backward induction to the whole repeated game, we see that this stage-result holds for the whole repeated game, that is, the optimal pricing should not take into account the adaptation costs (see the remark at the end of section 2.4.3). In other words, CSDs which try to pass on their own adaptation costs to their clients will be penalized by earning less than if they had just set the equilibrium price they would have set in the absence of adaptation costs[8]:

[8] It is interesting to note that Clearstream and Euroclear, the two largest European CSD groups, published documents in December 2012 (see Clearstream [4] and Euroclear [10]) that informs the clients their respective clients of an additional charge of EUR 0.094 and EUR 0.10, respectively, per instruction to finance the adaptation costs to T2S *before* their respective migration to T2S. (For the Euroclear Group, this charge applies so far only to the ESES platform, covering the Belgian, Dutch and French markets.) Clearstream furthermore announced that they would not add any margin on top of the direct T2S fee (c_{T2S} below). These decisions appear to confirm the prediction of the model that CSDs will not be able to recover adaptation costs by increasing fees after adaptation to the T2S environment. The model does not capture the possibility to raise fees in today's largely monopolistic market structures before the introduction of T2S.

Proposition 1 *In the finite model, no CSD should increase its prices in an attempt to cover its adaptation costs, whatever they are. If some CSD i deviates in at least one period of the repeated game from the price p_i^* it would have chosen in the absence of adaptation costs, it will make strictly less profits than by playing p_i^*. p_i^* is only a function of the costs $(\widetilde{c_i}, \widetilde{c_j})$ of all the CSDs at this period, the different price-elasticities $(\gamma_{ii}, \gamma_{ij}, \gamma_{ji}, \gamma_{jj})$ and the sizes (α_i, α_j) of the demand for each settlement service.*

Our price-setting model has thus clear-cut behavioural consequences concerning the pricing of the settlement services in the periods subsequent to the adaptation period, suggesting that in a competitive setting CSDs will not increase the price of their settlement services if they try to maximize their profits. In particular, it asserts that if players are rational and profit-seeking, then the costs reduction (that is, a decrease in $\widetilde{c_i}$ and $\widetilde{c_j}$) will be correctly reflected in lower prices (since p_i is a decreasing function of $\widetilde{c_i}$ and $\widetilde{c_j}$), and not kept at the pre-reshaping level in an attempt to cover the previous adaptation costs, least of all increased.

Definition 1 *We say that CSDs engage in price competition, or that there is price competition, if they always play the Nash equilibrium (p_i^*, p_j^*) at each price-setting game, given the current costs.*

Hence, price competition rules out tacit collusion to leave prices higher than the unique competitive equilibrium of the stage-game.

2.4.3 Optimal degree of reshaping chosen in the first period of the game

The optimal degree of reshaping in the first period of the game depends on the trade-off between the total profits a CSD can make after reshaping and the adaptation cost for the reshaping.

Total profits: Using the formula expressing $\pi_i = \pi_i(p_i^*(p_j), p_j)$ from the above section and replacing p_j by the equilibrium price p_j^* chosen by CSD j, we find that the profits derived by CSD i in a price-setting stage game where the equilibrium prices are (p_i^*, p_j^*) are selected is:

$$\pi_i^* = (A_i + B_i\widetilde{c_j} - D_i\widetilde{c_i})^2$$

where

$$\begin{aligned}
A_i &= \frac{1}{2\sqrt{\gamma_{ii}}}(\alpha_i + \frac{\gamma_{ij}}{4\gamma_{ii}\gamma_{jj} - \gamma_{ij}\gamma_{ji}}(2\gamma_{ii}\alpha_j + \gamma_{ji}\alpha_i)) \\
B_i &= \frac{\sqrt{\gamma_{ii}}\gamma_{ij}\gamma_{jj}}{4\gamma_{ii}\gamma_{jj} - \gamma_{ij}\gamma_{ji}} \\
D_i &= \sqrt{\gamma_{ii}}(\frac{2\gamma_{ii}\gamma_{jj} - \gamma_{ij}\gamma_{ji}}{4\gamma_{ii}\gamma_{jj} - \gamma_{ij}\gamma_{ji}})
\end{aligned}$$

A similar expression holds for π_j^*. Note that our assumption that $\gamma_{ii} \geq \gamma_{ji}$ and $\gamma_{jj} \geq \gamma_{ij}$ (Section 2.3) ensures that $D_i \geq 0$.

Remark: Note in passing that the quantity $A_i + B_i\widetilde{c_j} - D_i\widetilde{c_i}$, in the above expression for π_i^*, is positive for any costs $(\widetilde{c_i}, \widetilde{c_j})$ involved, because by the above $A_i + B_i\widetilde{c_j} - D_i\widetilde{c_i} = p_i^* - \widetilde{c_i} = \frac{1}{\gamma_{ii}}q_i^*$ and that our assumption on costs, that is, the CSA, precisely state that $p_i^* - \widetilde{c_i} \geq 0$ and $p_i^* - \widetilde{c_i} \geq 0$ for all "costs involved" (see Section 2.4.1).

The *total profits* are defined as the profits for the whole game, that is, the discounted sum, by an appropriate discount-factor δ, of the profits obtained at each stage game. For example, total profits for CSD i are

$$\pi_i^{tot} = \sum_{t=1}^{N} \delta^{t-1}\pi_{i,t}$$

provided that $\pi_{i,t}$ are the profits realised by CSD i on the t-th stage game. Note that, as usual, $\delta = \frac{1}{1+r}$, if r is the cost of capital mentioned earlier. Now the total profit π_i^{tot} derived by playing the Nash-equilibrium at each stage game is obtained by replacing in the sum each $\pi_{i,t}$ by the appropriate values of π_i^* (note that π_i^* depends on the fixed and variable costs involved at each period, and that these costs are the same for each period $t \geq 2$):

$$\pi_i^{tot} = \widetilde{\delta}(A_i + ((1-b_j)c_j + c_{T2S})B_i - ((1-b_i)c_i + c_{T2S}))D_i)^2 - C_{i,adapt}(b_i)$$

with

$$\widetilde{\delta} = \begin{cases} N & \text{if } \delta = 1 \\ \frac{1-\delta^N}{1-\delta} & \text{if } 0 \leq \delta < 1 \end{cases}$$

Adaptation costs: In the adaptation cost function $C_{i,adapt}(a_i, b_i)$, the parameter a_i captures the reshaping impact on the CSD's fixed cost; the parameter b_i captures the reshaping impact on the CSD's variable cost per transaction (see Table in Section 2.2).

Assume $C_{i,adapt}(a_i, b_i) = \xi_i b_i^2$ for simplicity and to focus on the tradeoff, in a competitive environment, between paying the lump sum today and benefiting from potential costs-per-transaction reductions tomorrow, and investing very little today while putting the T2S fees c_{T2S} on top of one's cost c_i. Note this yields an increasing, convex cost-function, as is probably the case in real life for the adaptation costs facing a CSD with the costs for reshaping the CSD's fixed cost part normalised to 0.[9]

Lemma and theorem for optimal degree of reshaping: If $\xi_i > \widetilde{\delta} c_i^2 D_i^2$, then the best-response function of CSD i at the first period of the game is

$$b_i^*(b_j) = \min(1, \max(0, \psi_i(A_i + ((1-b_j)c_j + c_{T2S})B_i - (c_i + c_{T2S})D_i))) \qquad (BR)$$

with

$$\psi_i = \frac{\widetilde{\delta} c_i D_i}{\xi_i - \widetilde{\delta} c_i^2 D_i^2}$$

But what if $\xi_i < \widetilde{\delta} c_i^2 D_i^2$? The following Lemma presents the answer:

Lemma 1 *Assume $\xi_i < \widetilde{\delta} c_i^2 D_i^2$. Then if CSDs engage in price competition, the best-response function $b_i^*(b_j)$ is the constant function $b_i^*(b_j) = 1$, which represents a complete reshaping decision from CSD i whatever the degree b_j CSD j choose to reshape.*

The proof of Lemma 1 is contained in Annex 7.1.3.

Remark: Note $\xi_i < \widetilde{\delta} c_i^2 D_i^2$ indicates low adaptation costs compared to the square of the CSD transaction costs times a specified function of cross-elasticities $D_i = D_i(\gamma_{ii}, \gamma_{ij}, \gamma_{ji}, \gamma_{jj})$. Lemma 1 shows these low costs result in a complete reshaping ($b_i = 1$) being always more appropriate, regardless of the other CSD's decision concerning its own degree of reshaping b_j.

Together with the best-response expression in case $\xi_i > \widetilde{\delta} c_i^2 D_i^2$ given by (BR), this allows to prove the following theorem:

[9] Of course, it would be possible to solve the model with other, more general forms for the adaptation-cost function, in particular, choosing costs functions that also depend on a_i or on the size of the price-independent demand α_i. For our adaptation cost function $C_{i,adapt}(a_i, b_i) = \xi_i b_i^2$, since there is no dependency on a_i then trivially at equilibrium $a_i = 1$. Indeed, since decreasing fixed costs do not affect the pricing, as previously noted in Proposition 1, profits are maximized when the fixed costs are minimized, which only occurs in $a_i = 1$.

Theorem 1 *If all CSDs are playing the Nash equilibrium of each following stage-game of the repeated game, then the optimal degrees (b_i^*, b_j^*) of reshaping are given by the following formulas:*
(i) If $\xi_i > \widetilde{\delta} c_i^2 D_i^2$ and $\xi_j > \widetilde{\delta} c_j^2 D_j^2$, and assuming the quantities (b_i^{**}, b_j^{**}) defined by

$$\begin{aligned} b_i^{**} &= \frac{\psi_i(A_i + (c_j + c_{T2S})B_i - (c_i + c_{T2S})D_i) - B_i\psi_i\psi_j c_j(A_j + (c_i + c_{T2S})B_j - (c_j + c_{T2S})D_j)}{1 - B_i B_j \psi_i \psi_j c_i c_j} \\ b_j^{**} &= \frac{\psi_j(A_j + (c_i + c_{T2S})B_j - (c_j + c_{T2S})D_j) - B_j\psi_i\psi_j c_i(A_i + (c_j + c_{T2S})B_i - (c_i + c_{T2S})D_i)}{1 - B_i B_j \psi_i \psi_j c_i c_j} \end{aligned}$$

belong to $[0,1]$, then $(b_i^*, b_j^*) = (b_i^{**}, b_j^{**})$.
(ii) If $\xi_i < \widetilde{\delta} c_i^2 D_i^2$ and $\xi_j < \widetilde{\delta} c_j^2 D_j^2$) then $b_i = b_j = 1$.
(iii) If $\xi_i > \widetilde{\delta} c_i^2 D_i^2$ and $\xi_j < \widetilde{\delta} c_j^2 D_j^2$ then $b_i = b_i^*(1) = \max(0, \min(\psi_i(A_i + c_{T2S}B_i - (c_i + c_{T2S}))D_i)))$ and $b_j = 1$.
(iv) If $\xi_i < \widetilde{\delta} c_i^2 D_i^2$ and $\xi_j > \widetilde{\delta} c_j^2 D_j^2$ then $b_i = 1$ and $b_j = b_j^*(1) = \max(0, \min(\psi_j(A_j + c_{T2S}B_j - (c_j + c_{T2S}))D_j)))$.

Proof: The proof consists solely of solving the system:

$$\begin{cases} b_i = b_i^*(b_j) \\ b_j = b_j^*(b_i) \end{cases}$$

in the four different cases, of which depends the form of the best-response function, as mentioned above. Detailed computations are provided in Annex 7.1.3.□

Now playing the Nash equilibrium of each stage game, that is, engaging in price competition, certainly yields a subgame-perfect Nash-equilibrium for our finite horizon game. Moreover, since the price-setting games G only have one Nash-equilibrium, and since this is also the case of the reshaping part of the game (step 1), backward-induction shows that this subgame-perfect Nash equilibrium is *de facto* the only one (see the Remark below for a more formal proof). Hence the optimal degree of reshaping in our finite game setting happen to be those given by Theorem 1. This concludes the analytical resolution of the finite game model.

Remark 1: The unique subgame-perfect Nash equilibrium can be derived more formally by following the generalised backward induction procedure described and proven in Mas-Colell et al. [18], p. 277: Let Γ_E denote the extensive form of the game which consists of the reshaping decision in the first period $t = 1$ of the game and subsequent price-setting games G_t in periods $t = 1, ..., N$. All variables are also indexed with t, so that the unique Nash equilibrium in the final subgame is $(p_{i,N}^*, p_{j,N}^*)$. As a consequence, $(p_{i,N-1}^*, p_{j,N-1}^*)$ is also the equilibrium in the reduced game of the preceding period $N-1$. Continuing this procedure until the first period $t = 1$ gives $(p_{i,1}^*, p_{j,1}^*)$ as the equilibrium. Finally, applying Theorem 1 yields the optimal degrees of reshaping (b_i^*, b_j^*) in $t = 1$. The uniqueness of the Nash equilibrium at each step implies the uniqueness of the subgame-perfect Nash equilibrium of Γ_E derived by backward induction.

Remark 2: The case where $\xi_i > \widetilde{\delta} c_i^2 D_i^2$ and $\xi_j > \widetilde{\delta} c_j^2 D_j^2$ includes potential "corner solutions" which are not covered by the statement of the theorem for simplicity. The complete derivation of all NE, including the so-called "corner-solutions", can be found in the proof of the theorem, in Appendix 7.1.3. Of course, it is these complete results established by distinguishing twelve different cases that should be used for any simulations. Still, the non-corner solution (b_i^{**}, b_j^{**}) quoted by the theorem appears in five out of twelve of these cases.

Remark 3: Notice that, as is often the case in a finite-horizon setting, the model as such does not allow for tacit collusion. Because all the subsequent stage-games which are price-setting games have only one Nash equilibrium and hence only one possible Nash equilibrium payoff vector, backward-induction predicts that repeating the Nash equilibrium of each stage-game is the only subgame perfect

Nash equilibrium of the whole game. Hence the N-period game does not add a lot of insights over the one-period game, that is, the game obtained by setting $N = 1$, except for quantifying the variation of the degree of reshaping to the time-horizon N of the CSDs. We will enlarge the spectrum of the possible strategies by moving to an infinite setting in Section 3, while re-using most of the results proved above.

2.4.4 Simulation results

The model developed previously can be used to produce numerical outputs. With sufficient data concerning prices, costs and settlement volumes, a precise number for the optimal degree of reshaping could be obtained. Because of the insufficient number of data points in our possession, we are not in a position to estimate the building blocks of the model in a methodologically correct way. Hence, the results of this section should better be understood as obtained by *assuming* a numerical range for the various parameters for a large CSD, rather than by *estimating* such ranges from reliable data.

Selection of parameter ranges: Estimations with very limited data on the annual settlement volumes and average prices per transaction of Clearstream and Euroclear[10] were used as a very rough indication to select a plausible order of magnitude for the α- and γ-parameters of the demand functions. Furthermore, we assume for the baseline model of the simulations that the corresponding fixed parameter values are identical for both markets, that is, that $\gamma_{11} = \gamma_{22}$, $\gamma_{12} = \gamma_{21}$, $\alpha_1 = \alpha_2$ and $c_1 = c_2$. The fixed value given to $\gamma_{11}, \gamma_{22}, \gamma_{12}, \gamma_{21}, \alpha_1$ and α_2 as well as the minimum and maximum values reported in the Table below are thus not directly linked to our data estimations.

For the whole the simulation part we choose to work with the costs and prices per settlement *instruction*, in line with the pricing conventions of most CSDs. A settlement transaction generally consists of two separate settlement instructions, so that we could have as well chosen to work with the costs and prices per settlement *transaction*, appropriately adjusting the parameters.

Data on current CSD costs for settlement and on their adaptations costs to T2S are difficult to obtain, too. For the current cost per instruction, i.e. the cost parameters c_1 and c_2, we assume the range between EUR 0.2 and EUR 0.6 as a rough approximation. These figures are based on observed prices for domestic settlement instructions and assume (significant) profit margins, so that the actual average cost of domestic and cross-border instructions is likely to be greater.

In T2S, a settlement instruction will cost EUR 0.15 plus fees for information services which depend on their use, which in turn will depend on the degree and way of a CSDs' adaptation towards T2S. The minimum usage of information services leads to an add-on of EUR 0.012 per instruction, whereas the currently expected average add-on will be EUR 0.042[11]. For the purpose of the simulation, $c_{T2S} = 0.192$ is used.

CSDs' adaptation costs will in reality be highly influenced by their size and their current IT architecture. The more modular their current architecture, the cheaper the adaptation to T2S. This is because T2S can more easily substitute the CSD's current settlement engine with only limited impact on the IT for other services. The more embedded the settlement engine in a CSD's overall architecture, the more costly reshaping becomes. These costs grow with the number of other services that a CSD provides. While these aspects are not directly captured by the model, they justify a very broad range for the adaptation costs $\xi_1 = \xi_2$ of a large CSD, which are assumed between EUR 5 million and EUR 50 million. This range may be set too high even for the largest European CSDs and covers costs for the adaptation of non-settlement related activities. Hence, in the absence of precise data concerning CSDs' costs, the parameters are chosen in a conservative way which reflects a significant variance[12] and is detrimental to reshaping.[13]

[10] The data are taken from the annual reports of Clearstream and Euroclear.

[11] These estimates are based on the report of a T2S pricing workshop with market participants in February 2011, see http://www.ecb.europa.eu/paym/t2s/progress/pdf/ag/mtg13/item-4-3-2011-02-23-report-of-the-pricing-workshop.pdf.

[12] The coefficient of variation, a normalised measure of variability, is greatest for the cost parameter ξ.

[13] In December 2012, Clearstream and Euroclear, the two largest European CSD groups, released publicly (see Clearstream [4] and Euroclear [10]) that they will charge a) EUR 25 million as a maximum to their users for the adaptation costs of Euroclear's ESES platform which covers the Belgian, Dutch and French markets and b) EUR 30 million as maximum "external portion" of Clearstream's T2S investment cost. Their respective total adaptation costs may be somewhat greater.

Thus the main purpose of this section is not to provide a clear-cut answer about the degree of optimal reshaping, but to:

- illustrate how simulations could be used to answer numerically the problem of optimal reshaping, if more data were to become available;
- (partially) solve the problem of the variation of the optimal degree of reshaping b_i with respect to its different arguments. Indeed, it is more illustrative to draw these graphs for some assumed parameter distributions than to try to solve analytically by computing derivatives whose signs are very difficult to determine given the many parameters involved. In particular, the adaptation costs were purposely chosen very high in order to get informative graphs. For example cutting down these high adaptation costs by a factor of 2 would result in a complete reshaping being optimal for almost all other ranges of parameters. This results in graphs being straight lines of equation $b_1 = 1$, a not-so-interesting graph to comment, although it has the strong policy implication of a complete reshaping being the best option.

Below is the Table which summarizes the parameters and parameter ranges assumed for all the simulations in this section.

	min value	max value	fixed value	coefficient of variation
γ_{11}, γ_{22}	150×10^6	230×10^6	190×10^6	0.12
γ_{12}, γ_{21}	90×10^6	150×10^6	120×10^6	0.14
α_1, α_2	100×10^6	180×10^6	140×10^6	0.16
c_1, c_2	0.2	0.6	0.4	0.29
c_{T2S}			0.192	0
N			1	0
$\xi_1 = \xi_2$[14]	5×10^6	50×10^6	27.5×10^6	0.47

Overview of figures: Each of the two-dimensional graph of Figure 1[15] was obtained by selecting a given parameter, displayed on the horizontal axis, and making it vary within the realistic range delimitated by the "min value" and the "max value" indicated in the above Table, while keeping all other parameters constant at the level indicated by the column "fixed values" of the same table. The vertical axis depicts the degree of optimal reshaping, b_1^*. Notably, the magnitude of the impact of each parameter has to be seen in conjunction with its assumed range of variability as evident from the coefficients of variation in the Table above.

Similarly, the three-dimensional graphs of Figure 2 were obtained by selecting two parameters to vary within their range while fixing the others. To interpret these graphs one has to keep in mind they represent the optimal degree of reshaping b_1^* of CSD 1, assuming the other CSD, CSD 2, also acts in an optimal way. Should the other CSD not follow a profit-maximizing rule, the optimal degree of reshaping of CSD 1 would be different, and would need to be computed using the best-response function $b_1^*(b_2)$ given in (BR) with b_2 as an input.

The histogram of Figure 3 shows the distribution of the optimal degree of reshaping b_1^* of CSD 1 assuming each parameter is chosen randomly and independently in the range described by the Table, for 100000 simulations, the uniform distribution being used. We see that, at least for those parameter distributions, complete reshaping is often optimal. A word of caution is warranted there: drawing randomly parameters from their assumed range can certainly give parameter constellations that make no sense. In particular, it could happen that the CSA does not hold. Hence in the histogram of Figure 3 we discarded those parameter constellations for which the CSA does not hold, which amounted to discarding roughly 9% of the total number of parameter constellations drawn.

[14] The same adaptation costs were assumed for the two CSDs at any point in the simulation, which implies that when ξ_1 increases then ξ_2 increases by the same amount.

[15] See Annex 7.9 for all figures based on the simulations.

Finally, the graphs of Figure 4 and 5 depict the results of Monte-Carlo simulations concerning the expected value of the optimal degree of reshaping, subject to knowing each of its parameter in turn. More precisely, for each point to be drawn, about 10 000 of randomly chosen constellations of parameters were chosen, and the average alone was represented. The ranges are still the same as indicated in the Table, and uniform distributions are assumed. The lowest curve of each graph indicates the variance of the simulations carried out to obtain the corresponding expected value of optimal reshaping on the higher curve. Because the shape of the graphs obtained are very similar to those of Figure 1, they provide a positive robustness-check for the sense of variation of the optimal degree of reshaping.

Results of the simulation exercise: We will comment the graphs of Figure 1 by moving from left to right and top to bottom.

The first and second graphs only underline that the greater the domestic market (α_1) or the foreign market (α_2), the higher the optimal degree of reshaping (for a fixed amount of adaptation costs). This is because higher volume of instructions in the market make it easier for a given CSD to recover from its adaptation costs, by benefiting of lower costs per instructions applied to larger volumes. Hence larger markets have higher degree of optimal reshaping than smaller markets, holding other parameters constant. Also, the size of the foreign market is much less an incentive to reshape than the size of the domestic market, as can be seen from the scale of the vertical axis of these graphs.

The third graph indicates that the higher the costs of settlement for CSD 1, c_1, the more incentives it has to reshape, and that there seems to be a linear relation between these two quantities. This confirms intuition, as reshaping cuts part of these costs per settlement instruction.

The fourth graph shows the optimal degree of reshaping b_1 of CSD 1 as a function of the costs per instruction c_2 of its competitor. We see that b_1 is first a (slightly) increasing function of the costs c_2, probably because reshaping allows CSD 1 to put lower prices and draw part of the demand for settlement services originally directed towards CSD 2 to its own market. This positive relation holds up to a certain threshold, with decreasing marginal increases of b_1 as c_2 increase. These decreasing marginal increases are probably due to the strategic decision of CSD 2 itself. Indeed, as we have seen, higher costs per instruction will prompt CSD 2 to reshape to decrease those costs. Past the threshold this effect dominates and the optimal degree of reshaping decreases with higher competitor costs. The optimal degree of reshaping for CSD 1 then becomes flat, corresponding to the situation where CSD 2 has completely reshaped anyway so cannot decrease any further its costs. Hence, when the cost of CSD 2 continues to increase, this has no impact anymore on CSD 1 decision to reshape. Nevertheless, note that all these variations, although interesting to comment analytically, are negligible (vertical scale).

The fifth graph shows that the higher the adaptation cost of a complete reshaping ξ_i, the less useful it is to reshape; this is obvious since the adaptation cost is the price to pay to reshape, which counter-balances future benefits. Hence the higher these costs, the less interesting it is to reshape, holding all other parameters constant. The interesting point is the particular form of this function, which decreases quickly first and then very slowly. This particular shape is probably linked to the convex form assumed for the adaptation cost function. Also noteworthy is the existence of the threshold under which complete reshaping is the optimal solution.

The last four graphs give an idea of the variation of the optimal degree of reshaping with respect to different market elasticities, and hence different equilibrium prices before reshaping.

For example, the sixth graph shows that the optimal degree of reshaping b_1 happens to be a decreasing function of γ_{11}: The effect from cost-saving on each instruction can compensate the relative fewer gains of market share due to lower prices, making reshaping more attractive for lower price sensitivity.[16] Notice that if we were to increase the overall price sensitivity of demand in market 1 by increasing both the parameter γ_{11} and the parameter γ_{12}, the optimal degree of reshaping would increase, as can be seen in the third three-dimensional graph of Figure 2.

The seventh graph indicates that the more CSD 1 is able to capture demand from the market of CSD 2, that is, the more it provides substitutable services, the higher the degree of optimal reshaping of CSD 1. More potential business seems always a good incentive to reshape, as low prices is the main determinant of demand in the building block of our model.

The eighth graph indicates that variations of γ_{21}, which is a parameter characterizing the demand for settlement services of CSD 2, and not of CSD 1, has nevertheless an impact on the optimal degree of reshaping of CSD 1. [17]

The last graph could be explained using the same argument of indirect effects stemming from the reshaping decision of another CSD. Indeed, higher γ_{22} implies less reshaping from CSD 2, as for the case of CSD 1 with γ_{11} explained in the sixth graph, which per se would suggests a higher incentive to reshape for CSD 1. However this effect is dominated by the fact that the price set by CSD 2, even in the absence of reshaping, would be much lower, which, if all other parameters are held constant, translates into less demand for CSD 1 settlement services, and hence less incentive to reshape for CSD 1.

Two other graphs could be drawn: choosing $N > 1$, one could draw the optimal degree of reshaping b_1 as a function of the discount factor δ applied to its future cash flows, and as a function of the time horizon N. These graphs show that the preference of the CSDs for the present has a meaningful impact on the optimal degree of reshaping. Hence, the more patient a given CSD, the more reshaping it will undertake today to derive benefits from it tomorrow, all other things being equal. An interesting point to notice is that making N very large does not necessarily make a complete reshaping the optimal solution: for some parameters values the optimal degree of reshaping converges instead towards a given value inferior to 1.

3 Extending the game to infinity

Although the finite game-theoretic setting described in the previous section is useful in simulations (to add the parameter N or time-horizon of the involved CSDs), and for understanding the model more easily, it is not very rich in terms of subgame perfect Nash equilibrium: as noted earlier, because the game is finite and that each stage game yields only one Nash equilibrium, solving by backward induction to find the subgame-perfect Nash equilibrium yields a single strategy, that is, the one where the unique Nash equilibrium of each stage game is played repeatedly, independently of the past. Hence the myopic profit-maximization strategy, where CSDs set the prices yielding the higher payoffs at each period of the game, is the only subgame-perfect Nash equilibrium.

To allow for other, intertemporal strategies, we need to translate the game to an infinite setting, which basically just means choosing to repeat it indefinitely while assuming $\delta < 1$ to ensure finite payoffs. Of course, playing at each step the Nash equilibrium of the stage game still yields a subgame-perfect Nash

[16] The expectation could be that higher price-sensitivity to its own price would make the price-reduction allowed by cost-reduction through reshaping even more attractive. Nevertheless, this neglects that prices at equilibrium (and thus the volumes at equilibrium) would, in a market with a higher γ_{11}, already be lower than in a market with low γ_{11}. This can be checked analytically by computing the first derivative of p_i^* with respect to γ_{11}, which is negative. Of course, cost-reduction also applies, and reshaping will drop costs, and thus enable lower prices, in the same way as in a low γ_{11} scenario. The graph hence pin point that the positive effect of lowering prices, and hence of reshaping, is less strong in the high γ_{11} scenario than in the low γ_{11} scenario.

[17] Since q_1 is unaffected by $\gamma_{21}$by definition, this impact can only be indirect, due to the strategic decision of CSD 2 to reshape. Indeed, by symmetry, we know from the graph of b_1 as a function of γ_{12} that, when faced with a higher γ_{21}, CSD 2 will reshape more. Hence compared to the low γ_{21} case prices of CSD 2 will be lower and CSD 1, faced with a lower demand for settlement services than in the low γ_{21} case, but still with the same adaptation cost, will find it less profitable to reshape. Hence its lower degree of optimal reshaping. Also interesting is that this effect disappears below a certain threshold (not shown in this graph).

equilibrium, since any player unilaterally deviating from such strategy will be worse off, but there may now be other subgame-perfect Nash equilibrium. Nevertheless, most of other results, such as the formula for the degree of optimal reshaping assuming all CSDs play the Nash equilibrium of each subsequent stage game, still hold for the infinite case. In particular, Lemma 1 as well as Theorem 1 hold as such in the new infinite setting, by taking $\widetilde{\delta} = 1/(1-\delta)$, that is, the limit of the previous $\widetilde{\delta}$ when N tends to infinity. Note the assumption included in Lemma 1 and in Theorem 1 that it gives the optimal reshaping *assuming all players play the Nash equilibrium of each price-setting game afterwards*.

3.1 Delaying or prompting the Reshaping: tacit collusion in the model

3.1.1 Motivation for a new game

In the previous model, CSDs are only allowed to reshape at a single point in time, whereas reshaping is a private decision that is part of the overall business strategy of the CSDs, and thus may be taken at different point in time for each of the CSDs. Importantly, because T2S opens up the domestically-oriented, fragmented and thus largely monopolistic European settlement markets, the decision when to reshape of a given CSD may be dependent on other CSDs' choices. Hence, to further study the possibility of reshaping, we should allow CSDs to reshape at any period of the model. To reach this aim we modify slightly the previous rules of the game as follows:

3.1.2 The new game

Consider the stage game $G_{reshape}$ made up of the three following consecutive steps:

1) CSDs that have not reshaped so far (simultaneously) choose either to reshape now by some degree $b_i > 0$ or not to reshape yet ($b_i = 0$).

2) All CSDs observe the reshaping decisions of others and (simultaneously) set the price of the settlement service they provide.

3) They earn the associated payoffs as implied by the price-setting stage game described in Section 2.3.

The game $G_{reshape}$ is repeated an infinite number of times. The payoff of the whole game is just the discounted payoffs of each stage game, with a discount factor δ (strictly) inferior to 1. Note that the stage-game $G_{reshape}$ is not exactly independent of the past since CSDs which chose to reshape in the past cannot choose to reshape to a different degree in the future. That is, each CSD can only reshape once and the reshaping cannot then be changed in subsequent periods of the game. This impediment is of course unrealistic for a CSD which would like to increase its reshaping but much more close to reality for a CSD who would like to *decrease* its degree of reshaping and recover part of the adaptation costs it had paid in some previous period: this would, as in real-life, not be feasible in our model. Indeed, adaptation costs are costs paid in a lump-sum that cannot be recovered even if a given CSD changes its mind and downsizes the degree of reshaping it wanted. That being said, to allow only reshaping to occur in a precise period and not in progressive (positive) steps is a limitation to our model. Nevertheless, the time frame, that is, the meaning of the "period" of the model, could easily be interpreted as a large enough lapse of time such that reshaping indeed only occurs during one precise period, for each CSD. Moreover, the results we are about to derive assume far-sighted CSDs, which discount the future very little, hence it is unlikely that those results would be affected by the possibility of CSDs to reshape in steps.

An important consequence for a game-theoretic point of view is that although our infinite game consists in playing $G_{reshape}$ over and over an infinite number of times, it does not fall into the usual category of the repeated games because the population of CSDs able to make a decision at step 1) of $G_{reshape}$ is subject to change depending on previous history. For example, all subgames where all CSDs have previously chosen to reshape (to such or such degree) consist solely of playing the price-setting game given this reshaping, that is, step 2 and 3 of $G_{reshape}$, and are thus different from the first $G_{reshape}$ played in the game.

3.2 Two main theorems

3.2.1 Tacit collusion to delay reshaping

We will first show that even for large values of δ, when one would expect that the incentive to reshape is too hard for the CSDs to resist (since they favour future revenues almost as much as the current lump sum they would have to pay for the adaptation costs, and these future revenues tend to infinity when δ tends to 1), there is a subgame-perfect Nash equilibrium in which no CSD reshape when the ratio of the costs $\frac{c_i}{c_j}$ of the involved CSDs satisfies a precise inequality. Indeed, denote by $f_i = f_i(\gamma_{ii}, \gamma_{ij}, \gamma_{ji}, \gamma_{jj})$ the function of the four different price-elasticities of the model defined by:

$$f_i = \frac{B_i}{D_i} = \frac{\gamma_{ij}\gamma_{jj}}{2\gamma_{ii}\gamma_{jj} - \gamma_{ij}\gamma_{ji}}.$$

Then we have:

Theorem 2 *Assume $\frac{1}{f_j} < \frac{c_i}{c_j} < f_i$. Then there exists $\overline{\delta} \in]0,1[$ such that for any discount factor $\delta \in]\overline{\delta}, 1[$, collusion for no reshaping can be sustained in a subgame perfect Nash-equilibrium. The result holds even when price competition is assumed.*

That is, there exists a subgame-perfect Nash equilibrium in which no CSD ever reshape when playing it.

The proof of Theorem 2 is given in Annex 7.2.1 and uses the following trigger strategy (S):

(S) *Do not reshape if no player has reshaped so far. If another player has reshaped and you have not reshaped, then reshape. At each price-setting game (step 2) play the unique Nash equilibrium, given all the costs involved.*

Remark 1: Since $f_i = \frac{B_i}{D_i} = \frac{\gamma_{ij}\gamma_{jj}}{2\gamma_{ii}\gamma_{jj}-\gamma_{ij}\gamma_{ji}}$, the condition $\frac{1}{f_j} < \frac{c_i}{c_j} < f_i$ of the theorem can be rewritten directly in terms of the elasticities as $2\frac{\gamma_{jj}}{\gamma_{ji}} - \frac{\gamma_{ij}}{\gamma_{ii}} < \frac{c_i}{c_j} < \frac{\gamma_{ij}\gamma_{jj}}{2\gamma_{ii}\gamma_{jj}-\gamma_{ij}\gamma_{ji}}$. The condition for CSD i to be dissuaded to reshape by CSD j own potential reshaping is:

$$c_j f_i > c_i$$

We see that the larger the costs of CSD j, the more likely this is going to occur. Indeed, CSD j can more convincingly threaten to reshape and engage in tougher price competition if it is currently operating at higher costs than if it is already efficient. Now, for the strategy profile to be a Nash equilibrium, a similar inequality is required for CSD i costs, that is, $c_i f_j > c_j$.

Remark 2: It is possible to determine explicitly a value of $\overline{\delta}$ for which Theorem 2 holds. Indeed, (see Annex 7.2.2) we have that $c_i < c_j f_i$ is a necessary and sufficient condition for the inequality ($*$) of the Proof of Theorem 2 (given in Annex 7.2.1) to hold. we see that the result of the theorem holds for $\overline{\delta} := \max(1 - \frac{c_i^2 D_i^2}{\xi_i}, 1 - \frac{c_j^2 D_j^2}{\xi_j}, \frac{B^2-A^2-\xi_i}{B^2-C^2-\xi_i}, \frac{B^2-A^2-\xi_j}{B^2-C^2-\xi_j})$ if $\xi < B^2 - C^2$ and for $\overline{\delta} := \max(1 - \frac{c_i^2 D_i^2}{\xi_i}, 1 - \frac{c_j^2 D_j^2}{\xi_j})$ if $\xi > B^2 - C^2$.

Remark 3: Using the expression for $\overline{\delta}$ provided by Remark 2, we see that for adaptation costs high enough, more precisely for $\xi_i > B^2 - C^2$ and $\xi_j > B^2 - C^2$, any further increase in both adaptation costs translates into an increase of our value of $\overline{\delta}$, since $\overline{\delta} := \max(1 - \frac{c_i^2 D_i^2}{\xi_i}, 1 - \frac{c_j^2 D_j^2}{\xi_j})$ and $1 - \frac{c_i^2 D_i^2}{\xi_i}$ (resp.$1 - \frac{c_j^2 D_j^2}{\xi_j}$) are increasing with ξ_i (resp. with ξ_j). This *suggests* that in that case, high adaptation costs do not increase but decrease the chance of collusion among CSDs. More precisely, the more costly the adaptation, the higher the discount factor above which there can be collusion (in the proof), hence the less likely collusion in that setting. This is rather counter-intuitive and stems from the fact that to make tacit collusion possible in our non-cooperative setting, CSDs must be credible when they frighten the others to reshape themselves should any CSD move towards reshaping. This tacit threat limits the benefits derived from reshaping for any deviating CSD. Yet, if reshaping costs are too high, then the threat become less credible, and the possibility of collusion based on that threat diminishes. Hence, a higher $\overline{\delta}$, that is, a greater valuation for the future cash flows compared to the current ones, need to be assumed to make such a threat credible. Now if $\xi_i < B^2 - C^2$ and $\xi_j < B^2 - C^2$, the effect of an increase in adaptation costs is not clear-cut and, even assuming that no better collusion strategies can be designed than our trigger strategy, depends on how the various parameters of the model compare to each other's. Indeed, $\frac{B^2-A^2-\xi_i}{B^2-C^2-\xi_i}$ is, when $\frac{c_i}{c_j} < f_i$, decreasing in ξ_i, hence if $\overline{\delta} := \max(1 - \frac{c_i^2 D_i^2}{\xi}, 1 - \frac{c_j^2 D_j^2}{\xi}, \frac{B^2-A^2-\xi_i}{B^2-C^2-\xi_i}, \frac{B^2-A^2-\xi_j}{B^2-C^2-\xi_j}) = \frac{B^2-A^2-\xi_i}{B^2-C^2-\xi_i} > \max(1 - \frac{c_i^2 D_i^2}{\xi}, 1 - \frac{c_j^2 D_j^2}{\xi}, \frac{B^2-A^2-\xi_j}{B^2-C^2-\xi_j}) =: \delta_0$ then an increase in ξ_i will first translate into a decrease of $\overline{\delta}$ (until $\overline{\delta}$ reaches δ_0) indicating that higher adaptation costs increase the chance of collusion. This is of course because, holding all other factors constant, the higher the adaptation costs a given CSD i faces for itself, the less profitable it is for the CSD to reshape and thus the more likely it is going to collude to delay the reshaping (see inequality $(*)$ in Proof of Theorem 2 given in Annex 7.2.1).

Remark 4: Under the conditions of Theorem 2, that is, when $c_i < f_i c_j$ and $c_j < f_j c_i$, the strategy (S) is, by the proof of the theorem, a Nash equilibrium for high discount factors. Another obvious Nash equilibrium under such conditions is for both CSDs to reshape completely at the first period of the game. Call this strategy (1). Now, it can readily be checked from the form of the one-period payoff function that, for high discount factors, the strategy (S) is (strictly) Pareto-dominated by the strategy (1) if both f_i and f_j are superior to 1. If both f_i and f_j are inferior to 1, then the converse holds: the strategy (S) (strictly) Pareto-dominates the strategy (1). Nevertheless we will prove in the discussion that in such a case (S) cannot be a subgame perfect Nash equilibrium[18]. When f_i is inferior to 1 and f_j superior to 1 or the reverse, then none of this two equilibrium Pareto-dominates the other, since one of the player is always worse off in a situation than in the other.

Remark 5: Theorem 2 corresponds to a situation where higher aggregate costs per transaction translates into higher profit margins for CSDs. Indeed, replacing f_i by its expression as a function of the elasticities it can easily be shown that $c_i < f_i c_j$ is equivalent to the profit margin per transaction $p_i^* - \widetilde{c}_i$ being higher when no CSDs reshape than when both fully reshape.

3.2.2 A sufficient condition for immediate reshaping

Theorem 2 gives a condition, namely that $c_i < f_i c_j$ and $c_j < f_j c_i$, under which there is possibility of collusion not to reshape among CSDs, when they value the future relatively to the present enough. The natural question is then to determine if, when the condition is not fulfilled, reshaping by at least one of the CSD always happens. Interestingly, it turns out to be the case:

Theorem 3 *Assume $c_i > f_i c_j$. Then for any discount-factor close enough to 1, CSD i will always completely reshape in any subgame perfect Nash equilibrium consistent with price competition. In particular, in the presence of price competition, the other CSD j cannot deter CSD i from reshaping in any credible way. More precisely, if $\xi_i < \pi_i(1,0)^2 - \pi_i(1,1)^2$, then for $\delta \geq 1 - \frac{c_i^2 D_i^2}{\xi_i}$ CSD i will always completely reshape, while if $\xi_i > \pi_i(1,0)^2 - \pi_i(1,1)^2$, CSD i will always completely reshape for $\delta \geq \max(1 - \frac{c_i^2 D_i^2}{\xi_i}, \frac{\pi_i(1,0)^2 - \pi_i(0,0)^2 - \xi_i}{\pi_i(1,0)^2 - \pi_i(1,1)^2 - \xi_i})$.*

[18] We need to establish that the conditions $c_i < f_i c_j$ and $c_j < f_j c_i$ are also necessary for (S) to be a subgame perfect Nash equilibrium, and this will be a consequence of the next section 's results. The fact that (S) as a subgame perfect Nash equilibrium cannot Pareto-dominate (1) then easily follows).

Proof. Since $\delta \geq 1 - \frac{c_i^2 D_i^2}{\xi_i}$, we have, by a reasoning similar to the proof of Theorem 2, that whatever the degree $b_j > 0$ CSD j choose to reshape, or if it refrains from reshaping ($b_j = 0$), $b_i = 1$ is the profit maximization choice of CSD i, should it choose to reshape. Note that it does not mean, *a priori*, that CSD i will choose to reshape: it could, in principle, as in Theorem 2, be dissuaded to do so by other CSDs' strategies.

The rest of the proof consists in showing this is never the case. Hence, the results will be that CSD i does choose to reshape, and that it chooses $b_i = 1$, that is, a complete reshaping, when it does so.

Now, the worst *credible* punishment CSD j could inflict on CSD i in order to induce it not to reshape is to reshape as soon as CSD i did by a degree $b_j^*(b_i)$ and engage in price competition afterwards, that is, play repeatedly the Nash equilibrium of the price-setting game. This would yield a payoff for CSD i of $\pi_i(1, b_j^*(b_i))$ instead of the $\pi_i(0,0)$ it received when it was not reshaping or of the $\pi_i(1,0)$ it received when it chose to reshape, in all subsequent price-setting stage games of the repeated game. Since $\pi_i(1, b_j^*(b_i)) \geq \pi_i(1,1)$, if we show that CSD i will still prefer to reshape if it earns, from the moment CSD j punishes it by also reshaping, $\pi_i(1,1)$ instead of $\pi_i(1, b_j^*(b_i))$, we will have proved that CSD i best action is indeed to reshape. But from the proof of Theorem 2, we know that whenever the inverse inequality of $(*)$ holds, CSD i is better off not playing the collusion strategy and reshaping, even though CSD j will punish it by reshaping itself and by engaging in harder price competition afterwards.

Assume by contradiction that $(*)$ holds. Then since by the proof of Remark 2 of Theorem 2 we have that $(*)$ holds for high values of δ if, and only if, $c_i < f_i c_j$, we immediately get a contradiction with $c_i > f_i c_j$. This proves the first part of Theorem 3.

To be more precise and prove the second part of Theorem 3, providing explicit values for the range of δ for which no infinite delaying of the reshaping decision can occur, we need to make a more precise use of Remark 2 of Theorem 2 by separating two cases:

If $\xi_i < B^2 - C^2$ then for $(*)$ to hold for some δ, in particular for some $\delta \geq 1 - \frac{c_i^2 D_i^2}{\xi_i}$, we need $c_i < f_i c_j$, which gives a contradiction with our assumption $c_i > f_i c_j$. Hence $(*)$ does *not* hold for *any* $\delta \geq 1 - \frac{c_i^2 D_i^2}{\xi_i}$, which gives our result than for $\delta \geq 1 - \frac{c_i^2 D_i^2}{\xi_i}$ CSD i will always reshape.

If $\xi_i > B^2 - C^2$ then we need $c_i \leq f_i c_j$ for $(*)$ to hold for $\delta \geq \frac{B^2 - A^2 - \xi_i}{B^2 - C^2 - \xi_i}$. Since by our assumptions $c_i > f_i c_j$, we deduce that for $\delta \geq \max(1 - \frac{c_i^2 D_i^2}{\xi_i}, \frac{B^2 - A^2 - \xi_i}{B^2 - C^2 - \xi_i})$, CSD i will always reshape. This concludes the proof. $\square$

In order to understand Theorem 3 better, we look at some of its consequences in a symmetric setting. We first define explicitly what we mean by "symmetric":

Definition 2 *A market is said symmetric if, and only if, $\alpha_i = \alpha_j$, $\gamma_{ij} = \gamma_{ji}$ and $\gamma_{ii} = \gamma_{jj}$. Two CSDs' transaction costs are symmetric if, and only if, $c_i = c_j$.*

Remark: The condition $\alpha_i = \alpha_j$ is not necessary for any of the propositions below. Indeed, an interesting point to notice is that none of the conditions given in the theorem refers to the parameters α_i and α_j, but only to elasticities. Hence, it is the sensitivity to prices more than the size of the market (of which (α_i, α_j) is a proxy) that matters. We included the condition $\alpha_i = \alpha_j$ in the Definition to reflect a completely symmetric demand function for both CSDs in a symmetric market.

Proposition 2 *Assume that market and CSDs transaction costs are symmetric, that $\gamma_{ij} < \gamma_{ii}$ and that CSDs engage in price competition. Then for any discount-factor $\delta \geq 1 - \frac{c_i^2 D_i^2}{\xi_i}$, both CSDs will completely reshape in any subgame perfect Nash equilibrium.*

Proof. It is enough to show that the condition $c_i > f_i c_j$ of Theorem 3 is satisfied. Because $c_i = c_j$ this condition is equivalent to $1 > f_i$. By definition of $f_i = \frac{\gamma_{ij}\gamma_{jj}}{2\gamma_{ii}\gamma_{jj} - \gamma_{ij}\gamma_{ji}}$ this is equivalent to $\gamma_{ij}(\gamma_{jj} + \gamma_{ji}) < 2\gamma_{ii}\gamma_{jj}$, which is true since by symmetry and the usual assumptions we have $\gamma_{ij} < \gamma_{ii}$ and $\gamma_{jj} + \gamma_{ji} < 2\gamma_{jj}$. Hence by Theorem 3 both CSDs will completely reshape in any subgame perfect Nash equilibrium consistent with price competition. $\square$

Note the condition $\gamma_{ij} < \gamma_{ii}$ only translates the property that a price level increase should lead to a lower *aggregate* demand, as explained in Remark 2 of Section 2.3, and is hence a rather natural assumption. From Proposition 2 we can easily get a similar result that predicts no collusion and the reshaping of at least one of the two CSDs in case the CSDs transaction costs are not symmetric:

Proposition 3 *Assume the market is symmetric and that $\gamma_{ij} < \gamma_{ii}$, and that CSDs engage in price competition. Then for all discount-factor $\delta \geq \max(1 - \frac{c_i^2 D_i^2}{\xi_i}, 1 - \frac{c_j^2 D_j^2}{\xi_j})$, the CSD with higher costs will reshape in any subgame perfect Nash equilibrium.*

Proof. Without loss of generality, we can assume CSD i has the highest costs, i.e. $c_i \geq c_j$. Then $\frac{c_i}{c_j} \geq 1$ while by Proposition 2 $1 > f_i$. Hence $\frac{c_i}{c_j} > f_i$ and Theorem 3 implies CSD i will reshape. $\square$.

Admittedly, all these non-collusion results assume price competition in the subsequent periods. Nevertheless, allowing for price collusion does not dissuade CSDs from reshaping, as can be seen in Annex 7.7. Simply put, this is because allowing for price-collusion allows the CSDs to pocket the unit-transaction cost-reductions instead of reflecting them in their prices.

3.2.3 Discussion of the main results

Theorem 2 asserts that for some range of parameters – more precisely, when $c_j f_i > c_i$ and $c_i f_j > c_j$ – there can be a perpetual delaying of the reshaping decision by the CSDs, resulting in none of the CSDs ever reshaping. This is the case because by not reshaping as long as the others do not reshape, a given CSD avoids to pay the adaptation costs, while the extra-rent it would extract through cost-reduction and earning market shares by reshaping seem no longer worth it when it knows that the other CSDs would then also reshape and compete with it on the same grounds. Of course, there is a first mover advantage: the first CSD to reshape would make more profits in the following period. But when the discount factor exceeds a certain threshold, this short-term benefit becomes irrelevant since the high discount factor gives an almost equal weight to the future than to the present.

Also worth noticing, the tacit collusion for not reshaping described by Theorem 2 holds for any value of δ higher than some threshold $\overline{\delta}$, which means no reshaping can potentially occur for far-sighted CSDs which attribute a high value to future payoffs. Hence contrary to our original model of Section 2, where the decision to reshape or not depended on the trade-off between paying now high adaptation costs and benefiting more from transaction costs reduction later, or avoiding this investment now and bearing the high costs per transaction in the subsequent periods of the model, the decision concerning reshaping depends here entirely on the expected strategy of the other CSD[19].

The suboptimal implication of Theorem 2 in terms of efficiency and investors' welfare should be accessed keeping in mind that Theorem 2 does not indicate that the *only* credible (subgame perfect) outcome is the one in which reshaping is indefinitely delayed: even under the assumptions that $c_j f_i > c_i$ and $c_i f_j > c_j$, there may be many more subgame perfect Nash equilibria. For example, one of these other possible subgame perfect Nash equilibrium is the one were CSDs reshape by their respective optimal degree as specified by Theorem 1 at the very first period of the model, and then set prices accordingly. Call this strategy (1), and call (S) any strategy yielding a perpetually delayed reshaping. It can easily be proved that strategy (S) never Pareto-dominates (1) as a Nash-equilibrium. Indeed by the previous results if (S) is a NE it means that $c_i < c_j f_i$ and $c_j < c_i f_j$. But for (S) to Pareto-dominate (1) we must have, by Remark 4 of Section 3.2.1, $f_i < 1$ and $f_j < 1$. This implies $c_i < c_j f_i < c_j$ and $c_j < c_i f_j < c_i$, a

[19] A comparison with the *infinite* version of our first model (Section 2), where the decision to reshape can only be taken in the first period of the model, gives an insight of how much the decision to delay reshaping in Theorem 2 relies on the mere possibility of CSDs to freely choose the time of their reshaping. Typically, in a setting where CSDs can only take the decision to reshape at the first period of the model, not only would there be no tacit collusion to delay reshaping, but, if $\delta \geq 1 - \frac{c_i^2 D_i^2}{\xi}$ – as is assumed in the Proof of Theorem 2 – then by Lemma 1 the reshaping would be *complete* (meaning CSD i would choose $b_i = 1$).

contradiction. On the contrary, we know by Remark 4 of Section 3.2.1 that if $f_i > 1$ and $f_j > 1$, which is the case for example if costs are symmetric ($c_i = c_j$) and if (S) is a NE, that the strategy (1), which implies reshaping in the first period, Pareto-dominates strategy (S). Hence both player will always be better off by reshaping immediately than by delaying ad infinitum the reshaping.

Besides, some other factors are likely to mitigate the results of Theorem 2:

- first, Theorem 2 applies only to the specific constellations of parameters satisfying the rather restrictive conditions that $c_j f_i > c_i$ and $c_i f_j > c_j$. Moreover, we proved in Theorem 3 that these conditions are also necessary, that is, if it ever happens that $c_i \geq f_i c_j$ or that $c_j \geq f_j c_i$, then at least one of CSD i or j will reshape, whatever the other CSD's strategy: there can be no deterrence for no-reshaping in that context, and (S) is not a subgame perfect Nash equilibrium anymore.

- second, Theorem 2 applies as such only for a market with no new entrants. A new entrant could create its infrastructure by optimally adjusting it to T2S following a Greenfield approach, which would be similar to a complete reshaping of an existing CSD. Its investment costs would play a similar role as the adaptation costs in the model, while its transaction cost would be likely to be less than the average cost of the market. Now, if we want to predict the outcome for the whole CSD industry, then as noted previously in Section 2.3, we have to interpret CSD j as representing the whole CSD industry, minus CSD i. The new entrant CSD would drive down the average costs for the whole settlement industry except CSD i, c_j. This dynamic effect of new entrance is thus pushing the inequality $c_j f_i > c_i$, which as we saw is necessary for collusion not to reshape among existing CSDs, towards the opposite inequality, that is, towards $c_j f_i \leq c_i$. As soon as this inequality is satisfied, it is Theorem 3, instead of Theorem 2, which applies, meaning CSD i will itself re-shape (whatever the other CSDs' plans). Hence, new entrants have a positive effect on the decision of reshaping by other, pre-existing, CSDs.

- third, because of the technical irreversibility of reshaping, the strategy supporting the continual delaying of reshaping can only be a "pure" trigger strategy, hence the resulting equilibrium is very fragile: should any player ever deviate from it, all players will reshape and the suboptimal situation is replaced by the optimal one, where costs per transactions are reduced both for CSDs, through a decrease of the costs in the model ($\widetilde{c_i}$ and $\widetilde{c_j}$), and for market participant, through a decrease of the prices (p_i and p_j), they are being charged.

Let us discuss further the condition under which tacit collusion is possible, but in terms of the various price-elasticities involved instead of in terms of the costs as in Remark 1 of Theorem 2. By definition, $f_i = f_i(\gamma_{ii}, \gamma_{ij}, \gamma_{ji}, \gamma_{jj}) = \frac{\gamma_{ij}\gamma_{jj}}{2\gamma_{ii}\gamma_{jj} - \gamma_{ij}\gamma_{ji}}$. Computing the first partial derivatives with respect to the relevant parameters, we see that f_i is decreasing in both parameters γ_{ii} and γ_{jj} while it is increasing in both cross-elasticities γ_{ij} and γ_{ji}. Hence, the higher the substitution / competition effect, i.e. the higher γ_{ij} and γ_{ji} compared to γ_{ii} and γ_{jj}, the more likely we end up with $c_i < f_i c_j$ (taking c_i and c_j as given here), and hence with the possibility that CSD j can influence CSD i strategy by refraining from reshaping but threatening to do so should CSD i reshape. If a similar condition holds for CSD i vis-a-vis CSD j, the result is a possible tacit collusion not to reshape. As noted in Remark 1 of Theorem 2, the threat for CSD j to reshape (and punish CSD i by competing at lower prices) is only credible if CSD j current costs are high (more precisely, if $c_j > \frac{c_i}{f_i}$). The deterrence comes precisely from the fact that should CSD j reshape, it will get costs similar to those of CSD i which will thus lose the competitive advantage it had benefited from before. Now, the elasticities γ_{ij} and γ_{ji} are precisely what allow this competitive advantage to be used. Hence, the higher the substitution effects, the more convincing CSD j threat to reshape and draw down its costs, and the more likely the collusion, should a similar condition hold for CSD i vis-a-vis CSD j. Conversely, the lower the substitution effects, the less likely the threat and influence of CSD j decision on CSD i. The extreme case would be completely independent and non-competing CSDs, i.e. $\gamma_{ij} = \gamma_{ji} = 0$, then the decisions to reshape or not by the CSDs are completely independent from one another and solely based on the respective costs.[20]

Hence our model provides the result, counter-intuitive at first sight, that greater competition can actually hinder the reshaping process rather than favouring it. However, this result is based on the assumption that substitution and thus competition effects are constant over time. We expect that explicitly allowing for increasing substitution / competition effects over time, which is expected as a consequence of the introduction of T2S, the CSD regulation and further harmonisation initiatives, would reverse this result. We leave the explicit modelling of increasing competition over time to future research.

4 Introducing delayed observability of the reshaping decision

The previous theorems were derived in a setting where the decision to reshape of other CSDs is immediately observable by all the CSDs in the next step: this comes from our definition of the game in Section 3.1.2. For example, we proved that, in such a game, tacit collusion to delay infinitely the reshaping can be sustained in a subgame perfect Nash equilibrium, even though reshaping would be done straight away if the decision to reshape could only be taken at a single point in time like in the infinite version of the model of Section 2. Important to note is that to design the trigger strategy (S) which is essential to the proof of Theorem 2, CSDs need to be aware of other CSDs decision to reshape. But one could argue that such an information is internal to the reshaping CSD and should not be known to the market (except by interpreting possible price-signals).

Suppose we modify the rules of the game to allow a given CSD to become aware of the decision of other CSD to reshape only N_0 periods after that decision was taken. Then, a similar result holds in this game of imperfect information, which provides some evidence of the robustness of Theorem 2 when the decision to reshape or not become eventually known by the market, possibly after some lag.

Theorem 4 *Assume $\frac{1}{f_j} < \frac{c_i}{c_j} < f_i$ and that the game is modified such that any given CSD only becomes aware of the decision of other CSDs' decision to reshape after playing N_0 times the other stage-game. Then there exists $\overline{\delta} \in]0,1[$ such that for any discount factor $\delta \in]\overline{\delta}, 1[$, tacit collusion for no reshaping can be sustained in a subgame perfect Nash-equilibrium, even if we assume CSDs engage in price competition.*

The Proof is given in Annex 7.2.3.

Remark: In the proof, inequality $(**)$ becomes[21] $\pi_i(0,0) > (1-\delta^{N_0})\pi_i^{ab} + \delta^{N_0}\pi_i(b_i,1)$ when no adaptation costs are assumed ($\xi_i = \xi_j = 0$). This assumption is only made in this Remark to allow to solve $(**)$[22]:

$$\delta > \sqrt[N_0]{\frac{\pi_i^{ab} - \pi_i(0,0)}{\pi_i^{ab} - \pi_i(b_i,1)}} =: \overline{\delta}$$

Hence the higher N_0, i.e. the less observable the reshaping, the greater the value of $\overline{\delta}$ as obtained in the proof of the theorem. This *suggests*[23] that the less observable the reshaping, the more difficult it becomes to sustain tacit collusion.

We now provide a theorem similar to Theorem 3 in which the decision to reshape always occurred in a delayed observability setting.

[20] Similarly to this paragraph, Annex 7.8 shows more formally that the set of parameters for which tacit collusion is sustainable in a high competition / substitution environment is larger than in a low competition / substitution environment.

[21] Note in passing that $\pi_i(0,0) > (1-\delta^{N_0})\pi_i^{ab} + \delta^{N_0}\pi_i(b_i,1)$ is also a sufficient condition for $(**)$ to be true. Hence when $\delta > \overline{\delta}$ then $(**)$ is also true, whatever the adaptation costs, and the proof of the theorem applies.

[22] Indeed, there is no closed-form formula for solving general polynomials of degree more than 4 (Galois).

[23] This only *suggests* it, and does not *prove* it, since nothing indicates that the threshold obtained above is an optimal one (that is, the lowest one). There may exist a threshold independent of N_0 above which tacit collusion occurs in a subgame perfect Nash equilibrium (possibly the strategy profile (S_{N_0}) defined in the proof of Theorem 4, or another strategy profile).

Theorem 5 *Assume $c_i > f_i c_j$. Then for any high-enough discount-factor δ, CSD i will always completely reshape in any subgame perfect Nash equilibrium consistent with price competition. In particular, the other CSD j cannot deter CSD i from reshaping in any credible way.*

Proof. First we prove that for a high-enough discount-factor δ, CSD i, if it ever chooses to reshape, will choose to reshape completely, that is, will choose $b_i = 1$.

By the proof of Theorem 4 we have that the profit of CSD i, should it choose to reshape by some degree b_i, would be:

$$\pi_i^{dev}(b_i) = \frac{1-\delta^{N_0}}{1-\delta}\pi_i^{ab} - \xi_i b_i^2 + \frac{\delta^{N_0}}{1-\delta}\pi_i(b_i, b_j^*(b_i))$$

Note that, trivially, $\pi_i(b_i, b_j^*(b_i))$ is maximum for $b_i = 1$. We have to prove that $\pi_i^{dev}(b_i)$ is also maximal for $b_i = 1$, knowing that π_i^{ab} is a function of b_i too. From the expression of $\pi_i^{dev}(b_i)$ above we have that for any $b_i \in [0, 1]$:

$$(1-\delta)\pi_i^{dev}(b_i) - \delta^{N_0}\pi_i(b_i, b_j^*(b_i)) = (1-\delta^{N_0})\pi_i^{ab} - (1-\delta)\xi_i b_i^2$$

Since $\pi_i(b_i, b_j, p_i, p_j)$ is bounded, so is π_i^{ab}. Furthermore $\xi_i b_i^2$ is also bounded, since $\xi_i b_i^2 \in [0, \xi_i]$. Hence when δ tends to 1, the right-hand side of this equation tends to 0, which proves that whatever b_i, $(1-\delta)\pi_i^{dev}(b_i) - \delta^{N_0}\pi_i(b_i, b_j^*(b_i))$ tends to 0 when δ tends to 1. Clearly, $(1-\delta)\pi_i^{dev}(b_i)$ reaches its maximum as a function of b_i whenever $\pi_i^{dev}(b_i)$ do so. Assume now by contradiction that $\pi_i^{dev}(b_i)$ is maximum for a $\widehat{b_i} < 1$. Then by applying the previous result to both $b_i = 1$ and $b_i = \widehat{b_i}$ we get that both $(1-\delta)\pi_i^{dev}(1) - \delta^{N_0}\pi_i(1, b_j^*(1))$ and $(1-\delta)\pi_i^{dev}(\widehat{b_i}) - \delta^{N_0}\pi_i(\widehat{b_i}, b_j^*(\widehat{b_i}))$ tends to 0 when δ tends to 1, hence the sum of these two quantities, that is,

$$(1-\delta)(\pi_i^{dev}(1) - \pi_i^{dev}(\widehat{b_i})) - \delta^{N_0}(\pi_i(1, b_j^*(1)) - \pi_i(\widehat{b_i}, b_j^*(\widehat{b_i})))$$

also tends to 0 when δ tends to 1. But since $\pi_i(b_i, b_j^*(b_i))$ is maximum *only at* $b_i = 1$, and that $\widehat{b_i} < 1$, we have that $\pi_i(1, b_j^*(1)) - \pi_i(\widehat{b_i}, b_j^*(\widehat{b_i})) =: \eta > 1$, hence $\delta^{N_0}(\pi_i(1, b_j^*(1)) - \pi_i(\widehat{b_i}, b_j^*(\widehat{b_i})))$ tends to η when δ tends to 1. Hence

$$(1-\delta)(\pi_i^{dev}(1) - \pi_i^{dev}(\widehat{b_i}))$$

tends to $\eta > 0$ when δ tends to 1. In particular there exists a δ such that $(1-\delta)(\pi_i^{dev}(1) - \pi_i^{dev}(\widehat{b_i}))$ is close enough to η to be strictly superior to $\eta/2$ and hence strictly positive. Hence $\pi_i^{dev}(1) > \pi_i^{dev}(\widehat{b_i})$, a contradiction with the definition of $\widehat{b_i}$ and the assumption that $\widehat{b_i} < 1$. This proves that $\widehat{b_i} = 1$ and that $\pi_i^{dev}(b_i)$ is maximum for $b_i = 1$.

Note that it does not mean, *a priori*, that CSD i will indeed choose to reshape: it could, in principle, as in Theorem 4, be dissuaded to do so by the other CSD strategy.

The rest of the proof consists in showing this is never the case. Hence, the results will be that CSD i do choose to reshape, and that it chooses $i = 1$, that is, a complete reshaping, when it does so.

Now, the worst *credible* punishment CSD j could inflict on CSD i in order to induce it not to reshape is to reshape as soon as CSD i did by a degree $b_j^*(b_i)$ and engage in price competition afterwards, that is, play repeatedly the Nash equilibrium of the price-setting game. This would yield a payoff for CSD i of $\pi_i(1, b_j^*(b_i))$ instead of the $\pi_i(0, 0)$ it received when it was not reshaping or of the $\pi_i(1, 0)$ it received when it chose to reshape, in all subsequent price-setting stage game of the repeated game. Since $\pi_i(1, b_j^*(b_i)) \geq \pi_i(1, 1)$, if we show that CSD i will still prefer to reshape if it earns, from the moment CSD j punishes it by also reshaping, $\pi_i(1, 1)$ instead of $\pi_i(1, b_j^*(b_i))$, we will have proved that CSD i best action is indeed to reshape. But from the proof of Theorem 4, we know that whenever the inverse inequality of $(*)$ holds, CSD i is better off not playing the collusion strategy and reshaping, even though CSD j will punish it by reshaping itself and engaging in harder price competition afterwards. Hence we only need to prove that under our assumptions (that is, that $c_i > f_i c_j$), the inverse inequality of $(*)$ holds.

$c_i > f_i c_j$ implies successively:

$$\begin{array}{rcl} D_i c_i - B_i c_j & > & 0 \\ \pi_i(0,0) & < & \pi_i(1,1) \\ \pi_i^{ab} - \pi_i(0,0) & > & \pi_i^{ab} - \pi_i(1,1) \\ \delta & < & 1 < \sqrt[N_0]{\dfrac{\pi_i^{ab} - \pi_i(0,0)}{\pi_i^{ab} - \pi_i(1,1)}} \\ \delta^{N_0} & < & \dfrac{\pi_i^{ab} - \pi_i(0,0)}{\pi_i^{ab} - \pi_i(1,1)} \\ \pi_i(0,0) & < & (1-\delta^{N_0})\pi_i^{ab} + \delta^{N_0}\pi_i(1,1) \end{array}$$

for all discount factor δ. Now since $(1-\delta)\xi_i$ tends to 0 when δ tends to 1, we have that, for δ high enough:

$$\pi_i(0,0) < (1-\delta^{N_0})\pi_i^{ab} + \delta^{N_0}\pi_i(1,1) - (1-\delta)\xi_i$$

and this is precisely the inverse inequality of $(*)$.$\square$

We assumed in this section that the reshaping decision is not observable for N_0 periods, but that it became public knowledge after. Hence, one could argue that working under the assumption of *complete* non-observability of the reshaping decision, where no other CSD than the reshaping CSD itself know about its decision to reshape, would have a different effect on the possibility for CSDs not to reshape. However, there are strong limits to the extent that CSDs can hide information on a crucial strategic decision such as their adaptation to T2S. For example, most CSDs are either user-owned and thus largely transparent to their users or publicly listed companies with certain transparency requirements. Furthermore, the CSD adaptation to T2S may require some adaptation on the side of the CSDs' clients, so CSDs usually involve their clients in major changes to their IT infrastructure. Hence, their decision to reshape or not is likely to be known to the market, possibly after some finite lag N_0.

5 Analysing the question to join or not with the model

In the previous section we have been concerned exclusively with the modelling of the reshaping decision for the CSDs which had chosen to join in the first place. Yet, what about also modelling the decision of joining or not joining T2S?

We can adjust the previous model to address this question by changing the cost per transaction $\widetilde{c}_i$: instead of $\widetilde{c}_i = (1-b_i)c_i + c_{T2S}$ we change the interpretation of parameter $b_i \in [0,1]$ and set $\widetilde{c}_i = (1-b_i)c_i + b_i c_{T2S}$. Hence, $b_i = 0$ corresponds to the decision not to join (and thus not to reshape), yielding costs of $\widetilde{c}_i = c_i$. Any $b_i > 0$ corresponds to the decision to join T2S and to reshape by some degree b_i.The extreme case, $b_i = 1$, represents the decision to join and to *completely* reshape, yielding $\widetilde{c}_i = c_{T2S}$ as before. Hence, the table which gives the relevant cost structures for the price-setting game of each stage changes to:

period	fixed costs $\widetilde{C}_{i,fixed}$ of CSD i	cost per transaction $\widetilde{c}_i$ of CSD i
1	$(1-a_i)C_{i,fixed} + C_{i,adapt}(a_i,b_i)$	$(1-b_i)c_i + b_i c_{T2S}$
2	$(1-a_i)C_{i,fixed}$	$(1-b_i)c_i + b_i c_{T2S}$
...	...	...
N	$(1-a_i)C_{i,fixed}$	$(1-b_i)c_i + b_i c_{T2S}$

That is, the more b_i-reshaping, the more the costs per transaction converge towards the T2S fees. This is because $(1-b_i)c_i + b_i c_{T2S}$ is nothing else that a convex combination of the two types of costs, the CSD i costs c_i and the T2S costs c_{T2S}.

Now, in case $c_i < c_{T2S}$, we obtain the fact that once a CSD has chosen to join T2S, the more it reshapes, the more its costs per transaction increase. But this has no consequences for the equilibrium solutions, since if a CSD has (average) costs per transaction below the T2S fees, then in our model it will choose $b_i = 0$ and thus not join T2S.[24] Abstracting from this corner solution, we will in everything that concerns the modelling of joining or not joining always assume that

$$c_i > c_{T2S}.$$

Note that we can apply our previous model and obtain similar results by just substituting every occurrence of c_i by $c_i - c_{T2S}$ (and, of course, c_j by $c_j - c_{T2S}$) since $(1-b_i)c_i + b_i c_{T2S} = (1-b_i)(c_i - c_{T2S}) + c_{T2S}$. In particular, we get theorems similar to Theorem 2 and 3 of Section 3, but concerning the decision to join or not to join T2S:

Theorem 6 *Assume $\frac{1}{f_j} < \frac{c_i - c_{T2S}}{c_j - c_{T2S}} < f_i$. Then there exists $\overline{\delta} \in]0,1[$ such that for any discount factor $\delta \in]\overline{\delta}, 1[$, collusion for not joining T2S can be sustained in a subgame perfect Nash-equilibrium, even if price competition is assumed.*

That is, there exists a subgame-perfect Nash equilibrium in which no CSD ever join T2S when playing it. The corresponding subgame-perfect Nash equilibrium is of course the following trigger strategy (S'):

(S') *Do not join if no player has joined so far. If another player has joined and yourself has not joined, then join. At each price-setting game (step 2) play the unique Nash equilibrium, given all the costs involved.*

Theorem 7 *Assume $c_i \geq f_i(c_j - c_{T2S}) + c_{T2S}$. Then for any $\delta \geq 1 - \frac{c_i^2 D_i^2}{\xi_i}$, CSD i will always join T2S in any subgame perfect Nash equilibrium consistent with price competition. In particular, the other CSD j cannot deter CSD i from joining T2S in any credible way.*

Hence, translating a parameter of the reshaping model allows to study the decision to join or not. The model focuses on the possible transaction cost savings with T2S. There are other benefits of T2S for the CSDs and their clients that are outside the scope of this model (see, e.g., Eurosystem [6]). They can make participation in T2S beneficial independent of the direct settlement costs.

6 Conclusion

In this article we used a game theoretic approach to model the strategic decision of European CSDs to reshape towards T2S, in both a finite and in an infinite setting, and noted that a similar model and discussion can also apply for the CSD's decision to join or not to join T2S.

In the finite setting, we obtained an explicit formula for the degree of optimal reshaping as well as for the optimal prices set at each period. Trying to recover the fixed adaptation costs by increasing the price per transaction results in decreasing profits for CSDs compared with the optimal solution.

In the infinite setting we provided a condition for the model parameters under which two CSDs, which would normally reshape in the previous setting, delay their reshaping infinitely many times because they are hence able to avoid paying adaptation costs while the incentives to reshape (cost-reduction and earning market shares) are reduced by this tacit collusion. This result is not robust though, in particular with the possibility of new entrants on the market who may push average costs lower and force the pre-existing CSDs to reshape. Indeed, we provide a condition under which a given CSD will always reshape, no matter whether the other CSDs are reshaping or not. Decreasing average costs for the whole market push towards

[24] When the individual CSD's costs are below the T2S fees, the CSD will not join T2S in this second model because joining and reshaping could only result in higher costs. Since there may be non cost-based incentives for a given, cost-efficient CSD to join, the second model does not capture properly the decision about reshaping for all CSDs. Hence, we used the first model, which starts from the subset of CSD having joined T2S (no matter the reasons), as a more appropriate way to express the degree of optimal reshaping.

this condition, and away from tacit collusion. Furthermore, this result is robust with respect to a delayed observability of the reshaping decision.

In particular the results in the infinite horizon model provide interesting insights about an optimal, non-reversible investment decision. These results are applicable in a wider Industrial Organisation context and not limited to the reshaping decision of CSDs.

References

[1] M.-N. Cales, D. Chabert, W. Hichri, N. Marchand (2011), *The reform of the European securities settlement systems: towards an integrated financial market*, Groupe dánalyse et de théorie économique de Lyon-St Etienne (GATE) Working Paper No. 1129.

[2] P. Van Cayseele, C. Wuyts (2007), *Cost efficiency in the European securities settlement and depository industry*, Journal of Banking & Finance 31, no 10, 3058-3079.

[3] Cherbonnier, F., J.-C. Rochet (2010) *Vertical Integration and Regulation in the Securities Settlement Industry*, IDEI Working Paper, n. 663.

[4] Clearstream (2012), *Clearstream publishes pricing model for TARGET2-Securities*, 20 December 2012. `http://www.clearstream.com/ci/dispatch/en/listcontent/ci_nav/news/30_Press/44_2012/Content_Files/2012/press_121220.htm`.

[5] ECB (2007), *T2S Economic Feasibility Study*, `https://www.ecb.int/press/pr/date/2007/html/pr070308_2.en.html`.

[6] ECB (2008), *T2S Economic Impact Assessment*, `https://www.ecb.int/paym/t2s/about/keydocs/html/index.en.html`.

[7] ECB (2010), *The Payment System. Payments, Securities and Derivatives, and the Role of the Eurosystem.* Editor: Tom Kokkola. `www.ecb.int/pub/pdf/other/paymentsystem201009en.pdf`.

[8] N. Economides (1996), *The economics of networks*, International Journal of Industrial Organization, Elsevier, 14(6), 673-699.

[9] J. Eliashberg, R. Steinberg (1991), *Competitive strategies for two firms with asymmetric Production cost structures*, Management Science 37, no. 11, 1452-1473

[10] Euroclear (2012), *Réactualisation des tarifs*, 12 December 2012. `https://www.euroclear.com/site/publishedFile/2012-12-06_ESES_NL_T2S_recharge_tcm91-266540.pdf`.

[11] C. H. Fine, R. M. Freund (1990), *Optimal investment in product-flexible manufacturing capacity,* Management Science 36, no. 4, 449-466.

[12] J. Friedman (1971), *A non-cooperative equilibrium for supergames*, Review of Economic Studies 38, no. 1, 1–12.

[13] A. Giovannini (2001), *Cross-border clearing and settlements arrangement in the European Union*, `http://ec.europa.eu/internal\_market/financial-markets/docs/clearing/first\_giovannini\_report\_en.pdf`.

[14] C. Holthausen, J. Tapking (2007), *Raising rival's costs in the securities settlement industry*, Journal of Financial Intermediation, 16, 91-116.

[15] K. Kauko (2007), *Interlinking securities settlement systems: A strategic commitment?* Journal of Banking and Finance 31, 2962-2977.

[16] G. Knieps (2006), *Competition in the post-trade markets: A network economic analysis of the securities business*, Journal of Industry, Competition and Trade, vol 6, no 1, 45-60.

[17] T. V. Koeppl, C. Monnet (2007), *Guess what: It's the settlements! Vertical integration as a barrier to efficient exchange consolidation*, Journal of Banking & Finance, 31, 3013-3033.

[18] A. Mas-Colell, M. D. Whinston, J. R. Green (1995), *Microeconomic Theory*, Oxford University Press.

[19] E. Maskin, J.Tirole (1988), *A Theory of Dynamic Oligopoly, II: Price Competition, Kinked Demand Curves, and Edgeworth Cycles*, Econometrica 56, no. 3, 571-599

[20] C. Matutes, A. J. Padilla (1994), *Shared ATM networks and banking competition*, European Economic Review 38, 1113-11138.

[21] A. Milne (2007), *The industrial organization of post-trade clearing and settlement*, Journal of Banking and Finance 31, 2945-2961.

[22] C. Pirrong (2007), *The Industrial Organization of Execution, Clearing and Settlement in Financial Markets*, Working Paper, `http://www.cba.uh.edu/spirrong/`.

[23] J. C. Rochet (2006), *The Welfare Effects of Vertical Integration in the Securities Clearing and Settlement Industry*, Mimeo, IDEA, Toulouse University.

[24] H. Schmiedel, M. Malkamaki, J. Tarkka (2006). *Economics of Sale and Technological Developments in Securities Depository and Settlement Systems* Journal of Banking & Finance 30, 1783-1806.

[25] M. Spence (1984), *Cost reduction, competition, and industry performance*, Econometrica 52, no 1, 101-122.

[26] J. Tapking, J. Yang (2006) *Horizontal and vertical integration in securities trading and settlement*, Journal of Money, Credit and Banking 38, no 7, 1765-1795.

[27] J. Tirole (1988) *The theory of industrial organization*, The MIT Press, Cambridge.

[28] T. M. Valletti, C. Cambini (2005), *Investments and network competition*, The RAND Journal of Economics 36, no 2, 446-467.

[29] J. Vickers (1995), *Competition and Regulation in Vertically Related Markets*, Review of Economic Studies, 62, 1-17.

[30] I. Vogelsang (2003), *Price regulation of access to telecommunications networks*, Journal of Economic Literature 41, 830-862.

[31] S. D. Wu (2005), *Managing capacity in the high-tech industry: a review of literature*, The Engineering Economist 50, 125-158.

7 Annex

7.1 Computations for the dynamic N-period model for the degree of optimal reshaping

7.1.1 Expression of the prices at equilibrium

Profits for CSD i in a period where the fixed costs are $\widetilde{C}_{i,fixed}$ and the costs per transaction are $\widetilde{c_i}$ are:

$$\begin{aligned}\pi_i &= q_i(p_i - \widetilde{c_i}) - \widetilde{C}_{i,fixed} \\ &= (\alpha_i - \gamma_{ii}p_i + \gamma_{ij}p_j)(p_i - \widetilde{c_i}) - \widetilde{C}_{i,fixed}\end{aligned}$$

Now

$$\begin{aligned}\frac{\partial \pi_i}{\partial p_i} &= (\alpha_i - \gamma_{ii}p_i + \gamma_{ij}p_j) + (-\gamma_{ii})(p_i - \widetilde{c_i}) \\ &= -2\gamma_{ii}p_i + \alpha_i + \gamma_{ij}p_j + \gamma_{ii}\widetilde{c_i}\end{aligned}$$

Hence $\frac{\partial \pi_i}{\partial p_i} > 0$ if, and only if, $p_i < \frac{1}{2\gamma_{ii}}(\alpha_i + \gamma_{ij}p_j + \gamma_{ii}\widetilde{c_i})$. This proves that the best-response $p_i^*(p_j)$ of CSD i to a price p_j from CSD j is:

$$p_i^*(p_j) = \frac{1}{2\gamma_{ii}}(\alpha_i + \gamma_{ii}\widetilde{c_i} + \gamma_{ij}p_j)$$

Denoting by (p_i^*, p_i^*) the equilibrium prices of the stage-game G, we have:

$$\begin{cases} p_i^* = p_i^*(p_j) = \frac{1}{2\gamma_{ii}}(\alpha_i + \gamma_{ii}\widetilde{c_i} + \gamma_{ij}p_j) \\ p_j^* = p_j^*(p_i) = \frac{1}{2\gamma_{jj}}(\alpha_j + \gamma_{jj}\widetilde{c_j} + \gamma_{ji}p_i) \end{cases}$$

Hence

$$p_i^* = \frac{1}{2\gamma_{ii}}(\alpha_i + \gamma_{ii}\widetilde{c_i} + \gamma_{ij}(\frac{1}{2\gamma_{jj}}(\alpha_j + \gamma_{jj}\widetilde{c_j} + \gamma_{ji}p_i)))$$

$$4\gamma_{ii}\gamma_{jj}p_i^* = 2\gamma_{jj}\alpha_i + 2\gamma_{jj}\gamma_{ii}\widetilde{c_i} + \gamma_{ij}(\alpha_j + \gamma_{jj}\widetilde{c_j} + \gamma_{ji}p_j)$$

$$(4\gamma_{ii}\gamma_{jj} - \gamma_{ij}\gamma_{ji})p_i^* = 2\gamma_{jj}\alpha_i + 2\gamma_{jj}\gamma_{ii}\widetilde{c_i} + \gamma_{ij}\alpha_j + \gamma_{ij}\gamma_{jj}\widetilde{c_j}$$

$$p_i^* = \frac{1}{4\gamma_{ii}\gamma_{jj} - \gamma_{ij}\gamma_{ji}}(2\gamma_{jj}\alpha_i + 2\gamma_{ii}\gamma_{jj}\widetilde{c_i} + \gamma_{ij}\alpha_j + \gamma_{ij}\gamma_{jj}\widetilde{c_j})$$

Similarly,

$$p_j^* = \frac{1}{4\gamma_{ii}\gamma_{jj} - \gamma_{ij}\gamma_{ji}}(2\gamma_{ii}\alpha_j + 2\gamma_{ii}\gamma_{jj}\widetilde{c_j} + \gamma_{ji}\alpha_i + \gamma_{ji}\gamma_{ii}\widetilde{c_i})$$

These formulas define the equilibrium prices only when they are superior to the costs involved. Indeed, if $p_i^* < \widetilde{c_i}$ then CSD i optimal response is not to settle any transactions. We will always assume, in what follows, that this is not the case, i.e. that it is always more profitable for a CSD to settle a transaction than to refuse to settle it. This allows to get rid of (unrealistic) corner solutions. Hence, we will always assume $p_i^* \geq \widetilde{c_i}$ and $p_j^* \geq \widetilde{c_j}$ for any costs involved where $p_i^* = p_i^*(\widetilde{c_i}, \widetilde{c_j})$ and $p_j^* = p_j^*(\widetilde{c_i}, \widetilde{c_j})$ are the expressions, dependent on the cost level $(\widetilde{c_i}, \widetilde{c_j})$, given above. This assumption is thus a joint set of assumptions on the parameter of the model and will be used in other part of the article.

7.1.2 Expression of the profits at equilibrium

Replacing in the expression of the profits p_i by $p_i^*(p_j)$ gives:

$$\begin{aligned}
\pi_i(p_i^*(p_j), p_j) &= (\alpha_i - \gamma_{ii}p_i + \gamma_{ij}p_j)(p_i - \widetilde{c_i}) - \widetilde{C}_{i,fixed} \\
&= (\alpha_i - \frac{1}{2}(\alpha_i + \gamma_{ii}\widetilde{c_i} + \gamma_{ij}p_j) + \gamma_{ij}p_j)(\frac{1}{2\gamma_{ii}}(\alpha_i + \gamma_{ii}\widetilde{c_i} + \gamma_{ij}p_j) - \widetilde{c_i}) - \widetilde{C}_{i,fixed} \\
&= \frac{1}{\gamma_{ii}}\left(\frac{1}{2}(\alpha_i - \gamma_{ii}\widetilde{c_i} + \gamma_{ij}p_j)\right)^2 - \widetilde{C}_{i,fixed}
\end{aligned}$$

Then replacing p_j by its expression at equilibrium yields:

$$\begin{aligned}
\pi_i &= \frac{1}{4\gamma_{ii}}\left(\alpha_i - \gamma_{ii}\widetilde{c_i} + \gamma_{ij}p_j\right)^2 - \widetilde{C}_{i,fixed} \\
&= \frac{1}{4\gamma_{ii}}\left(\alpha_i - \gamma_{ii}\widetilde{c_i} + \frac{\gamma_{ij}}{4\gamma_{ii}\gamma_{jj} - \gamma_{ij}\gamma_{ji}}(2\gamma_{ii}\alpha_j + 2\gamma_{ii}\gamma_{jj}\widetilde{c_j} + \gamma_{ji}\alpha_i + \gamma_{ji}\gamma_{ii}\widetilde{c_i})\right)^2 - \widetilde{C}_{i,fixed} \\
&= \frac{1}{4\gamma_{ii}}(\alpha_i + \frac{\gamma_{ij}}{4\gamma_{ii}\gamma_{jj} - \gamma_{ij}\gamma_{ji}}(2\gamma_{ii}\alpha_j + \gamma_{ji}\alpha_i) \\
&\quad +(\frac{2\gamma_{ii}\gamma_{ij}\gamma_{jj}}{4\gamma_{ii}\gamma_{jj} - \gamma_{ij}\gamma_{ji}} - \gamma_{ii})\widetilde{c_j} + \frac{\gamma_{ij}\gamma_{ji}\gamma_{ii}}{4\gamma_{ii}\gamma_{jj} - \gamma_{ij}\gamma_{ji}}\widetilde{c_i})^2 - \widetilde{C}_{i,fixed} \\
&= (\frac{1}{2\sqrt{\gamma_{ii}}}(\alpha_i + \frac{\gamma_{ij}}{4\gamma_{ii}\gamma_{jj} - \gamma_{ij}\gamma_{ji}}(2\gamma_{ii}\alpha_j + \gamma_{ji}\alpha_i)) \\
&\quad + \frac{\sqrt{\gamma_{ii}}\gamma_{ij}\gamma_{jj}}{4\gamma_{ii}\gamma_{jj} - \gamma_{ij}\gamma_{ji}}\widetilde{c_j} + \frac{1}{2\sqrt{\gamma_{ii}}}(\frac{\gamma_{ii}\gamma_{ij}\gamma_{ji}}{4\gamma_{ii}\gamma_{jj} - \gamma_{ij}\gamma_{ji}} - \gamma_{ii})\widetilde{c_i})^2 - \widetilde{C}_{i,fixed}
\end{aligned}$$

Now

$$\begin{aligned}
\frac{1}{2\sqrt{\gamma_{ii}}}(\frac{\gamma_{ii}\gamma_{ij}\gamma_{ji}}{4\gamma_{ii}\gamma_{jj} - \gamma_{ij}\gamma_{ji}} - \gamma_{ii}) &= \frac{\sqrt{\gamma_{ii}}}{2}(\frac{\gamma_{ij}\gamma_{ji} - 4\gamma_{ii}\gamma_{jj} + \gamma_{ij}\gamma_{ji}}{4\gamma_{ii}\gamma_{jj} - \gamma_{ij}\gamma_{ji}}) \\
&= \sqrt{\gamma_{ii}}(\frac{\gamma_{ij}\gamma_{ji} - 2\gamma_{ii}\gamma_{jj}}{4\gamma_{ii}\gamma_{jj} - \gamma_{ij}\gamma_{ji}}) \\
&= -\sqrt{\gamma_{ii}}(\frac{2\gamma_{ii}\gamma_{jj} - \gamma_{ij}\gamma_{ji}}{4\gamma_{ii}\gamma_{jj} - \gamma_{ij}\gamma_{ji}})
\end{aligned}$$

Let

$$\begin{aligned}
A_i &= \frac{1}{2\sqrt{\gamma_{ii}}}(\alpha_i + \frac{\gamma_{ij}}{4\gamma_{ii}\gamma_{jj} - \gamma_{ij}\gamma_{ji}}(2\gamma_{ii}\alpha_j + \gamma_{ji}\alpha_i)) \\
B_i &= \frac{\sqrt{\gamma_{ii}}\gamma_{ij}\gamma_{jj}}{4\gamma_{ii}\gamma_{jj} - \gamma_{ij}\gamma_{ji}} \\
D_i &= \sqrt{\gamma_{ii}}\frac{2\gamma_{ii}\gamma_{jj} - \gamma_{ij}\gamma_{ji}}{4\gamma_{ii}\gamma_{jj} - \gamma_{ij}\gamma_{ji}}
\end{aligned}$$

Now profits can be written in a very compact way as:

$$\pi_i = (A_i + B_i\widetilde{c_j} - D_i\widetilde{c_i})^2 - \widetilde{C}_{i,fixed}$$

Note in passing that $A_i + B_i\widetilde{c_j} - D_i\widetilde{c_i}$ is positive for any costs $(\widetilde{c_i}, \widetilde{c_j})$ involved, because by the above $A_i + B_i\widetilde{c_j} - D_i\widetilde{c_i} = p_i - \widetilde{c_i} = \frac{1}{\gamma_{ii}}q_i \geq 0$ by assumption (CSA). This fact will be used later.

Total profits, as defined in Section 2.4.3, are thus:

$$\begin{aligned}
\pi_i^{tot} &= (1 + \delta + ... + \delta^{N-1})\{(A_i + ((1-b_j)c_j + c_{T2S})B_i - ((1-b_i)c_i + c_{T2S})D_i)^2 - (1-a_i)C_{i,fixed}\} \\
&\quad - C_{i,adapt}(a_i, b_i) \\
&= \widetilde{\delta}(A_i + ((1-b_j)c_j + c_{T2S})B_i - ((1-b_i)c_i + c_{T2S})D_i)^2 - \widetilde{\delta}(1-a_i)C_{i,fixed} - C_{i,adapt}(a_i, b_i)
\end{aligned}$$

with

$$\widetilde{\delta} = \begin{cases} N & \text{if } \delta = 1 \\ \frac{1-\delta^N}{1-\delta} & \text{if } 0 \leq \delta < 1 \end{cases}$$

7.1.3 Expression of the degree of reshaping chosen and Proof of Lemma 1 and Theorem 1

Now assuming in the previous expression of π_i^{tot} that $C_{i,adapt}(a_i, b_i) = \xi_i b_i^2$ and $C_{i,fixed} = 0$ for simplicity, we can compute the derivative of π_i^{tot} with respect to the degree of reshaping b_i:

$$\frac{\partial \pi_i^{tot}}{\partial b_i} = 2\widetilde{\delta} c_i D_i (A_i + ((1-b_j)c_j + c_{T2S})B_i - ((1-b_i)c_i + c_{T2S})D_i) - 2\xi_i b_i$$

Hence $\frac{\partial \pi_i^{tot}}{\partial b_i} > 0$ is equivalent to:

$$c_i D_i (A_i + ((1-b_j)c_j + c_{T2S})B_i - (c_i + c_{T2S}))D_i) > \frac{\xi_i - \widetilde{\delta} c_i^2 D_i^2}{\widetilde{\delta}} b_i$$

We can now prove Lemma 1, which is re-state here for convenience for the reader:

Lemma 1 *Assume $\xi_i < \widetilde{\delta} c_i^2 D_i^2$. Then if all CSDs are playing the Nash equilibrium of each following stage-game of the repeated game, the best-response function $b_i^*(b_j)$ is the constant function $b_i^*(b_j) = 1$, which represents a throughout reshaping decision from CSD i whatever the degree b_j CSD j choose to reshape.*

Proof of Lemma 1:

If $\xi_i < \widetilde{\delta} c_i^2 D_i^2$ then the sign of $\frac{\partial \pi_i^{tot}}{\partial b_i}$ indicate that π_i^{tot} is first decreasing, then increasing, hence its maximum is reached either at $b_i = 0$ or at $b_i = 1$ (and possibly at both). But

$$\pi_i^{tot}(0, b_j) < \pi_i^{tot}(1, b_j)$$

is successively equivalent to

$$\begin{aligned} \widetilde{\delta}(A_i + ((1-b_j)c_j + c_{T2S})B_i - (c_i + c_{T2S})D_i)^2 &< \widetilde{\delta}(A_i + ((1-b_j)c_j + c_{T2S})B_i - c_{T2S}D_i)^2 - \xi_i \\ \widetilde{\delta}(A_i + ((1-b_j)c_j + c_{T2S})B_i - (c_i + c_{T2S})D_i)^2 + \xi_i &< \widetilde{\delta}(A_i + ((1-b_j)c_j + c_{T2S})B_i - c_{T2S}D_i)^2 \end{aligned}$$

But the last line is true since we can write:

$$\begin{aligned} & \widetilde{\delta}(A_i + ((1-b_j)c_j + c_{T2S})B_i - (c_i + c_{T2S})D_i)^2 + \xi_i \\ \leq \; & \widetilde{\delta}(A_i + ((1-b_j)c_j + c_{T2S})B_i - (c_i + c_{T2S})D_i)^2 + \widetilde{\delta} c_i^2 D_i^2) \\ \leq \; & \widetilde{\delta}\{(A_i + ((1-b_j)c_j + c_{T2S})B_i - (c_i + c_{T2S})D_i)^2 + (c_i D_i)^2\} \\ \leq \; & \widetilde{\delta}\{(A_i + ((1-b_j)c_j + c_{T2S})B_i - (c_i + c_{T2S})D_i + (c_i D_i))^2 \\ = \; & \widetilde{\delta}\{(A_i + ((1-b_j)c_j + c_{T2S})B_i - c_{T2S}D_i)^2 \end{aligned}$$

The last inequality above was obtained using $a^2 + b^2 \leq (a+b)^2$ for any a, b such that $ab \geq 0$. Here $a = A_i + ((1-b_j)c_j + c_{T2S})B_i - (c_i + c_{T2S})D_i \geq 0$ (since this is equal to a positive factor times the quantity at equilibrium of the stage game) and $b = c_i D_i \geq 0$. Note that this result is independent of the level of reshaping b_j chosen by the other CSD j. This concludes the proof of Lemma 1. □

Assume now that $\xi_i > \widetilde{\delta} c_i^2 D_i^2$. Then the maximum $b_i^{**}(b_j)$ of $\pi_i^{tot}(b_j)$ where b_i is not restricted to $[0, 1]$ is:

$$b_i^{**}(b_j) = \frac{\widetilde{\delta} c_i D_i}{\xi_i - \widetilde{\delta} c_i^2 D_i^2}(A_i + ((1-b_j)c_j + c_{T2S})B_i - (c_i + c_{T2S}))D_i)$$

Note $b_i^{**}(b_j)$ can be greater than 1 or smaller than 0. Letting $\psi_i = \frac{\widetilde{\delta} c_i D_i}{\xi_i - \widetilde{\delta} c_i^2 D_i^2}$, we have

$$\begin{aligned} b_i^{**}(b_j) &= \psi_i(A_i + ((1-b_j)c_j + c_{T2S})B_i - (c_i + c_{T2S}))D_i) \\ &= \psi_i(A_i + (c_j + c_{T2S})B_i - (c_i + c_{T2S})D_i) - \psi_i c_j B_i b_j \end{aligned}$$

Because $\xi_i > \widetilde{\delta} c_i^2 D_i^2$, π_i^{tot} is first increasing, then decreasing, and reaches its (unrestricted) maximum in $b_i^{**}(b_j)$. It is then easy to see that the best-response function $b_i^*(b_j)$ is 1 if $b_i^{**}(b_j) > 1$ and 0 if $b_i^{**}(b_j) < 0$, whereas it is $b_i^{**}(b_j)$ when $b_i^{**}(b_j) \in [0,1]$. This can be summarized by writing:

$$b_i^*(b_j) = \min(1, \max(0, b_i^{**}(b_j)))$$

that is,

$$b_i^*(b_j) = \min(1, \max(0, \psi_i(A_i + (c_j + c_{T2S})B_i - (c_i + c_{T2S})D_i) - \psi_i c_j B_i b_j))$$

We can now proceed to prove Theorem 1.

Proof of Theorem 1:

We will use Lemma 1 extensively to prove part (ii), (iii) and (iv) of the theorem.

(ii) Assume $\xi_i < \widetilde{\delta} c_i^2 D_i^2$ and $\xi_j < \widetilde{\delta} c_j^2 D_j^2$. Then by Lemma 1 applied to both CSD i and j we have $b_i^*(b_j) = 1$ for any b_j and $b_j^*(b_i) = 1$ for any b_i. Hence at equilibrium $b_i^* = 1 = b_j^*$.

(iii) Assume $\xi_i > \widetilde{\delta} c_i^2 D_i^2$ and $\xi_j < \widetilde{\delta} c_j^2 D_j^2$. Lemma 1 applied to CSD j gives $b_j^*(b_i) = 1$ for any b_i, and thus at equilibrium CSD i plays its best-response (as previously computed above) $b_i^*(b_j) = \min(1, \max(0, \psi_i(A_i + (c_j + c_{T2S})B_i - (c_i + c_{T2S}))D_i)))$.

(iv) is similar to (iii) interchanging the roles of i and j. Indeed, if we assume $\xi_i < \widetilde{\delta} c_i^2 D_i^2$ and $\xi_j > \widetilde{\delta} c_j^2 D_j^2$. Lemma 1 applied to CSD i gives $b_i^*(b_j) = 1$ for any b_j, and thus at equilibrium $b_j^*(b_i) = \min(1, \max(0, \psi_i(A_i + (c_j + c_{T2S})B_i - (c_i + c_{T2S}))D_i)))$.

We are thus only left with the task of proving point (i) of Theorem 1.

Assume $\xi_i > \widetilde{\delta} c_i^2 D_i^2$ and $\xi_j > \widetilde{\delta} c_j^2 D_j^2$. Consider the solution (b_i^{**}, b_j^{**}) obtained from solving:

$$\begin{cases} b_i^{**} = b_i^{**}(b_j^{**}) \\ b_j^{**} = b_j^{**}(b_i^{**}) \end{cases}$$

That is, we solve for a Nash equilibrium the unrestricted game. The system is equivalent to

$$\begin{cases} b_i^{**} = \psi_i(A_i + (c_j + c_{T2S})B_i - (c_i + c_{T2S})D_i) - \psi_i c_j B_i b_j^{**} \\ b_j^{**} = \psi_j(A_j + (c_i + c_{T2S})B_j - (c_j + c_{T2S})D_j) - \psi_j c_i B_j b_i^{**} \end{cases}$$

Substituting b_j^* by its expression in function of b_i^* (given by the second equation) in the first equation of the system gives:

$$\begin{aligned} b_i^{**} &= \psi_i(A_i + (c_j + c_{T2S})B_i - (c_i + c_{T2S})D_i) - \psi_i c_j B_i(\psi_j(A_j + (c_i + c_{T2S})B_j - (c_j + c_{T2S})D_j) - \psi_j c_i B_j b_i) \\ b_i^{**} &= \frac{\psi_i(A_i + (c_j + c_{T2S})B_i - (c_i + c_{T2S})D_i) - \psi_i \psi_j c_j B_i(A_j + (c_i + c_{T2S})B_j - (c_j + c_{T2S})D_j)}{1 - \psi_i \psi_j c_i c_j B_i B_j} \end{aligned}$$

This is the optimal, unrestricted, degree b_i^{**} of reshaping chosen by CSD i at equilibrium, assuming $\psi_i \psi_j c_i c_j B_i B_j \neq 1$[25]. A similar formula, interchanging the role of i and j, holds for b_j^{**}.

Now, whenever b_i^{**} and b_j^{**} both belong to $[0,1]$, the solution (b_i^*, b_j^*) of the *restricted* game, that is, where b_i^* and b_j^* are constraint to belong to $[0,1]$, is just the solution of the unrestricted game. That is,

[25] There is, of course, a graphical interpretation to the condition $1 \neq \psi_i \psi_j c_i c_j B_i B_j$. Indeed, if $1 = \psi_i \psi_j c_i c_j B_i B_j$ then the two best-response functions $b_i^{**}(b_j)$ and $b_j^{**}(b_i)$ are parallel when drawn in the (b_i, b_j)-plane, hence they either intercept in all their points or in none. A discussion on what happens in this precise case is carried out in Appendix 7.1.4. For the rest of the article we always assume $1 \neq \psi_i \psi_j c_i c_j B_i B_j$.

$(b_i^*, b_j^*) = (b_i^{**}, b_j^{**})$. This intuitive result can be proved more formally: by definition, (b_i^{**}, b_j^{**}) solves the system

$$\begin{cases} b_i^{**} = b_i^{**}(b_j^{**}) \\ b_j^{**} = b_j^{**}(b_i^{**}) \end{cases}$$

Now, because $b_i^{**}(b_j^{**}) = b_i^{**}$ belongs to $[0,1]$, $b_i^*(b_j^{**}) = \min(1, \max(0, b_i^{**}(b_j^{**}))) = b_i^{**}(b_j^{**}) = b_i^{**}$. Similarly, $b_j^* = b_j^{**}$.

This proves Theorem 1 as such. We know move to provide a complete resolution (including corner-solutions) of the Nash equilibrium of game.

We have proved previously that

$$b_i^{**}(b_j) = \psi_i(A_i + (c_j + c_{T2S})B_i - (c_i + c_{T2S})D_i) - \psi_i c_j B_i b_j$$

For convenience, let

$$\alpha = \psi_i(A_i + (c_j + c_{T2S})B_i - (c_i + c_{T2S})D_i)$$

and

$$\beta = \psi_i c_j B_i$$

Hence

$$b_i^{**}(b_j) = \alpha - \beta b_j$$

Let us look at the signs of the quantities involved. Since $\xi_i > \widetilde{\delta} c_i^2 D_i^2$, we have that $\psi_i = \frac{\widetilde{\delta} c_i D_i}{\xi_i - \widetilde{\delta} c_i^2 D_i^2} > 0$, and by the CSA applied with the costs $(c_i + c_{T2S}, c_j + c_{T2S})$, we have $A_i + (c_j + c_{T2S})B_i - (c_i + c_{T2S})D_i > 0$. Hence, $\alpha > 0$. Similarly, because $\psi_i > 0$ and that the CSA applied with the costs $(c_i + c_{T2S}, c_{T2S})$ gives $A_i + c_{T2S}B_i - (c_i + c_{T2S}) > 0$, we get that $b_i^{**}(1) = \alpha - \beta > 0$.

Similarly, define $\widetilde{\alpha}$ and $\widetilde{\beta}$ such that

$$b_j^{**}(b_i) = \widetilde{\alpha} - \widetilde{\beta} b_i$$

for all b_i. We also have, through the CSA, that $\widetilde{\alpha} > 0$ and $\widetilde{\alpha} - \widetilde{\beta} > 0$.

We can now proceed to a disjunction aimed at capturing all the different types of corner solutions that could occur.

First assume $\alpha < 1$ and $\widetilde{\alpha} < 1$. In that case $b_i^*(b_j) = b_i^{**}(b_j)$ for all $b_j \in [0,1]$ and similarly $b_j^*(b_i) = b_j^{**}(b_i)$. Hence both best-response functions are straight-lines on their whole domain of definition (which is $[0,1]$), and since $b_j^*(b_i)$ is below $b_i^*(b_j)$ at $b_i = \alpha - \beta$ while it is above at $b_i = \alpha$, the two segments intersect in some point in the interior of $[0,1]^2$. Because the unrestricted best-responses $b_i^{**}(b_j)$, $b_j^{**}(b_i)$ coincide with $b_i^*(b_j)$, $b_j^*(b_i)$ this point is precisely (b_i^{**}, b_j^{**}). This proves that the Nash equilibrium, in that case, is unique, and is (b_i^{**}, b_j^{**}).

Assume $\alpha < 1$ and $\widetilde{\alpha} > 1$. Consider the intersection of the two best-response functions with the straight line D of equation $b_i = \alpha - \beta$. Because $b_i^*(1) = \alpha - \beta$, D intersects $b_i^*(b_j)$ at the point of coordinate $(\alpha - \beta, 1)$. Assume $b_j^{**}(\alpha - \beta) \geq 1$. Then $b_i^{**}(b_j)$ and $b_j^{**}(b_i)$ do not intersect in the interior of $[0,1]^2$ and $b_j^*(\alpha - \beta) = 1$, implying that the unique intersection point of the two best-response function is precisely $(\alpha - \beta, 1)$, that is, $(b_i^{**}(1), 1)$. Hence, $(b_i^*, b_j^*) = (b_i^{**}(1), 1)$ is the unique NE. Assume now $b_j^{**}(\alpha - \beta) < 1$. Then $b_j^*(b_i)$ is below $b_i^*(b_j)$ at $b_i = \alpha - \beta$ while it is above at $b_i = \alpha$, and as in the case where $\alpha < 1$ and $\widetilde{\alpha} < 1$ we conclude that (b_i^{**}, b_j^{**}) is the unique NE.

Assume $\alpha > 1$ and $\widetilde{\alpha} < 1$. By the same reasoning as above, interchanging the role of i and j, we get that if $b_i^{**}(\widetilde{\alpha} - \widetilde{\beta}) \geq 1$, then the unique NE is $(b_i^*, b_j^*) = (1, b_j^{**}(1))$ whereas if $b_i^{**}(\widetilde{\alpha} - \widetilde{\beta}) < 1$, it is $(b_i^*, b_j^*) = (b_i^{**}, b_j^{**})$.

The last case is when $\alpha > 1$ and $\widetilde{\alpha} > 1$. To tackle it we need further disjunctions.

First assume $\widetilde{\alpha} - \widetilde{\beta} > 1$ and $\alpha - \beta > 1$. This implies that $b_i^{**}(b_j) > 1$ and $b_j^{**}(b_i) > 1$ for all b_j in $[0,1]$ and all b_i in $[0,1]$. Hence the two best-response functions are $b_i^{**}(b_j) = 1$ and $b_j^{**}(b_i) = 1$ for all b_j in $[0,1]$ and all b_i in $[0,1]$. They intersect in the unique NE, at $(b_i^*, b_j^*) = (1,1)$.

Second assume $\widetilde{\alpha} - \widetilde{\beta} > 1$ and $\alpha - \beta < 1$. This implies that $b_i^{**}(b_j) > 1$ for all b_j in $[0,1]$ and thus $b_i^{**}(b_j) = 1$ for all b_j in $[0,1]$. This best-response thus intersect $b_i^{**}(b_j)$ at a unique NE, which is $(b_i^*, b_j^*) = (b_i^{**}(1), 1)$.

The third case, in which $\widetilde{\alpha} - \widetilde{\beta} < 1$ and $\alpha - \beta > 1$, is similar to the second case, interchanging i and j. It thus yields a unique NE at $(b_i^*, b_j^*) = (1, b_j^{**}(1))$.

The last case is when $\widetilde{\alpha} - \widetilde{\beta} < 1$ and $\alpha - \beta < 1$. Let A be the point of coordinate $(x_A, 1)$ such that $b_j^*(x_A + \varepsilon) < 1$ for any $\varepsilon > 0$ whereas $b_j^*(x_A) = 1$. Similarly, let B be the point of coordinate $(1, y_B)$ the point such that $b_i^*(y_B + \varepsilon) < 1$ for any $\varepsilon > 0$ whereas $b_i^*(y_B) = 1$. Because x_A solves $1 = b_j^{**}(x_A)$ and y_B solves $1 = b_i^{**}(y_B)$, we easily get $x_A = \frac{\widetilde{\alpha}-1}{\widetilde{\beta}}$ and $y_B = \frac{\alpha-1}{\beta}$. It is then convenient to carry out the discussion depending on the relative value of x_A compared to $\alpha - \beta$ and of y_B compared to $\widetilde{\alpha} - \widetilde{\beta}$.

If $\widetilde{\alpha} - \widetilde{\beta} > y_B$ and $\alpha - \beta < x_A$, then there is a unique NE at $(b_i^*, b_j^*) = (b_i^{**}(1), 1)$.

If $\widetilde{\alpha} - \widetilde{\beta} > y_B$ and $\alpha - \beta > x_A$, then there is a unique NE at $(b_i^*, b_j^*) = (b_i^{**}, b_j^{**})$.

If $\widetilde{\alpha} - \widetilde{\beta} < y_B$ and $\alpha - \beta > x_A$, then there is a unique NE at $(b_i^*, b_j^*) = (1, b_j^{**}(1))$.

If $\widetilde{\alpha} - \widetilde{\beta} < y_B$ and $\alpha - \beta < x_A$, then each of the previous points, that is, $(b_i^{**}(1), 1)$, (b_i^{**}, b_j^{**}) and $(1, b_j^{**}(1))$, are NE, and these are the only NE, since the best-response functions intercept in exactly these three points.

This concludes the disjunction. □

7.1.4 Case where $1 = \psi_i \psi_j c_i c_j B_i B_j$

We assume in this section $1 = \psi_i \psi_j c_i c_j B_i B_j$ and re-use the notations of the previous section. The slope of $b_i^{**}(b_j)$ relatively to the plane (b_i, b_j) is thus precisely $\frac{-1}{\beta}$ and the slope of $b_j^{**}(b_i)$ in this same plane is $\widetilde{-\beta}$. Hence the slopes are equal and the (unrestricted) best-response are parallel if, and only if, $\frac{-1}{\beta} = -\widetilde{\beta}$, i.e. if and only if $1 = \beta\widetilde{\beta}$. Because $\beta = \psi_i c_j B_i$ and $\widetilde{\beta} = \psi_j c_i B_j$, this is equivalent to $1 = \psi_i c_j B_i \psi_j c_i B_j$, that is, precisely our condition. Here we discuss what happens for $b_i^*(b_j)$ and $b_j^*(b_i)$ in that case – the rest of the article always making the implicit assumptions that $1 \neq \psi_i c_j B_i \psi_j c_i B_j$.

We can separate three cases.

Assume first that $\widetilde{\alpha} > \frac{\alpha}{\beta}$. This means the best-response straight-line $b_j^{**}(b_i)$ is "above" $b_j^{**}(b_i)$ in the (b_i, b_j) plane. Now if $b_i^*(1) \geq 0$ the best-response $b_i^*(b_j)$ and $b_j^*(b_i)$ intersect in a single NE at $(b_i^*, b_j^*) = (b_i^{**}(1), 1)$. Otherwise, they do not intersect and there is no NE.

Assume that $\widetilde{\alpha} < \frac{\alpha}{\beta}$. This means the best-response straight-line $b_j^{**}(b_i)$ is "below" $b_j^{**}(b_i)$ in the (b_i, b_j) plane. Now if $b_j^*(1) \geq 0$ then the best-response $b_i^*(b_j)$ and $b_j^*(b_i)$ intersect in a single NE at $(b_i^*, b_j^*) = (1, b_j^{**}(1))$. Otherwise, they do not intersect and there is no NE.

Assume $\widetilde{\alpha} = \frac{\alpha}{\beta}$. Then the two best-responses form a single straight-line and any point of $b_j^{**}(b_i)$ in the interior of $[0,1]^2$ is a NE. There is first an infinity of NE in that case.

To conclude, note that in this case too the CSA still implies $\alpha > 0$ and $\alpha - \beta > 0$ as well as $\widetilde{\alpha} > 0$ and $\widetilde{\alpha} - \widetilde{\beta} > 0$, hence there are no other cases to explore than the one previously stated.

7.2 Proofs of Section 3 results and Theorems

7.2.1 Proof of Theorem 2

Assume $\frac{1}{f_j} < \frac{c_i}{c_j} < f_i$ and consider the following trigger strategy (S):

(S) *Do not reshape if no player has reshaped so far. If another player has reshaped and yourself has not reshaped, then reshape. At each price-setting game (step 2) play the unique Nash equilibrium, given the costs.*

To prove this strategy profile yields a subgame-perfect Nash equilibrium, we show it satisfies the definition of a subgame-perfect Nash equilibrium, that is, it induces a Nash equilibrium on every subgame of the whole game. The whole game consists of two types of subgames: those where at least one player has reshaped in one of the previous periods, and those where no player has reshaped so far.

Consider a subgame of the first type. Then the strategy induced by (S) on this subgame consists of reshaping at stage 1) in case it has not been done yet, and playing the Nash equilibrium of the price-setting game for each stage game. As argued in Section 2.4, this constitutes a Nash equilibrium of the subgame. Indeed, since player's actions do not take the past into account anymore and play repeatedly the Nash equilibrium of the price setting game, any player unilaterally deviating from it would not be better of, by definition of a Nash equilibrium of the stage game and of the payoff for the subgame.

Consider a subgame of the second type. Then the strategy profile induced by (S) on this subgame is precisely (S) itself, and in particular it implies not to reshape at stage 1) of the first period of the subgame, since no one has ever reshaped before. Now does it constitutes a Nash equilibrium of the subgame? Consider first the value V of playing this induced strategy for player i, assuming all other players also play it. V is equal to the sum of the immediate payoff $\pi_i(0,0)$ plus the discounted value of playing this strategy on the next periods of the subgame - which happen conveniently to be equal to our whole subgame. Hence we can write

$$V = \pi_i(0,0) + \delta V$$

which yields

$$V = \frac{\pi_i(0,0)}{1-\delta}$$

Now assume player i would unilaterally deviate from this strategy by reshaping to some extent $b_i > 0$. This entails some immediate reshaping costs $\xi_i b_i^2$. Since in the first period of our subgame other players do not reshape, CSD i immediate payoff ends up being $\pi_i(b_i,0) - \xi_i b_i^2$. But in the following periods all CSDs also reshape and CSD i profit is $\pi_i(b_i, b_j^*(b_i))$. Hence the total profits $\pi_i^{dev}(b_i)$ obtained from unilaterally deviating from the induced strategy by playing $b_i > 0$ at step 1) of the first period of the subgame are

$$\begin{aligned}
\pi_i^{dev}(b_i) &= \pi_i(b_i,0) - \xi_i b_i^2 + \delta\pi_i(b_i, b_j^*(b_i)) + \delta^2\pi_i(b_i, b_j^*(b_i)) + ... \\
&= \pi_i(b_i,0) - \xi_i b_i^2 + \frac{\delta}{1-\delta}\pi_i(b_i, b_j^*(b_i))
\end{aligned}$$

which can be written as:

$$\pi_i^{dev}(b_i) = (A_i + (c_j + c_{T2S})B_i - ((1-b_i)c_i + c_{T2S}))D_i)^2 + \frac{\delta}{1-\delta}(A_i + c_{T2S}B_i - ((1-b_i)c_i + c_{T2S}))D_i)^2 - \xi_i b_i^2$$

$$\frac{\partial \pi_i^{dev}}{\partial b_i} = 2c_i D_i(A_i + (c_j + c_{T2S})B_i - ((1-b_i)c_i + c_{T2S}))D_i) + \frac{2\delta}{1-\delta}c_i D_i(A_i + c_{T2S}B_i - ((1-b_i)c_i + c_{T2S}))D_i) - 2\xi_i b_i$$

Rearranging the expression we see that $\frac{\partial \pi_i^{dev}}{\partial b_i} > 0$ if, and only if:

$$(\xi_i - \frac{c_i^2 D_i^2}{1-\delta})b_i < c_i D_i(\frac{1}{1-\delta}A_i + (c_i + \frac{1}{1-\delta}c_{T2S})B_i - \frac{1}{1-\delta}(c_i + c_{T2S})D_i)$$

Now assuming $\delta > 1 - \frac{c_i^2 D_i^2}{\xi_i}$ yields $\xi_i - \frac{c_i^2 D_i^2}{1-\delta} < 0$. Hence the maximum of π_i^{dev} is attained either at $b_i = 0$ or at $b_i = 1$. To decide which value of b_i yields this maximum we just have to compare $\pi_i^{dev}(0)$ with $\pi_i^{dev}(1)$. Assume also $\xi_i < \frac{c_j^2 D_j^2}{1-\delta}$, that is, the similar condition for CSD j; then, because Lemma 1 applies for CSD j (note it does not apply as such for CSD i), a complete reshaping ($b_j = 1$) is always more appropriate for CSD j, whatever the CSD i decision concerning its own degree of reshaping b_i. In particular $b_j^*(0) = b_j^*(1) = 1$, which we will use below.

We will prove that the maximum of π_i^{dev} is attained at $b_i = 1$ using a argument inspired from the proof of Lemma 1 (see Annex 7.1.3). Indeed, $\pi_i^{dev}(0) < \pi_i^{dev}(1)$ is successively equivalent to

$$\begin{aligned}
\pi_i(0,0) + \frac{\delta}{1-\delta}\pi_i(0, b_j^*(0)) &< \pi_i(1,0) + \frac{\delta}{1-\delta}\pi_i(1, b_j^*(1)) - \xi_i \\
\pi_i(0,0) + \frac{\delta}{1-\delta}\pi_i(0,1) &< \pi_i(1,0) + \frac{\delta}{1-\delta}\pi_i(1,1) - \xi_i
\end{aligned}$$

Setting for convenience:

$$\begin{array}{rcl} A & = & A_i + (c_j + c_{T2S})B_i - (c_i + c_{T2S}))D_i \\ B & = & A_i + c_{T2S}B_i - (c_i + c_{T2S}))D_i \\ A' & = & A_i + (c_j + c_{T2S})B_i - c_{T2S})D_i \\ B' & = & A_i + c_{T2S}B_i - c_{T2S}D_i \end{array}$$

this amounts to

$$A^2 + \frac{\delta}{1-\delta}B^2 < A'^2 + \frac{\delta}{1-\delta}B'^2 - \xi_i$$

Since $\xi_i < \frac{c_j^2 D_j^2}{1-\delta}$, this last line is implied by

$$A^2 + \frac{\delta}{1-\delta}B^2 < A'^2 + \frac{\delta}{1-\delta}B'^2 - \frac{c_j^2 D_j^2}{1-\delta}$$

But this later condition is equivalent, given that $\frac{c_j^2 D_j^2}{1-\delta} = (1 + \frac{\delta}{1-\delta})c_j^2 D_j^2$, to

$$A^2 + c_j^2 D_j^2 + \frac{\delta}{1-\delta}(B^2 + c_j^2 D_j^2) < A'^2 + \frac{\delta}{1-\delta}B'^2$$

which is always true because

$$\begin{cases} A^2 + c_j^2 D_j^2 = A^2 + (c_j D_j)^2 \leq (A + c_j D_j)^2 = A'^2 \\ B^2 + c_j^2 D_j^2 = B^2 + (c_j D_j)^2 \leq (B + c_j D_j)^2 = B'^2 \end{cases}$$

(this is the case because both Ac_jD_j and Bc_jD_j are nonnegative).

This prove that π_i^{dev} is maximum for $b_i = 1$.

Hence not reshaping and conforming to the induced strategy in the subgame yields a higher payoff than deviating from the strategy profile for player i if

$$V > \pi_i^{dev}(1)$$

idem est if

$$\frac{\pi_i(0,0)}{1-\delta} > \pi_i^{dev}(1) = \pi_i(1,0) - \xi_i + \frac{\delta}{1-\delta}\pi_i(1,1)$$

which is equivalent to

$$\pi_i(0,0) > (1-\delta)(\pi_i(1,0) - \xi_i) + \delta\pi_i(1,1) \quad (*)$$

Now if $\pi_i(0,0) > \pi_i(1,1)$, then there exists $\overline{\delta} \in]0,1[$ such that for any discount factor $\delta \in]\overline{\delta},1[$, inequality $(*)$ is true. Indeed, when δ tends toward 1, the right-hand side of $(*)$ tends to $\pi_i(1,1)$, so if $\pi_i(0,0) > \pi_i(1,1)$, then, since this inequality is strict, there exist a $\overline{\delta} \in]0,1[$ such that

$$\pi_i(0,0) > (1-\overline{\delta})(\pi_i(1,0) - \xi_i) + \overline{\delta}\pi_i(1,1) > \pi_i(1,1)$$

and then any $\delta \in]\overline{\delta},1[$ also satisfies this inequality.

But $\pi_i(0,0) > \pi_i(1,1)$ is equivalent to

$$A_i + (c_j + c_{T2S})B_i - (c_i + c_{T2S})D_i > A_i + c_{T2S}B_i - c_{T2S}D_i$$

idem est

$$c_jB_i - c_iD_i > 0$$

which is precisely the condition $\frac{c_i}{c_j} < \frac{B_i}{D_i}$. Hence if $\frac{c_i}{c_j} < \frac{B_i}{D_i}$, CSD i will not find it profitable to deviate from the strategy induced by (S) on our subgame. We have to assume similarly that $\frac{c_j}{c_i} < \frac{B_j}{D_j}$ such that CSD j will not find it profitable neither. Hence, under the assumptions of the Theorem, the strategy induced by (S) is a Nash equilibrium of the subgame.

This concludes the proof that (S) is a subgame-perfect Nash equilibrium of the whole game.□

7.2.2 Proof of Remark 2 concerning Theorem 2

We want to derive from the proof of Theorem 2 an explicit value for δ. Setting for convenience:

$$\begin{aligned}
A &= \sqrt{\pi_i(0,0)} = A_i + (c_j + c_{T2S})B_i - (c_i + c_{T2S})D_i \\
B &= \sqrt{\pi_i(1,0)} = A_i + (c_j + c_{T2S})B_i - c_{T2S}D_i \\
C &= \sqrt{\pi_i(1,1)} = A_i + c_{T2S}B_i - c_{T2S}D_i
\end{aligned}$$

inequality $(*)$ of the proof of Theorem 2 is equivalent to:

$$\begin{aligned}
A^2 &> (1-\delta)(B^2 - \xi_i) + \delta C^2 \\
A^2 - B^2 &> \delta(C^2 - B^2 + \xi_i) - \xi_i \\
\delta(B^2 - C^2 - \xi_i) &> B^2 - A^2 - \xi_i
\end{aligned}$$

If $\xi_i < B^2 - C^2$ then $(*)$ is thus equivalent to:

$$\delta > \frac{B^2 - A^2 - \xi_i}{B^2 - C^2 - \xi_i}$$

This is only possible for at least *one* value of δ if $\frac{B^2-A^2-\xi_i}{B^2-C^2-\xi_i} < 1$ *idem est* if $B^2 - A^2 - \xi_i < B^2 - C^2 - \xi_i$, which amounts to $A > C$, that is, to $c_jB_i - c_iD_i > 0$, which is true since by assumption $c_i < c_jf_i$. $(*)$ then holds for any $\delta > \frac{B^2-A^2-\xi_i}{B^2-C^2-\xi_i}$. Hence looking at the Proof of Theorem 2 again we see that setting $\overline{\delta} := \max(1 - \frac{c_i^2D_i^2}{\xi_i}, 1 - \frac{c_j^2D_j^2}{\xi_j}, \frac{B^2-A^2-\xi_i}{B^2-C^2-\xi_i}, \frac{B^2-A^2-\xi_j}{B^2-C^2-\xi_j})$ gives an explicit value for the $\overline{\delta}$ mentioned in the theorem.

If $\xi_i > B^2 - C^2$ then $(*)$ is equivalent to:

$$\delta < \frac{B^2 - A^2 - \xi_i}{B^2 - C^2 - \xi_i}$$

This only happens for all large enough values of δ if $\frac{B^2-A^2-\xi_i}{B^2-C^2-\xi_i} \geq 1$ *idem est* if $B^2 - A^2 - \xi_i \leq B^2 - C^2 - \xi_i$, which amounts again to the assumption $c_i \leq c_jf_i$. Then it is here enough to set $\overline{\delta} := \max(1 - \frac{c_i^2D_i^2}{\xi_i}, 1 - \frac{c_j^2D_j^2}{\xi_j})$ as an explicit value for $\overline{\delta}$.

So in both situation we see that $(*)$ holds for all values of δ close enough to 1 if, and only if, $c_i < c_jf_i$. More precisely we proved that $c_i < c_jf_i$ is equivalent to: "for any ξ_i there exists $\overline{\delta} \in]0,1[$ such that for any $\delta \in]\overline{\delta}, 1[$, $(*)$ holds."

This Remark will be useful for the proof of Theorem 3.

Note in passing that $\frac{B^2-A^2-\xi_i}{B^2-C^2-\xi_i}$ can be simplified by rewriting it as:

$$\frac{(B-A)(B+A) - \xi_i}{(B-C)(B+C) - \xi_i} = \frac{c_iD_i(c_jB_i - c_iD_i) - \xi_i}{c_jB_i(2A_i + (c_j + 2c_{T2S})B_i - 2c_{T2S}D_i) - \xi_i}$$

7.2.3 Proof of Theorem 4

Let N_0 be the time necessary for a given CSD to become aware of another CSD's reshaping. We adapt the proof of Theorem 2 to this new game of imperfect information by using the following strategy (S_{N_0}):

(S_{N_0}) *If it is not common belief that at least one player has reshaped so far, then do not reshape and play what you believe is the Nash equilibrium of the price-setting game (that is, play the Nash equilibrium of the price-setting game where costs corresponds to costs without reshaping for all of the players). If* *it is common belief that some player has reshaped, then reshape completely and always play the Nash equilibrium of each price-setting game.*

To prove this strategy profile yields a subgame-perfect Nash equilibrium, we show, as in the proof of Theorem 2, that it satisfies the definition of a subgame-perfect Nash equilibrium, that is, it induces a Nash equilibrium on every subgame of the whole game. The whole game consists of two types of subgames: those where it is known by all CSDs that at least one CSD has reshaped in one of the previous periods, and those where no CSD is commonly known to have reshaped so far.

Consider first a subgame T of the first type. Then the strategy induced by (S_{N_0}) on the subgame T consists of reshaping at stage 1) of the first stage-game $G_{reshape}$ of T in case it has not been done yet, and then playing repeatedly the Nash equilibrium of the price-setting game at each subsequent stage-game. Consider a given CSD i and assume all other CSDs than i are playing according to the induced strategy. We proceed to determine CSD i's best-response – the goal being to prove that i's best-response is also to conform to the induced strategy.

First, since in the subgame T other CSDs are repeatedly playing the Nash equilibrium of the price-setting game, it is easy to see that CSD i's best strategy is to play each time the Nash equilibrium of the price-setting game given its costs, regardless of the degree b_i of reshaping it had possibly chosen and whenever it had possibly chosen to reshape. Indeed, if CSD i deviates in at least one period from playing the Nash equilibrium of the price-setting game, then, by definition of the Nash equilibrium of the price-setting game, CSD i profits on this period would be lower, and so will be its profits for the whole game, since the profits for the whole game are just the discounted payoffs of the profits at each stage game. Hence, in any of its best-responses, CSD i will always choose to play the Nash equilibrium of each price-setting game.

Second, by Lemma 1, if $\delta > \frac{c_i^2 D_i^2}{\xi_i}$, CSD i cannot avoid to reshape at some point. Indeed, Lemma 1 for $\delta > \frac{c_i^2 D_i^2}{\xi_i}$ predicts a higher profit for CSD i if it chooses to reshape in the first period of the model than if it choose to never reshape. Note that this does *not* imply, as such, that the optimal outcome is to reshape immediately (that is, in the first period of the subgame T), since Lemma 1 only compares the outcome of reshaping *in the first period* with no reshaping at all. But this proves that reshaping must indeed occurs in any best-response to other CSDs' strategy (possibly not in the first period of the subgame but in a later period). All that is left to prove is that the reshaping is complete (meaning $b_i = 1$ is chosen) and that it occurs in the first period of the subgame T, which we will call period 0. Let t_0 be the period in which reshaping occurs.

By Lemma 1 again, but applied to the subgame of T starting at time t_0 in which CSD i choose to reshape, we see that, because $\delta > \frac{c_i^2 D_i^2}{\xi_i}$, $b_i = 1$ is optimal. Hence reshaping, when it occurs, is always complete.

Assume that reshaping occurs at a period $t_0 > 0$. Then the additional profits from reshaping one period earlier would have been:

$$-\xi_i + \delta\xi_i + P$$

where P is the additional profits in one period obtained from a low cost situation compared to a high cost situation, *idem est*

$$\begin{aligned} P &= (A_i + c_{T2S}B_i - c_{T2S}D_i)^2 - (A_i + c_{T2S}B_i - (c_i + c_{T2S})D_i)^2 \\ &= 2c_iD_i(A_i + c_{T2S}B_i - c_{T2S}D_i) - (c_iD_i)^2 \\ &= c_iD_i\{(A_i + c_{T2S}B_i - c_{T2S}D_i) + (A_i + c_{T2S}B_i - (c_{T2S} + c_i)D_i)\} \end{aligned}$$

It can be seen, either from the first expression or from the last one (sum of two positive quantities), than $P > 0$. Hence, for any δ satisfying $1 > \delta > 1 - \frac{P}{\xi_i}$, we have $-\xi_i + \delta\xi_i + P > 0$, a contradiction since we assumed i was playing its best-response. We conclude that the reshaping occurs at time 0 in T. Hence CSD i's unique best-response is to conform too to the induced strategy. This shows that the strategy (S_{N_0}) induces a Nash equilibrium on any subgame of the first type.

Consider now a subgame U of the second type. Then the strategy profile induced by (S_{N_0}) on this subgame is precisely (S_{N_0}) itself, and in particular it implies not to reshape at stage 1) of the first period of the subgame, since no one has ever reshaped before. Now does it constitute a Nash equilibrium of the subgame? Consider first the value V of playing this induced strategy for CSD i, assuming all other CSDs also play it. V is equal to the sum of the immediate payoff $\pi_i(0,0)$ plus the discounted value of playing this strategy on the next periods of the subgame - which happen conveniently to be equal to our whole subgame. Hence we can write

$$V = \pi_i(0,0) + \delta V$$

which yields

$$V = \frac{\pi_i(0,0)}{1-\delta}$$

Now assume CSD i would unilaterally deviate from this strategy by reshaping to some extent $b_i > 0$. Since in the N_0 following periods of the subgame U other CSDs do not reshape and play the Nash equilibrium of the price-setting game in which they are unaware of CSD i new costs, the payoff at each of these periods is denoted by π_i^{ab} where the index ab stands for "abnormal", the abnormality being that other CSDs do not know the true costs of CSD i and are thus in reality not playing the Nash equilibrium of the *relevant* price-setting game. But after N_0 periods all CSDs also reshape and CSD i profit is then only $\pi_i(b_i, b_j^*(b_i))$. Hence the total profits $\pi_i^{dev}(b_i)$ obtained from unilaterally deviating from the induced strategy by playing $b_i > 0$ at step 1) of the first period of the subgame is:

$$\begin{aligned}\pi_i^{dev}(b_i) &= \pi_i^{ab} - \xi_i b_i^2 + \delta\pi_i^{ab} + ... + \delta^{N_0-1}\pi_i^{ab} + \delta^{N_0}\pi_i(b_i, b_j^*(b_i)) + \delta^{N_0+1}\pi_i(b_i, b_j^*(b_i)) + ... \\ &= \frac{1-\delta^{N_0}}{1-\delta}\pi_i^{ab} - \xi_i b_i^2 + \frac{\delta^{N_0}}{1-\delta}\pi_i(b_i, b_j^*(b_i))\end{aligned}$$

The maximum value of π_i^{ab} assuming a given b_i and that other CSDs follow the induced strategy could be computed explicitly, then $\pi_i^{dev}(b_i)$ could be derived with respect to b_i to find its maximum with respect to b_i and the best-possible deviation thus obtained: this was the method of the proof of Theorem 2. But to avoid additional computations in this proof we will use the fact that π_i^{ab} is bounded, since the *general* payoff-function $\pi_i(b_i, b_j, p_i, p_j)$ is itself bounded.

No reshaping and conforming to the induced strategy in the subgame yields a higher payoff than deviating from the strategy profile for CSD i if, and only if,

$$V > \pi_i^{dev}(b_i)$$

i. e.:

$$\frac{\pi_i(0,0)}{1-\delta} > \frac{1-\delta^{N_0}}{1-\delta}\pi_i^{ab} - \xi_i b_i^2 + \frac{\delta^{N_0}}{1-\delta}\pi_i(b_i, b_j^*(b_i))$$

which is equivalent to

$$\pi_i(0,0) > (1-\delta^{N_0})\pi_i^{ab} - (1-\delta)\xi_i b_i^2 + \delta^{N_0}\pi_i(b_i, b_j^*(b_i)) \quad (*)$$

Now notice that for $\delta \geq \frac{c_j^2 D_j^2}{\xi_j}$, applying Lemma 1 to the infinite subgame of our game starting when CSD j becomes aware of CSD i reshaping decision yields $b_j^*(b_i) = 1$ for all b_i. Note that the same Lemma cannot be applied as such to CSD i, even if $\delta \geq \frac{c_i^2 D_i^2}{\xi_j}$ (because of the first N_0 periods of the subgame in which CSD i benefits from a first-mover advantage). We will actually avoid to determine explicitly the reshaping degree of CSD i here[26].

Since $b_j^*(b_i) = 1$ inequality $(*)$ is equivalent to:

$$\pi_i(0,0) > (1-\delta^{N_0})\pi_i^{ab} - (1-\delta)\xi_i b_i^2 + \delta^{N_0}\pi_i(b_i, 1) \quad (**)$$

Now if $\pi_i(0,0) > \pi_i(b_i, 1)$, then there exists $\overline{\delta} \in]0,1[$ such that for any discount factor $\delta \in]\overline{\delta}, 1[$, inequality $(**)$ is true. Indeed, when δ tends toward 1, the right-hand side of $(**)$ tends to $\pi_i(b_i, 1)$

[26] Indeed, for this particular proof we do not need to find b_i explicitly, although it will be shown in the proof of Theorem 4 that $b_i = 1$ for values of δ high enough.

because π_i^{ab} is bounded and $1-\delta^{N_0}$ tends to 0. So if $\pi_i(0,0) > \pi_i(b_i,1)$, then, since this inequality is strict, there exist a $\overline{\delta} \in]0,1[$ such that

$$\pi_i(0,0) > (1-\overline{\delta}^{N_0})\pi_i(1,0) - (1-\delta)\xi_i b_i^2 + \overline{\delta}^{N_0}\pi_i(b_i,1) > \pi_i(b_i,1)$$

and then any $\delta \in]\overline{\delta},1[$ will also satisfies this inequality.

But $\pi_i(0,0) > \pi_i(b_i,1)$ is equivalent to

$$A_i + (c_j + c_{T2S})B_i - (c_i + c_{T2S})D_i > A_i + c_{T2S}B_i - ((1-b_i)c_i + c_{T2S})D_i$$

which is equivalent to

$$c_jB_i - b_ic_iD_i > 0$$

idem est

$$c_jB_i > b_ic_iD_i$$

Since $b_i \leq 1$ and $c_iD_i \geq 0$, we see that this last inequality is implied by the condition $c_jB_i > c_iD_i$, which is precisely the condition $\frac{c_i}{c_j} < \frac{B_i}{D_i} = f_i$ assumed in the theorem. Hence if $\frac{c_i}{c_j} < \frac{B_i}{D_i}$, CSD i will not find it profitable to deviate from the strategy induced by (S_{N_0}) on our subgame U. We have to assume similarly that $\frac{c_j}{c_i} < \frac{B_j}{D_j}$ such that CSD j will not find it profitable neither. Hence, under the assumptions of the theorem, the strategy induced by (S_{N_0}) is a Nash equilibrium of the subgame U.

This concludes the proof that (S) is a subgame-perfect Nash equilibrium of the whole game.$\square$

7.3 Complement about the building block of the model

The advantage of our choice for the underlying model is its simplicity and its generality. Basically, all it says is that the quantity demanded of settlement services is a linear, decreasing function of its own price and an increasing function in the prices of competitors. That demand of a given product decreases with its price and increase with the prices of its available substitutes is a fairly general assumption. Moreover linearity is often assumed in economics for simplicity reasons. Nevertheless the world is often not linear, hence in the case of a derivable function linearity should better be regarded as a first approximation.

Consider again the building block for our model:

$$\begin{cases} q_i = \alpha_i - \gamma_{ii}p_i + \gamma_{ij}p_j \\ q_j = \alpha_j - \gamma_{jj}p_j + \gamma_{ji}p_i \end{cases}$$

Writing $q_i = \alpha_i - \gamma_{ii}p_i + \gamma_{ij}p_j$ may strictly speaking allow for negative quantities: for too high prices p_i compared to the competitor, namely for $p_i > \frac{\alpha_i + \gamma_{ij}p_j}{\gamma_{ii}}$ the quantity q_i becomes negative. Of course, since it is not possible to produce negative quantities we should assume the firm would stop producing (or the CSD stop providing settlement services). Hence the model should better be re-written as:

$$q_i = \max(0, \alpha_i - \gamma_{ii}p_i + \gamma_{ij}p_j)$$

Now, a firm which has stopped producing and continues fixing too high a price should not influence the quantity sold by the other firm, that is, we do not want that q_j tends to infinity when p_i tends to infinity. A possible way to bypass this pitfall is to re-interpret the price p_i in the above equation as the price at which at least some transaction has been brought from i. Hence, while it is more convenient to let the price p_i fixed by CSD i be arbitrary, the price appearing in the formula determining the quantities will only be precisely p_i if:

$$p_i < \frac{\alpha_i + \gamma_{ij}p_j}{\gamma_{ii}} =: p_{i,\lim}$$

For $p_i > p^i_{\lim}$ it will just be $p^i_{\lim}$, and similarly interchanging i and j.

Hence the formula giving q_i becomes:

$$q_i = \max(0, \alpha_i - \gamma_{ii}p_i + \gamma_{ij}\min(p_j, p_{j,\lim}))$$

Remark: For the same reasons as $p_i < \frac{\alpha_i+\gamma_{ij}p_j}{\gamma_{ii}}$, we have

$$p_j < \frac{\alpha_j + \gamma_{ji}p_i}{\gamma_{jj}}$$

hence

$$p_i < \frac{1}{\gamma_{ii}}(\alpha_i + \gamma_{ij}\frac{\alpha_j + \gamma_{ji}p_i}{\gamma_{jj}})$$

This solves in

$$p_i < \frac{1}{\gamma_{ii}}(1 - \frac{\gamma_{ij}\gamma_{ji}}{\gamma_{ii}\gamma_{jj}})^{-1}(\alpha_i + \frac{\gamma_{ij}}{\gamma_{jj}}\alpha_j) =: p^i_{\max}$$

Hence formally we could (but do not need to) restrain the action space of the price-setting game of CSD i to the segment $[0, p^i_{\max}]$. Note we will still need to use the formula $q_i = \max(0, \alpha_i - \gamma_{ii}p_i + \gamma_{ij}\min(p_j, p_{j,\lim}))$ since not all elements in $[0, p_{j,\max}]$ implies a positive quantity for q_i. That is why there is no point in reducing the action set.

7.4 Link with the Hoteling and Salop model

The Salop model with n firms can be seen as n Hoteling models assembled one after the other around a circle. Clearly, a Salop model with only two firms separating the circle in two *equal* parts is equivalent to the Hoteling model. Hence, the Salop model is an extension of the Hoteling model.

The main reason why many articles make use of the Salop model with three firms only is, besides the complexity of computations involved for more firms, that with only three firms the Salop model allows competition between any two firms to take place. Indeed, for more than three firms, it becomes difficult to justify the use of such a spatial model, where two non-adjacent firms can only start competing when they (jointly) wipe out the entire share of all the firms standing between them. Moreover, it is combinatorially difficult to compute the equilibrium for more than three firms when, for example, location-specific network effects are added, as is the case, for example, in Matutes and Padilla [20], and Cales et al [1].

Our model extends to more than two firms; for an arbitrary number of firms, write

$$q_i = \alpha_i - \gamma_{ii}p_i + \sum_{j\neq i}\gamma_{ij}p_j$$

For example, for three firms we get:

$$\begin{aligned} q_i &= \alpha_i - \gamma_{ii}p_i + \gamma_{ij}p_j + \gamma_{ik}p_k \\ q_j &= \alpha_j - \gamma_{jj}p_j + \gamma_{ji}p_i + \gamma_{jk}p_k \\ q_k &= \alpha_k - \gamma_{kk}p_k + \gamma_{ki}p_i + \gamma_{kj}p_j \end{aligned}$$

Summing gives the aggregate demand, and as previously we see that it is inelastic when $-\gamma_{ii} + \gamma_{ij} + \gamma_{ik} = 0$ for all triplet (i, j, k).

In Cales et al model [1], demand can be expressed in a similar way: in the incompatibility case, take $\gamma_{ii} = 1/T$ and $\gamma_{ij} = 1/(2T)$ for any $i \neq j$. Since $-\gamma_{ii} + \gamma_{ij} + \gamma_{ik} = 0$, we find again that demand is unrealistic.

Now, Cales model, in the other compatibility scenario, incorporates network effects on the clients' utility. How can these effects be factored in our model?

An increase in utility (resulting from CSD i choice not yet modelled here), which increase the price a consumer is willing to pay, can be captured by an increase in α_i. Hence, the demand curve price is shifted to the right: at any given quantity clients are willing to pay more than before, because of the increased utility they derive from the services.

Now, contrary to the Salop modelling used by Cales et al, there is no competitive effect in this utility increase effect: at equilibrium we notice that not only π_i has increased, but also π_j (still π_j increases 4/15 less than π_i).

To get the same modelling we should assume $q_i = a_i + \alpha_i - \alpha_j - \alpha_k - \gamma_{ii}p_i + \gamma_{ij}p_j + \gamma_{ik}p_k$ where a_i is a constant chosen high enough to make the quantities positive and $\alpha_i, \alpha_j, \alpha_k$ are now devoted uniquely to the increased utility of belonging to a bigger network (that can be specified using such or such model, in particular we can give them the same form as in Cales et al article [1]).

7.5 Generalising to n CSDs

7.5.1 When CSDs are given the choice to reshape in the first period of the model only

Although throughout the article we favoured the view, when dealing with many CSDs, to focus on a given CSD i and interpret the other CSD j as representing the overall market (minus CSD i), the model could easily be extended to more than two firms, provided some symmetric assumptions are made in order to keep the complexity of the computations down. Indeed, the general model for n CSD has its demand for CSD i settlement services given by:

$$q_i = \alpha_i - \gamma_{ii}p_i + \sum_{j \neq i} \gamma_{ij}p_j$$

We will in the whole section assume that the cross-elasticities are all equal:

$$\gamma_{ij} = \gamma_{kl} := \gamma'$$

for any $i \neq j$ and $k \neq l$, as well as

$$\gamma_{ii} = \gamma_{jj} = \gamma$$

for any i, j.

With this equation demands become:

$$q_i = \alpha_i - \gamma p_i + \gamma' \sum_{j \neq i} p_j$$

and profits can be written as before as

$$\pi_i = (p_i - \widetilde{c}_i)q_i - \widetilde{C}_i$$

For a given vector a of length n always denote by a_{-i} the vector of length $n-1$ obtained from a by deleting the ith element of a. For example if p is the vector of prices $(p_1, ..., p_n)$ then p_{-i} the vector $(p_1, ..., p_{i-1}, p_{i+1}, ...p_n)$. The best-response function for the price-setting stage is:

$$p_i^*(p_{-i}) = \frac{1}{2\gamma}(\alpha_i + \gamma\widetilde{c}_i + \gamma' \sum_{j \neq i} p_j)$$

Solving the system $\{p_i^* = p_i^*(p_{-i}^*),\ i \in \{1, ..., n\}\}$(for example, by summing all equations to find an expression for $\sum_j p_j^*$) gives:

$$p_i^* = \frac{(2\gamma - (n-2)\gamma')(\alpha_i + \gamma\widetilde{c}_i) + \gamma' \sum_{j \neq i}(\alpha_j + \gamma\widetilde{c}_j)}{(2\gamma + \gamma')(2\gamma - (n-1)\gamma')}$$

We notice that at equilibrium $p_i^* - \widetilde{c}_i = \frac{1}{\gamma}q_i^*$. Hence $\pi_i^* = \gamma(p_i^* - \widetilde{c}_i)^2 - \widetilde{C}_i$, yielding

$$\pi_i^* = (A_i + B\sum_{j \neq i} \widetilde{c}_j - D\widetilde{c}_i)^2 - C_i$$

with

$$
\begin{aligned}
A_i &= \frac{\sqrt{\gamma}\{(2\gamma-(n-2)\gamma')\alpha_i+\gamma'\sum_{j\neq i}\widetilde{\alpha}_j\}}{(2\gamma+\gamma')(2\gamma-(n-1)\gamma')}\\
B &= \frac{\gamma\gamma'\sqrt{\gamma}}{(2\gamma+\gamma')(2\gamma-(n-1)\gamma')}\\
D &= \frac{\sqrt{\gamma}\{2\gamma^2-(n-2)\gamma\gamma'-(n-1)\gamma'^2\}}{(2\gamma+\gamma')(2\gamma-(n-1)\gamma')}
\end{aligned}
$$

We thus deduce π_i^{tot}:

$$
\pi_i^{tot}=\widetilde{\delta}(A_i+(\sum_{j\neq i}((1-b_j)c_j+c_{T2S}))B-((1-b_i)c_i+c_{T2S}))D)^2-\widetilde{\delta}(1-a_i)C_{i,fixed}-C_{i,adapt}(a_i,b_i)
$$

with

$$
\widetilde{\delta}=\begin{cases} N & \text{if } \delta=1\\ \frac{1-\delta^N}{1-\delta} & \text{if } 0\leq\delta<1\end{cases}
$$

Let $b=(b_1,...b_n)$. Assuming $C_{i,adapt}(a_i,b_i)=\xi_i b_i^2$, this expression can be derived with respect to b_i to find the best-response function in terms of reshaping in a N period game or in an infinite game with $\delta<1$. Notice in passing we do not assume symmetric adaptation costs, since ξ_i is allowed to depend on CSD i.

In particular, we have an equivalent to Lemma 1:

Lemma 2 *Assume $\xi_i<\widetilde{\delta}c_i^2D_i^2$. Then if CSDs engage in price competition, the best-response function $b_i^*(b_j)$ is the constant function $b_i^*(b_{-i})=1$, which represents a complete reshaping decision from CSD i whatever the degrees b_j, $j\neq i$ chosen by the other CSDs.*

Proof: Because $\xi_i<\widetilde{\delta}c_i^2D_i^2$ the function π_i^{tot} is first decreasing in b_i, then increasing. Hence its maximum is either reached at $b_i=0$ or at $b_i=1$. We prove that $\pi_i^{tot}(0,b_{-i})<\pi_i^{tot}(1,b_{-i})$ for any b_{-i}.

$\pi_i^{tot}(0,b_{-i})<\pi_i^{tot}(1,b_{-i})$ is equivalent to

$$
\widetilde{\delta}(A_i+(\sum_{j\neq i}c_j+c_{T2S})B-(c_i+c_{T2S})D)^2+\xi_i<\widetilde{\delta}(A_i+(\sum_{j\neq i}c_j+c_{T2S})B-c_{T2S}D)^2
$$

But since $a^2+b^2<(a+b)^2$ for any $ab>0$, we have

$$
\widetilde{\delta}(A_i+(\sum_{j\neq i}c_j+c_{T2S})B-(c_i+c_{T2S})D)^2+\widetilde{\delta}c_i^2D^2<\widetilde{\delta}(A_i+(\sum_{j\neq i}c_j+c_{T2S})B-(c_i+c_{T2S})D+c_iD)^2
$$

and this inequality implies the above, since $\xi_i<\widetilde{\delta}c_i^2D^2$.□

Assume now $\xi_i<\widetilde{\delta}c_i^2D^2$ for all i. Then π_i^{tot} reaches its maximum in some point $b_i^{**}(b_{-i})=\psi_i(A_i+(\sum_{j\neq i}c_j+c_{T2S})B-(c_i+c_{T2S})D))-\psi_iB\sum_{j\neq i}b_j$. Solving for the (unrestricted) solution to the system $\{b_i^{**}=b_i^{**}(b_{-i}^{**}),\ i\in\{1,...,n\}\}$ yields:

$$
b_i^{**}=\frac{((1-\frac{1}{\psi_i}\sum_{j\neq i}\psi_j)\beta_i+\sum_{j\neq i}\beta_j}{1+\psi_iB\sum_{j\neq i}\frac{\psi_j(1-\psi_iB)}{\psi_i(1-\psi_jB)}}
$$

with $\beta_i:=\psi_i(A_i+(\sum_{j\neq i}c_j+c_{T2S})B-(c_i+c_{T2S})D)$.

This proves the following theorem (note we will not consider all the corner solutions as in the two-CSDs case, but just assume the previously computed quantities belongs indeed to the action set of each CSD:

Theorem 8 *Assume CSDs are engaging in price competition, and that* $\xi_i > \widetilde{\delta} c_i^2 D_i^2$ *and* $b_i^{**} \in [0,1]$ *for each CSD* i*. Then the optimal degree of reshaping is given by* $b_i^* = b_i^{**}$*.*

Simpler formula, assuming symmetric variable costs for CSDs. In order to provide a less complex formula, we could have assumed perfect symmetry of the CSDs and markets. In particular, CSDs are assumed to have the same costs c_i. We can derive very easily a simpler expression for b_i^* under the assumptions of Theorem 8. Indeed by symmetry $b_i^{**} = b_j^{**}$ for any CSD i, j hence,

$$\begin{aligned} b_i^{**} &= \beta_i - \psi_i B \sum_{j \neq i} b_j^{**} \\ &= \beta_i - \psi_i B(n-1) b_i^{**} \end{aligned}$$

Hence

$$b_i^{**} = \frac{\psi_i}{1 + (n-1)\psi_i B}(A_i + ((n-1)B - D)(c_i + c_{T2S})$$

One can check this is indeed, for $n = 2$, an equivalent formula to the one given by Theorem 1, assuming $\gamma_{jj} =: \gamma$, $\gamma_{ij} := \gamma'$ for any $i \neq j$.

Remark: using this formula, by computing the derivative of b_i^{**} with respect to the number n of CSDs, one can show that provided the market are large enough (i.e α_i high enough), or that the costs $c_i + c_{T2S}$ are high enough, the degree of optimal reshaping is an increasing function of the number of CSDs in the market. More precisely, in the symmetric case b_i^{**} is a (strictly) increasing function of n if, and only if,

$$(c_i + c_{T2S})(1 - \psi_i D) > A_i \psi_i$$

7.5.2 When CSDs are given the choice to reshape at any period: collusion theorem and immediate reshaping

Using very similar arguments to the ones employed in the proofs of Theorems 2 and 3, one could prove:

Theorem 9 *If* $c_i < f_i(\sum_{j \neq i} c_j)$ *for all* i *and the discount factor is high enough, a collusion-type theorem holds: there exists a subgame perfect Nash equilibrium where CSDs indefinitely delay reshaping. The strategy sustaining it is the same as in Theorem 2.*

Theorem 10 *If CSD* i*'s costs satisfy* $c_i > f_i(\sum_{j \neq i} c_j)$*, and the discount factor is high enough, and assuming CSDs engage in price competition, then CSD* i *will always reshape in any subgame perfect Nash equilibrium.*

Note that, mathematically, it *looks* as if increased number of CSDs would tend to re-enforce the possibility of tacit collusion not to reshape, as $\sum_{j \neq i} c_j$ grows with n. Nevertheless the fact that the strategy sustaining collusion in the proof of the theorem needs this assumption do not mean there is no other strategies, with less stringent assumptions, that sustain collusion. Hence strictly speaking nothing can be deduced about the likelihood of collusion from the assumption $c_i < f_i(\sum_{j \neq i} c_j)$. Moreover, the strategy employed in the proof, that is, that each CSD assumes other CSDs' strategy is to reshape only in case some competitor has reshaped, become a less and less realistic assumption as the number of competitors in the market grows. Divergence of views on the market as well as on other market participants strategy is likely to drive the real process outside of the equilibrium path of strategy (S), resulting in CSDs reshaping.

7.6 Robustness checks relatively to the shape of the adaptation-costs function

Most of the results of the paper are robust, at least qualitatively, to the shape of the adaptation-cost function assumed. More precisely, all results from Section 3.1 onwards, because they apply to far-sighted CSDs, do not depend on the shape of the adaptation cost function - as far as only finite values are allowed.

The results of Theorem 1 as well as of Lemma 1, depends quantitatively on the adaptation costs function. Nevertheless we can still get similar qualitative results for general cost functions. For example, there is a rough equivalent to Lemma 1, which shows that for high enough discount factors, or for gently increasing adaptation cost functions, the optimal degree of reshaping is always a complete reshaping, whatever the other CSD reshaping:

Proposition 4 *Assume price competition, and that the adaptation cost function is derivable, and that its derivative is bounded from above by M. Then there exists a discount factor $\overline{\delta}_i$ such that in any $\delta > \overline{\delta}_i$, the best-response function $b_i^*(b_j)$ is the constant function $b_i^*(b_j) = 1$, which represents a complete reshaping decision from CSD i whatever the degree b_j CSD j choose to reshape. More precisely, one can choose $\overline{\delta}_i = 1 - (2c_i D_i(A_i + c_{T2S}B_i - (c_i + c_{T2S})D_i))/M$.*

Proof. Deriving π_i^{tot} with respect to the degree of reshaping b_i gives:

$$\begin{aligned}\frac{\partial \pi_i^{tot}}{\partial b_i} &= 2\widetilde{\delta} c_i D_i(A_i + ((1-b_j)c_j + c_{T2S})B_i - ((1-b_i)c_i + c_{T2S})D_i) - \frac{\partial C_{i,adapt}}{\partial b_i}(b_i) \\ &\geq 2\widetilde{\delta} c_i D_i(A_i + ((1-b_j)c_j + c_{T2S})B_i - ((1-b_i)c_i + c_{T2S})D_i) - M\end{aligned}$$

which is always positive by our choice of δ. Hence, the maximum of π_i^{tot} is attained for $b_i = 1$. ■

This is a very intuitive results, since it basically says that if CSDs are far-sighted enough, they will want to reshape completely. In particular it contains the quadratic cost function studied in the article, as well as any adaptation-cost function having a continuous first derivative.

Interestingly, choosing a linear adaptation cost function $C_{i,adapt}(b_i) = (\xi_i - C_i)b_i + C_i$ (such that $C_{i,adapt}(1) = \xi_i$) gives a discrete, 0 or 1 best-response correspondence. Indeed, setting $f_i(b_j) = c_i^2 D_i^2 + 2c_i D_i\{A_i + ((1-b_j)c_j + c_{T2S})B_i - (c_i + c_{T2S})D_i\}$ for convenience, one can show:

$$b_i^*(b_j) = \begin{cases} 1 \text{ if } \frac{\xi_i - C_i}{\widetilde{\delta}} \leq f_i(b_j) \\ 0 \text{ if } \frac{\xi_i - C_i}{\widetilde{\delta}} \geq f_i(b_j) \end{cases}$$

(note that for the b_j on the frontier, playing 0 or 1 are equivalent, and this is why we used the term correspondence instead of function). Note in passing that the higher the initial reshaping cost $C_{i,adapt}(0)$, the higher the reshaping, as the slope of the resulting adaptation cost function is lower.

This leads to the equilibra (b_i^*, b_j^*) indicated in the Table below:

	$\frac{\xi_i - C_i}{\widetilde{\delta}} \leq f_i(1)$	$f_i(1) \leq \frac{\xi_i - C_i}{\widetilde{\delta}} \leq f_i(0)$	$f_i(0) \leq \frac{\xi_i - C_i}{\widetilde{\delta}}$
$\frac{\xi_j - C_j}{\widetilde{\delta}} \leq f_j(1)$	$(1,1)$	$(0,1)$	$(0,1)$
$f_j(1) \leq \frac{\xi_j - C_j}{\widetilde{\delta}} \leq f_j(0)$	$(1,0)$	$(0,1)$ and $(1,0)$	$(0,1)$
$f_j(0) \leq \frac{\xi_j - C_j}{\widetilde{\delta}}$	$(1,0)$	$(1,0)$	$(0,0)$

Interestingly, the condition $f_i(1) \leq \frac{\xi_i - C_i}{\widetilde{\delta}} \leq f_i(0)$ and $f_j(1) \leq \frac{\xi_j - C_j}{\widetilde{\delta}} \leq f_j(0)$ allows for two Nash equilibria; since computation show that in this case the CSD having reshaped is better off than the other (assuming the same initial parameters for both CSDs), we can distinguish a first mover advantage there. Of course, starting with any parameter distribution and increasing the discount factor drives us towards the first cell of the Table, where both completely reshape.

7.7 Nash equilibrium with collusion in prices

7.7.1 Introduction

In this paper we focused our attention on tacit collusion in the decision to reshape towards T2S. Most of the literature since Friedman [12] has concentrated on tacit collusion in prices. This paper has instead focused on the question of the consequences of price competition on the incentives for reshaping towards T2S, including potential collusion therein. This focus can be justified because the European securities settlement industry will benefit from increased competition by three complementary factors after the introduction of T2S: first, the introduction of T2S will overcome a number of technical barriers to cross-border settlement and provide a single platform on which CSDs will be able to compete; second, the CSD regulation will ensure common legal standards for securities settlement in the European Union, including free choice of CSDs for users and a 'passport'-concept for authorised CSDs to provide their services in other EU member states; and third, a number of other initiatives contribute to further harmonisation of technical, legal and regulatory aspects and the associated business processes[27].

Expectations of strong price competition after the introduction of T2S appear to be confirmed by a recent announcement by Clearstream, with roughly 40% of the expected settlement volumes the biggest single CSD group in T2S, that Clearstream would not add any margin on top of the direct T2S fee (c_{T2S} in our model) after their migration to T2S (see Clearstream [4]). In order not to mix the incentives for the investment / reshaping decision given price competition with the effects of increases in price competition, we have assumed no changes in price competition over time. The degree of price competition is indirectly captured in the cross-sensitivity parameter γ_{12} and γ_{21}. We expect that explicitly modelling an increase in competition over time would further increase the benefits of reshaping towards T2S, but leave the explicit modelling for future research.

Another interesting theoretical question, closely linked to the literature following Friedman [12], is to find other Nash-equilibria of the whole game for which collusion also appears at the price level. Characterising all Nash equilibria seems an ambitious task, as many strategies become possible when the assumption of price competition is removed. Hence, we will only consider a few such equilibria, and under some simplifying assumptions. This nevertheless helps cast some light on the more complex real world situation where CSDs can potentially collude both in the timing of the reshaping, in its degree, and in the price levels fixed after reshaping. Price-collusion does not necessarily imply a greater potential for collusion in delaying the decision to reshape. This is because by colluding on prices CSDs can pocket the unit-cost-reduction stemming from reshaping rather than passing them to the market.

7.7.2 Simplifying assumptions

1) For simplicity, we will assume all CSDs' parameters are symmetric Hence we will adopt the same notations as in Annex 7.5 with

$$\gamma_{ij} = \gamma_{kl} =: \gamma'$$

for any $i \neq j$ and $k \neq l$, and

$$\gamma_{ii} = \gamma_{jj} =: \gamma$$

for any i, j, but also:

$$\begin{aligned} \alpha_i &= \alpha_j =: \alpha \\ c_i &= c_j =: c \end{aligned}$$

for any i, j.

2) We will also restrict ourselves to *symmetric* strategies. This makes the resolution easier and can to some extent be justified on the grounds of symmetric conditions, although of course symmetric parameters do not dispel the possibility of the existence of non-symmetric NE.

[27] See, e.g., the third progress report of the T2S Harmonisation Steering Group, available at `http://www.ecb.europa.eu/paym/t2s/pdf/Third_T2S_Harmonisation_Progress_Report.pdf`.

7.7.3 Analytical resolution

By definition, in an *explicit* collusion situation, prices are set jointly so as to maximize the aggregate profit. Hence in a symmetric setting we can expect each CSD's individual profits explicitly colluding to be usually higher than in a competitive equilibrium. Nevertheless, because of demand elasticity it could happen that such profits are just equal to, and no greater than, the ones made in the competitive equilibrium. Lemma 3 below gives the condition for which profits are *strictly* higher, as this will be needed for the proof of Theorem 12.

Lemma 3 *Assume the same current unit costs for both CSDs, that is, $\widetilde{c}_i = \widetilde{c}_j =: \widetilde{c}$. Then the one-stage profit in a (explicit) price collusion of the price-setting game is greater than the (unique) NE of the price setting game if, and only if,*

$$\alpha\left(\frac{1}{2(\gamma-\gamma')^{\frac{3}{2}}} - \frac{\sqrt{\gamma}}{2\gamma+\gamma'}\right) > \left(\frac{1}{2}\frac{1}{\sqrt{\gamma-\gamma'}} + \frac{\sqrt{\gamma}(\gamma\gamma' - 2\gamma^2 - \gamma'^2)}{4\gamma^2-\gamma'^2}\right)\widetilde{c} \qquad (C)$$

Proof: We assume the same current unit costs for both CSDs. This would be the consequence, for example, of same individual unit-cost as implied by our symmetric assumption *and* of the same degree of prior reshaping. Then $\widetilde{c}_i = c_{T2S} + (1-b_i)c = c_{T2S} + (1-b_j)c = \widetilde{c}_j =: \widetilde{c}$, and profits in the price setting stage obtained **under competition** can be expressed as:

$$\pi_i^* = (A + B\widetilde{c}_j - D\widetilde{c}_i)^2 = (A + (B-D)\widetilde{c}_i)^2 =: \pi^*.$$

with

$$\begin{aligned} A &= \frac{\alpha\sqrt{\gamma}}{2\gamma+\gamma'} \\ B &= \frac{\gamma^{\frac{3}{2}}\gamma'}{(4\gamma^2-\gamma')} \\ D &= \frac{\sqrt{\gamma}(2\gamma^2-\gamma'^2)}{4\gamma^2-\gamma'^2} \end{aligned}$$

Let's assume **price-collusion** in one of the stage price setting game (note this cannot result in a NE of the one-period game). Because we are only looking at symmetric equilibria, this is tantamount to looking at identical prices $p_i = p_j = p$ that a monopoly consisting of both CSDs would set. Profits to each of the CSD for setting a common price p is

$$\pi = (\alpha + (\gamma - \gamma')p)(p - \widetilde{c})$$

The maximum is obtained for

$$p = \frac{\widetilde{c}}{2} + \frac{\alpha}{2(\gamma-\gamma')}$$

This results in an individual profit of:

$$\pi = \frac{1}{4}\frac{1}{\gamma-\gamma'}\left(\frac{\alpha}{\gamma-\gamma'} - \widetilde{c}\right)^2$$

Now, profits can be higher in the price-collusion case if, and only if,

$$\pi > \pi^*$$

This can be re-written as:

$$\frac{1}{2}\frac{1}{\sqrt{\gamma-\gamma'}}\left(\frac{\alpha}{\gamma-\gamma'} - \widetilde{c}\right) > A + (B-D)\widetilde{c}$$

Replacing A, B and C by their value and re-arranging yields precisely the condition (C).

Remark 11 *If we denote by p^* the price corresponding to price competition, we can express the difference of prices of p^* and p, as*

$$p - p^* = \frac{\gamma'}{2(2\gamma+\gamma')}\widetilde{c} + \frac{\alpha\gamma'}{2(\gamma-\gamma')(2\gamma-\gamma')}$$

Hence, the higher the unit-cost, the higher the potential price charged by the de facto monopoly compared to the competitive equilibrium price. p is here the price which maximizes profits in a monopoly. Nevertheless as in what follows the (tacit) collusion is obtained not by assuming a monopoly setting prices but by indefinitely repeating the price-setting game and punishing the other CSD in case it deviates by playing the competitive price, and any price between p^* and p can be sustained in the equilibria we will define.

Theorem 12 *Assume that $\widetilde{c} := c_{T2S}$ satisfies (C). Let p be any price strictly higher than $p^* = \frac{\alpha}{2\gamma-\gamma'} + \frac{\gamma c_{T2S}}{2\gamma+\gamma'}$ but no greater than $\frac{c_{T2S}}{2} + \frac{\alpha}{2(\gamma-\gamma')}$. Then for discount factors high enough, there exists a subgame perfect Nash equilibrium where CSDs collude in prices after reshaping.*

Proof: Let $p \in]\frac{\alpha}{2\gamma-\gamma'} + \frac{\gamma c_{T2S}}{2\gamma+\gamma'}; \frac{c_{T2S}}{2} + \frac{\alpha}{2(\gamma-\gamma')}]$. Consider the following strategy, that we denote by (S_{col}): *"If you have not reshaped yet, then fully reshape immediately. Then, if someone has not reshaped yet then play p^* in each of the subsequent price-setting stage. Otherwise play p as long as everyone plays p, and play p^* indefinitely otherwise."*

We will prove (S_{col}) is a subgame-perfect Nash equilibrium. We need first the following Lemma:

Lemma 4 *For high enough discount factors, a given CSD will fully reshape at the beginning of the game whether it plans to follow the strategy (S_{col}) or to defect. That is, it will always choose to reshape with a degree of reshaping equal to 1 in the first period of the game.*

Proof of Lemma: Let b be the degree of reshaping which maximizes total profits of CSD i in a game where price-collusion is assumed to happen.

$$\begin{aligned}
\pi_i^{tot} &= \widetilde{\delta}(A + ((1-b)c_j + c_{T2S}))B - ((1-b)c_i + c_{T2S}))D)^2 \\
&= \widetilde{\delta}(A + ((1-b)c + c_{T2S}))(B-D))^2
\end{aligned}$$

by symmetry. But

$$\begin{aligned}
D - B &= \frac{\sqrt{\gamma}}{4\gamma^2 - \gamma'^2}(2\gamma^2 - \gamma'^2 - \gamma\gamma') \\
&= \frac{\sqrt{\gamma}}{4\gamma^2 - \gamma'^2}(\gamma^2 - \gamma'^2 + \gamma(\gamma-\gamma'))
\end{aligned}$$

with $\gamma^2 - \gamma'^2 \geq 0$ and $\gamma - \gamma' \geq 0$ since $\gamma \geq \gamma'$ by assumption. Hence $D - B \geq 0$ and π_i^{tot} is maximum for $b = 1$.

Now assume a given CSD wants to deviate from the strategy. Because the CSD non-deviating from it will then switch to the non-cooperative mode (playing the competitive price equilibrium of each subsequent price-setting game), the optimal degree of reshaping is precisely the one already derived in the corpus of this paper. When δ tends to 1, $\widetilde{\delta} = 1/(1-\delta)$ tends to $+\infty$ and so $\xi_i < \widetilde{\delta}c_i^2 D_i^2$. Hence, by Lemma 1:

$$b^* = 1$$

This completes the proof of Lemma 4.

We now resume the proof of Theorem 12. Let H be a subgame of the game. Let always be i the CSD playing on the considered subgame H according to the strategy (S'_{col}) induced by (S_{col}) on H.

If H is **on** the equilibrium path, it can only be of one of the following type:

- type A: H contains the first stage *idem est* H is the whole game. Then, by Lemma 4, playing $b_j = 1$ is optimal for CSD j. After the reshaping phase, because CSD i is following (S_{col}), we are left with a simple infinitely repeated game where one of the player (CSD i) is playing a simple grim trigger strategy to promote playing $p > p^*$. Because condition (C) ensures playing p indeed results in *strictly* higher payoffs than playing the competitive equilibrium price p^* for each of the stage game, the trigger strategy is a NE on this subgame for high-discount factors. Hence the best-response answer, for high enough discount factors, is for CSD j to play p too, and punish i if it deviates. Overall, CSD j is thus also conforming with (S_{col}) on H.

- type B: H does not contain the first stage. Because H lies on the equilibrium path by assumption the reshaping for both CSDs has already occurred in the first stage of the game. Hence H again only consists of the price-setting game being repeated an infinite number of times. Condition (C) coupled with high enough discount factors imply the simple grim trigger strategy induced by the strategy (S_{col}) on the subgame H is a Nash equilibrium.

If H is **off** the equilibrium path, it can be of the following type:

- type C: the history leading to H is one where the CSD i has not reshaped yet then since we assume i will follow the strategy (S'_{col}) induced by (S_{col}) on H, CSD i will immediately reshape with $b_i = 1$. Discarding the history leading to H gives a game identical to the whole game, that is, to the type A subgame. The same reasoning applies and CSD j's best-response is again to conform with (S_{col}).

- type D: the history leading to H is one where the CSD i has reshaped. Then either j has reshaped at the *same* point in time and played p from that point in time till now, and then i's strategy is a grim-trigger strategy for which the strategy induced by (S_{col}) on H is again j's, or not. If j has played another move than the one advocated by (S_{col}) during the history leading to H then i is playing the competitive price equilibrium p^* of the price-setting stage indefinitely. Clearly, j's best-response is p^*, since p^* is the unique Nash equilibrium of the stage game. This achieves the proof of subgame perfection.

Remark 13 *The following strategy also sustains the price-collusion Nash equilibrium, and leads to the same equilibrium path: "Reshape in the first period of the game completely. If everyone has reshaped in the first period and played until now p in the price-setting game then play p. Otherwise play p^*." The proof is somehow simpler than for the strategy (S_{col}) we chose since it is clear Type C and D collapse together. Nevertheless the strategy (S_{col}) is more adapted to allow a delay in the decision to reshape. Also, as soon as p is chosen close to the optimal $p^* = \frac{\widetilde{c}}{2} + \frac{\alpha}{2(\gamma-\gamma')}$ where p^* is the price-setting stage game Nash equilibrium for current costs $\widetilde{c}$, there is little incentive for CSDs to delay the reshaping: for high discount rates a Pareto-dominating subgame-perfect Nash equilibrium is indeed to reshape immediately and completely, while price-colluding on each subsequent price-setting stage by playing exactly $p(\widetilde{c}) = \frac{\widetilde{c}}{2} + \frac{\alpha}{2(\gamma-\gamma')}$ for $\widetilde{c} = c_{T2S} + c$. Thus, the unit-cost reductions obtained from reshaping are partly absorbed by the CSDs themselves and not passed to the market. Because $\widetilde{c}$ decrease with reshaping, the market still has a decreased colluding price compared to a (hypothetical, not modelled here) situation of prior price-collusion with higher costs. If the prior situation was closer to $p^*(c)$, the prices may increase as a result of the new price-collusion.*

7.8 Discussion on the effect of greater competition on the relevance tacit collusion theorem

The model captures competition effects through the cross-elasticities γ_{ij}, which reflect a (fixed) substitution effect assumed between the two different markets. Indeed, as noticed earlier, for $\gamma_{ij} < \gamma_{jj}$, an increase of one unit of the price p_j leads to a fall of γ_{jj} of volume of CSD j, but to an increase of volume of γ_{ij} for CSD i. In the extreme case where $\gamma_{ij} = \gamma_{jj}$, an increase of one unit of the price p_j leads to no fall in the aggregate volume of settlements, which can be interpreted as "the whole demand for settlement services having left market j due to the price increase moves to market i". Hence, the closer γ_{ij} is to γ_{jj}, the greater the substitution effect. Now, it is interesting to wonder if a greater competition (substitution) effect invalidates the tacit collusion theorem. That is, does greater competition decrease the chances of tacit collusion to avoid reshaping? Interestingly, the answer is that it does quite the opposite: the distribution of individual CSDs' costs satisfying the condition of the tacit collusion theorem 2 becomes larger with greater competition. To show this, let us for convenience assume

$$\gamma_{ij} = \gamma_{kl} =: \gamma'$$

for any $i \neq j$ and $k \neq l$, and

$$\gamma_{ii} = \gamma_{jj} =: \gamma$$

for any i, j. The condition under which there is a subgame perfect Nash equilibrium in which CSDs indefinitely delay the moment of their reshaping, i.e. tacitly collude not to reshape, is:

$$\frac{1}{f_j} < \frac{c_i}{c_j} < f_i$$

with

$$f_i = f_j = \frac{\gamma\gamma'}{2\gamma^2 - \gamma'^2}$$

Hence the condition becomes, if we let $z = \frac{\gamma'}{\gamma}$

$$\frac{2}{z} - z < \frac{c_i}{c_j} < (\frac{2}{z} - z)^{-1}$$

Now as explained previously, greater competition is translated by the model parameter γ' coming closer to γ. That is, by an increase by some factor $a > 1$ of the ratio $z = \frac{\gamma'}{\gamma}$, which becomes $az > z$ (with, of course, $az < 1$). The range of admissible ratios $\frac{c_i}{c_j}$ actually becomes larger when z increases to az, because $\frac{2}{az} - az < \frac{2}{z} - z$, as well as $(\frac{2}{z} - z)^{-1} < (\frac{2}{az} - az)^{-1}$. This is because $\frac{2}{az} - az < \frac{2}{z} - z$ is equivalent to $2(\frac{1-a}{a}) < z^2(a-1)$ which is trivially true for $a > 1$ since $2(\frac{1-a}{a}) < 0 < z^2(a-1)$. This concludes the proof.

Therefore, the set of parameters for which tacit collusion is sustainable in a high competition / substitution environment is larger than in a low competition / substitution environment.

7.9 Graphs from simulations of Section 2.4.4

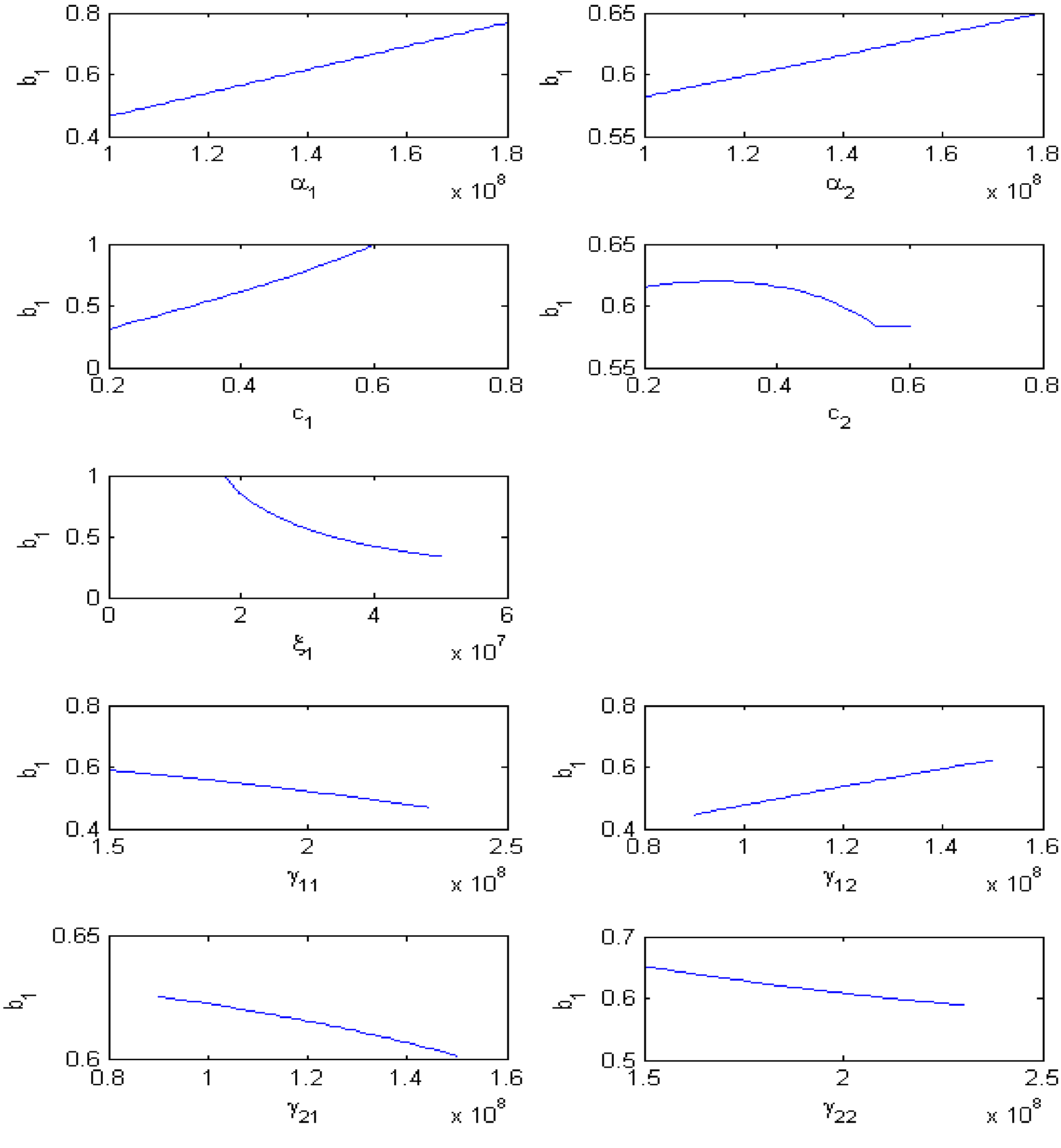


Figure 1: Two-dimensional graphs of the optimal reshaping function of CSD 1

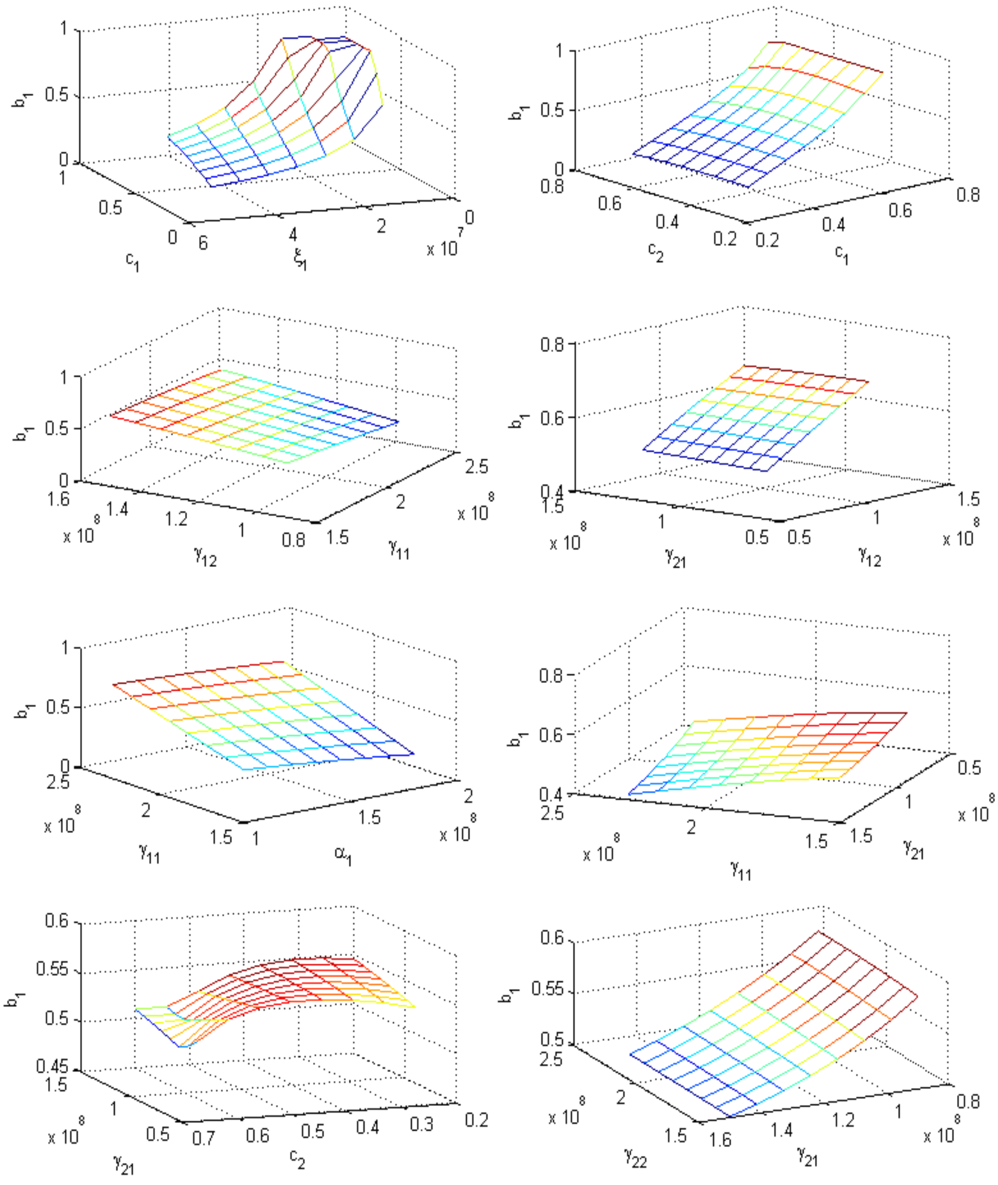


Figure 2: Three-dimensional graphs of the optimal reshaping function of CSD 1

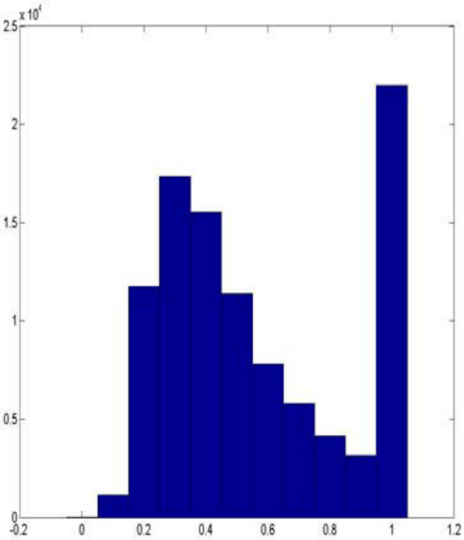


Figure 3: Histogram of the optimal reshaping function of CSD 1

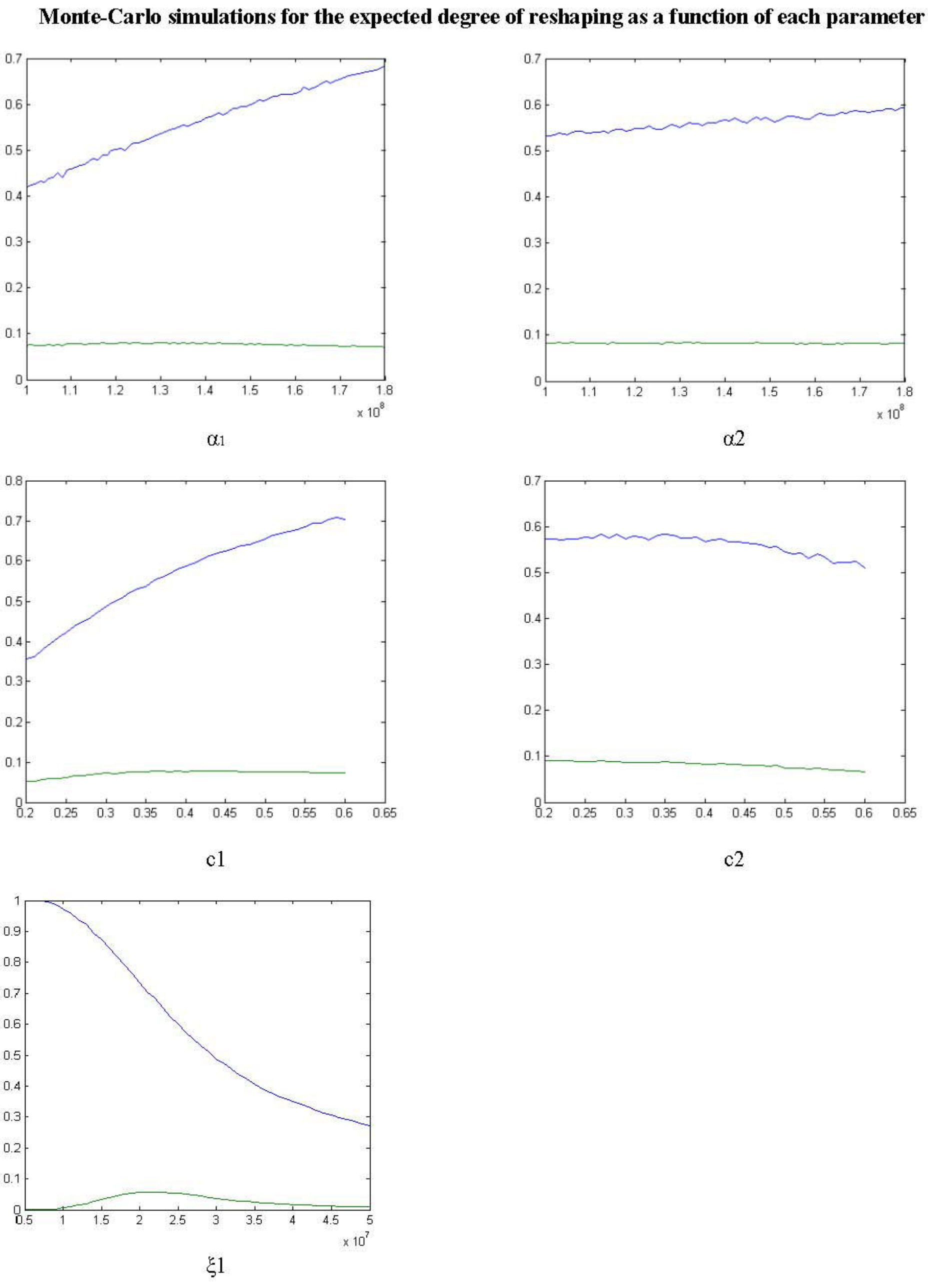
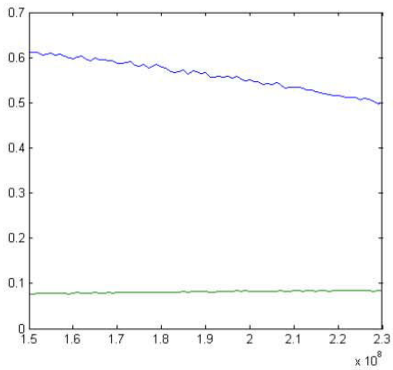
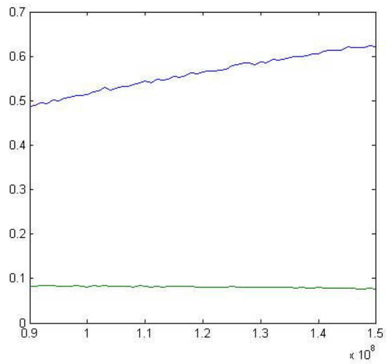


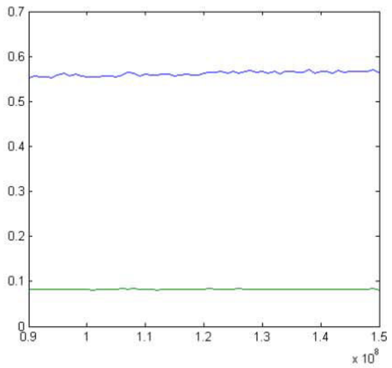
Figure 4: Average value (top line of each graph) and variance (bottom line of each graph) of the degree of reshaping (10000 simulations for each point)



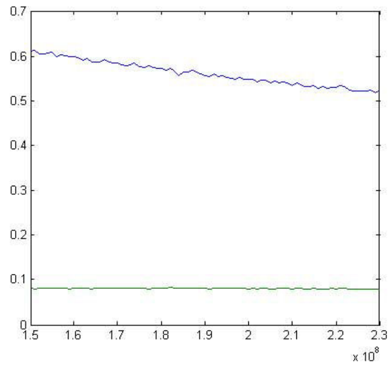
$\gamma 11$



$\gamma 12$



$\gamma\ 21$



$\gamma\ 22$

Figure 5: Average value (top line of each graph) and variance (bottom line of each graph) of the degree of reshaping (10000 simulations for each point)

Article 4: Endogenizing the network effect: how Monte-Carlo simulations and Graph Theory could benefit economic modelling[1]

1 Introduction

A general aspect of many market infrastructures or organisations is their inherent network effects: for example, additional market participants in a trading or settlement system can increase the benefit of participation for all other market participants. More participants implies each agent can do business with more counterparties, hence that liquidity increases and capital costs tend to decrease (see e.g. [14], chapter 5). The same can be said for telecommunication networks, which prompted in the 90s some industrial organisation studies to introduce network effects in their modelling through the use of a network utility function of a specific form, very often a simple linear function of the aggregate size of the network. In addition to these pure network effects, many industries exhibit significant economies of scale and scope (see [4] for a quantitative estimation of their relevance in the post-trading industry). These effects lead to decreasing average transaction costs as the number of users of a market infrastructure increases. They are also relevant for decisions to join existing infrastructures of previously fragmented markets, or to create new ones.

The reader can refer to Economides [15] for a description of general network effects in an industrial organisation setting, as well as to Economides and Salop [16], and to Park and Ahan [45] for interconnection pricing and joint ownership in network industries. An informal introductory discussion of these network effects in the setting of the post-trade industry, with some comparisons with other network industries, can also be found in Knieps [33].

The first main contribution of the article is to allow the network effect of a given industry to be endogenous to the model by providing explicit micro-fundation for it (Section 2). So far in industrial organisation theory the network effects on agents' utility have been *assumed* rather than *derived* from a model. We propose a completely different approach, based on a simple modelling of the networks involved, which allows to derive the network effects in an endogenous manner instead of assuming an *ex-ante* form of the its utility. Hence, instead of assuming an *ex-ante* special form for the contribution of the network effect in our agents' utility, we derive this additional utility is from the basic assumptions of the model (different costs, profit margins, pre-existing network and overall pattern of trade).

The second main contribution of this paper is to build on the results from Section 2 in order to study agents' decision to join or not a complete network, as opposed to staying in a pre-existing network (Section 3). The section thus illustrates how utility functions derived from network models can be used to reflect on the plausible migrations of agents which lies on different networks. A quantification of agents incentives to join the new network is obtained as a function of the involved networks operational costs, the individual agents' costs per transaction, the agents' adaptation costs, as well as the initial belief held by agents about the number of other agents having already joined the network. In particular, the paper illustrates the possible synchronisation problems faced by agents and the potential need for synchronisation from a given subset of agents in order to achieve a more profitable situation for all agents (Section 3). Policy implications are that more transparency concerning the agents' decision to join the network, as well as more detailed feasibility studies concerning the costs of adapting to the network, may allow to reach more efficient Nash equilibria.

The third main contribution of this paper relates to the determination of stable networks and equilibria in a game of network formation derived from our basic model (Section 4.2). In that sense this paper is

[1] A version of this paper has been submitted in October 2014 to the journal *Managerial and Decision Economics* under the title "Network effects and modelling agents' decision to join a network".

linked to the so-called emerging "network economics" literature, which has been successfully applied over the past decade to describe economic phenomenon as diverse as the social transmission of job information (Calvo-Armengol and Jackson [3]), free-trade agreements (Goyal and Joshi [23]), and co-authorship links (Jackson and Wolinsky [28]). Moreover, this stream of literature contains many articles that are basic models described in a generic setting and could thus potentially be applied to a variety of situations (see Bala and Goyal [1], Belleflamme and Bloch [2], Dutta and al [10], Goyal and Vega-Redondo [24], Jackson [26], [27]). For each trade that can be realised a chain of intermediaries of minimum costs is chosen, and the profits derived from realising the trade along that chain of intermediaries are equally shared by the various market participants involved in the chain. The rules of the new game, which extend the network utility pay-off derived in Section 2, have not yet been studied in the network economics literature. Simple as they seem, they introduce profound topological strategic considerations in the decision of market participants concerning the creation of links with one another. For example, creating a link towards a player which provides access to many other players not directly accessed would generally be attractive because of the plausible intermediation fees profits. Section 4.2 sheds light on such strategic interactions, and the structures that emerge in equilibrium, with a view to understand better the strategic considerations of these complex networks and their implication for the efficiency of the resulting global network structure.

2 Microfundations of network utility

2.1 Graph-theoretic models for networks

2.1.1 Definitions

We use a graph theoretic framework in which agents are represented by the nodes of some (possibly evoluting) network G_t. We denote by $V(G_t)$ the nodes of a given network G_t at time t. The cardinality of a network is by definition the cardinality of its set of nodes, that we often denote it by n. An edge, or link, is any (ordered) pair of nodes (A, B), with A and B in $V(G_t)$. It is denoted AB. Note that the order is important, because AB is different from BA if A and B are not the same node.

At any given period t, any edge AB is either *in the network*, that is, belongs to the set of edges $E(G_t)$ of the network G_t, and in this case we note $AB \in E(G_t)$, or *not in the network*, and we note $AB \notin E(G_t)$. Of course, an edge AB from the node A to the node B represents some (unidirectional) "link" the agent A has established with agent B. For example, if one thinks of the nodes as the Central Securities Depositories (CSD) of the post-trading industry, an edge between A and B could mean that A is an investor agent in B *idem est* holds an account with B that allows it to settle transactions of securities in the domestic market of B. Note the relation does not need to be symmetric, that is, A can have an account with B (and thus be able to settle trades for its clients in the domestic market of B) without the converse being true (B does not need to have an account with A). This is the reason why the pairs are ordered.

As an example, consider the hypothetical network on Figure 1. Note that while some pairs of nodes are linked in both directions (AB and FE), other links are just one directional. *Loops*, that is, edges which come from and arrive at the same node (AA, BB, CC, DD, EE, FF in Figure 1), can symbolise the fact that the agents are indeed able to carry out domestic transactions. For simplicity we will often not represent loops on figures.

Three networks have a particular significance when studying trade or settlement networks: the empty network, which is a network with no link, and generally reflects too high costs for establishing a link compared to the expected profits derived from it, the complete network, which contains all possible links and indicate a strong desire of disintermediation, and the complete star network, where one node, called the *center* of the star, is linked in both directions to all other nodes, which are called the *leaves* of the star. Hence any trade in a complete star network involves the center of the star. The star network has also been called the "hub and spoke" model in the context of the securities settlement industry, see for example [35]. Figure 2 represents these three particular types of networks.

Edges can either be in the network before the start of the simulation, that is, before period 1, or be

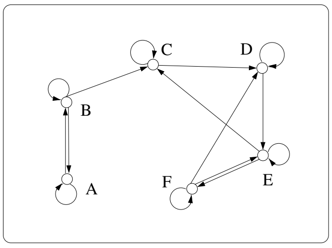


Figure 1: A network with six nodes and 14 edges.

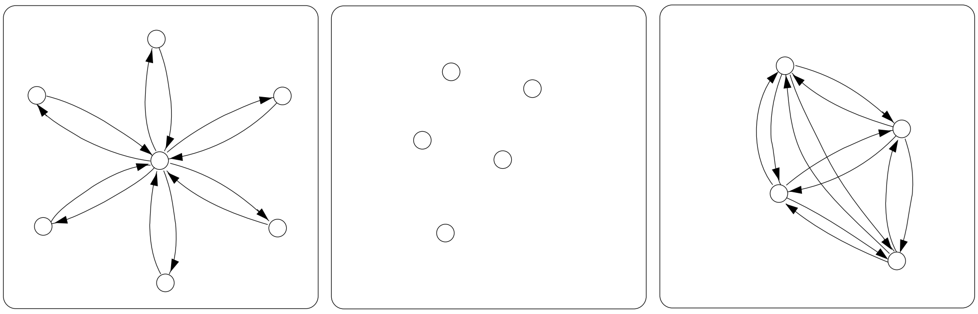


Figure 2: A complete star with 6 leaves, the empty network on 5 nodes and the complete network on 4 nodes.

built after, if the agents expect it to be profitable. At each period t node A can build any edge starting at A not already in the network, at a cost of w^t_{AB}.

When the link AB is in the network at time t_0, that is, was either in the original network or has been created by player A previously, it can be used at any time $t \geq t_0$ to realise the latent profit g^t_{AB} with a probability P^t_{AB}. Using the pre-existing link AB at time t a incurs a cost of c^t_{AB}. We call $g^t_{AB} - c^t_{AB}$ the *net profit*. For player A to build the link AB at time t incurs a fixed, one-off cost of w^t_{AB}. If we denote by E^t_{AB} the expectation of net profits from realising the link AB at time t, then a rational player A should create the link AB at time t if, and only if, $E^t_{AB} > w^t_{AB}$.

We assume the agents are rational investors with perfect information, thus being perfectly able to predict how much gains can be expected from establishing a new link to another agent. We also assume their time frame N is finite and fixed, although we can easily relax this assumption to allow for a moving window of N periods (N represents thus the time-horizon of the agents). Coefficient of actualisation of future gains, that is, discount rates, can be included in the model by simply taking the discounted value for the future gains (resp. costs) g^t_{AB} (resp. w^t_{AB}).

We present two closely related models which correspond to two different situations, more precisely, are the agents examining only the future expected values of profits they could earn by building some edges at each period or can they wait to know that at least one trade can be carried out through the edge they may build. In both models, running and operating the network incurs fixed costs, that we denote by C_{fixed}. Note that by fixed costs we actually mean any deterministic function, which does not depend on the probabilities involved. For example C_{fixed} could be taken to be any function of time, but not a function of the number of newly built edges (which is a random variable).

2.2 Input (for both models)

The input for both models consists of

- a pre-existing network G_0, that is, a set of n nodes (representing agents) of the network G_0, denoted by $V(G_0)$, and the set of edges $E(G_0)$ already in the network at the beginning of period 1,

- the number of periods of the simulation, i.e. the time-horizon of the model N,

- for each period $t = 1...N$, some probability distribution on the edges P^t. It means for any $t = 1...N$ we have both $\sum_{A,B} P^t_{AB} = 1$ and $P^t_{AB} \geq 0$ for all $A, B \in V(G)$.

- for each directed edge AB, and each period $t = 1...N$:

(i) the cost c^t_{AB} of using the pre-existing edge AB at time t,

(ii) the cost w^t_{AB} of creating edge AB at time t,

(iii) the gain g^t_{AB} of realising the transaction that uses the edge AB at time t.

Note that each of the three previous quantities could be a random variable, as soon as we know how to simulate them (for the simulation part) and to express their mean (for the analytical part). For example in the case where the expected profit from the transaction is not a fixed value but a random one, replace g^t_{AB} by the realised value $g^t_{AB}(\omega)$, where ω belongs to the universe Ω, in the algorithms below.

2.2.1 First model: optimisation algorithm when the idea of a trade is known after the possibility of building an edge

The first model can be described as follows: at the beginning of each period t each link (edge) not already formed is considered by the agent, which decides to establish a link (build an edge) if, and only if, the expected additional profits derived from this link exceed the costs of establishing it, *idem est* if, and only if, $E^t_{AB} > w^t_{AB}$. Then an *idea of trade* (arising from one of the agent's client) arises, involving a given edge e_t. If this edge is not in the network at that point, there is no way this trade could be realised. If the edge e_t is in the network (either because it was there right from the beginning or because it was constructed afterwards), then the initiating agent ponders if the costs of using that pre-established link exceeds the benefits it derives from the trade. If it does, the idea of trade is *realised*, the aggregate gains derived from this trade being $g^t_{AB} - c^t_{AB}$, otherwise it isn't.

Algorithm:

For each period t from 1 to N do:

1) For each edge AB not already build, look if $E^t_{AB} \geq w^t_{AB}$.
If $E^t_{AB} \geq w^t_{AB}$ then
  Pay w^t_{AB} to build the edge
Else
  do nothing.

2) Draw an idea of a trade according to the probability distribution P^t. We denote the idea of the trade arising at period t by e_t.

3) Look if edge e_t is in the network.
If e_t is in the network then
  Look if $g^t_{e_t} > c^t_{e_t}$.
  If $g^t_{e_t} > c^t_{e_t}$ then
    Carry out (realise) the transaction and earn a profit of $g^t_{e_t} > c^t_{e_t}$
  Else
    do nothing.
Else
  do nothing.

2.2.2 Second model: optimisation algorithm when the next idea of a trade is known before the possibility of building an edge

The second model differs in that not all non-existing links are considered at each period. On the contrary, the idea of a trade comes first, and there may or may not be a link to allow it to be realised. In case there is no link the agent can build the link if the expected additional profits derived from it (knowing a trade can immediately be realised using this link) exceeds the costs. Whenever a link corresponding to the idea of a trade is in place, the costs associated with using this link are compared to the profits derived from using it and the transaction is made if, and only if, the profits are larger.

Algorithm:

For each period t from 1 to N do:

1) Draw an idea of a trade according to the probability distribution P^t. We denote the idea of the trade arising at period t by e_t.

2) If the edge e_t is already in the network then go to step 4), otherwise continue to step 3).

3) Look if $E^t_{e_t} \geq w^t_{e_t}$ where $E^t_{e_t}$ is the expected gains from future trades involving the edge e_t, including the one that would like to be realised at period t.
If $E^t_{e_t} \geq w^t_{e_t}$ then
  Pay $w^t_{e_t}$ to build the edge
Else
  do nothing.

4) Look if $g^t_{e_t} > c^t_{e_t}$.
If $g^t_{e_t} > c^t_{e_t}$ then
  Carry out (realise) the transaction and earn a profit of $g^t_{e_t} - c^t_{e_t}$
Else
  do nothing.

Remark: for all A, B, E^t_{AB} depends on the time-horizon N.

2.2.3 A simple introductory example of a non-evoluting network

Simplifying assumptions for an analytical resolution Of course, each of the two general models for an evoluting network contains the simpler case of a non-evoluting network: just set all the costs w^t_{AB} for creating new edges higher than the possible profits derived from any single edge, and no edge will ever be built. Note that in such case the two distinct algorithms do exactly the same thing, collapsing in the single model detailed here.

To get a nice closed-form formula for the expected profits of switching to the network we make a number of simplifying assumptions, but we still want to distinguish between domestic trades (those involving edges of type AA, the *loops*), from cross-border trades (those involving edges AB with $A \neq B$).

Let α be the probability for an idea of a trade to be domestic. A natural proxy for α would thus be the ratio of domestic trades over the total number of trades (cross-border and domestic).
Hence, by definition of α:

$$\begin{aligned} P_{AA} &= \frac{\alpha}{n} \\ P_{AB} &= \frac{1-\alpha}{n(n-1)} \end{aligned}$$

Assume also for simplicity that:

$$\begin{aligned} \forall A, B &\in V(G): \quad P_{AA} = P_{BB} \\ \forall A, B, C, D &\in V(G) \;/\; A \neq B,\ C \neq D : P_{AB} = P_{CD} \end{aligned}$$

That is, all domestic trades are equiprobable and this is also the case for all cross-border trades. We will thus often write P_{AA} for the probability of a (any) domestic trade and P_{AB} for the probability of a (any) cross-border trade without stating precisely which are the agents involved. Similarly we suppose costs and gains derived from a transaction to be the same for each type of trades, and we denote by c_{dom} and g_{dom} the costs and gains from a domestic trade and by c_{cross} and g_{cross} the costs and gains from a cross-border trade. We also suppose $c_{cross} < g_{cross}$ and $c_{dom} < g_{dom}$. Concerning the network G we suppose that all loops, that is, edges of type AA, are already there (hence reflecting the fact that every agent is able to carry-out domestic trades), and that there are r other pre-existing edges which are not loops (of type AB with $A \neq B$).

In what follows we derive the utility of evolving networks from the central planner perspective, that is, the aggregate utility. We then apply the formula both to the general case of a evoluting network and to the specific case of a network which is complete, that is contains all the edges, and will be the new network proposed by the central planner to the players. Subtracting the quantities gives the marginal utility of the new proposed network compared to the existing (and evolving) one. This is thus a crucial input to understanding the utility of the new proposed network. We then take the point of view of an individual agent and study the decision of the agents to join or not the new proposed network.

Analytical resolution: a closed-form expression of the network benefits **Some notations**

For any real number r, we denote by $\lfloor r \rfloor$ the floor of r, that is, the largest integer smaller or equal to r.

We denote by Ω the universe and by $\omega \in \Omega$ an element of the universe.

For two sets $\mathcal{A}$ and $\mathcal{B}$ we write $\mathcal{A} \subseteq \mathcal{B}$ if $\mathcal{A}$ is included in $\mathcal{B}$, that is, if all elements of the set $\mathcal{A}$ also belong to $\mathcal{B}$, and in that case we denote by $\overline{\mathcal{A}}$ the complement of $\mathcal{A}$ in $\mathcal{B}$, that is, the set of all elements of $\mathcal{B}$ which are not in $\mathcal{A}$. We denote by $1_{\mathcal{A}}$ the indicator function of some a given set $\mathcal{A}$, ie:

$$1_{\mathcal{A}}(\omega) = \begin{cases} 1 \text{ if } \omega \in \mathcal{A} \\ 0 \text{ otherwise} \end{cases}$$

Analysis

Profits is a random variable that can be written:

$$profits = \sum_{t=1}^{N} (g^t_{e_t} - c^t_{e_t}) 1_{(e_t \in E(G))} - C_{fixed}$$

Remark: if we chose to discount future profits and costs by some discount factor δ^t, it's enough to simply replace all the $g^t_{e_t}$ by $\delta^t g^t_{e_t}$ and all the $c^t_{e_t}$ by $\delta^t c^t_{e_t}$ in the above formula and the results that follow.

Now taking expectation on both sides of the previous equation yields and assuming for simplicity gains and costs do not vary over time, and write $g_{dom} := g^t_{dom}$, etc., we get (see Annex 6):

$$E(profits) = \alpha N(g_{dom} - c_{dom}) + N(1-\alpha)\frac{r}{n(n-1)}(g_{cross} - c_{cross}) - C_{fixed}$$

This general formula of course applies to the particular case of the complete network proposed by the manager or central planner by simply setting $r = n(n-1)$. In that case applying the previous formula, and indexing the different parameters in a straightforward way, we get:

$$E(profits \text{ with the network}) = \alpha N(g_{dom} - c^{comp}_{dom}) + N(1-\alpha)(g_{cross} - c^{comp}_{cross}) - C^{comp}_{fixed}$$

where the index *comp* stands for "complete", meanning the parameters are those of the complete network being introduced by the manager or central planner. By simply subtracting the two equations we thus obtain an expression for the expected marginal utility of replacing the evolving network by the new complete network:

$$\alpha N(c_{dom} - c^{comp}_{dom}) + N(1-\alpha)\left\{(g_{cross} - c^{comp}_{cross}) - \frac{r}{n(n-1)}(g_{cross} - c_{cross})\right\} + C_{fixed} - C^{comp}_{fixed}$$

We discuss its interpretation in the next section.

Discussion on the benefit split-up of switching to a new complete network For discussing the merits or drawbacks of the network it is better to rewrite the previous expression as

$$\alpha N(c_{dom} - c_{dom}^{comp}) + N(1-\alpha)\left\{(1 - \frac{r}{n(n-1)})g_{cross} + \frac{r}{n(n-1)}c_{cross} - c_{cross}^{comp}\right\} + C_{fixed} - C_{fixed}^{comp}$$

The first term represents the benefits from domestic trades. These potential benefits rely solely on costs: if the costs of the complete network for domestic trades are lower than the one of the existing network, the first term will contribute positively to the aggregate marginal utility. This particular fact can point towards the need for the agents to adapt to the network (for example, to make their IT infrastructure compatible to the one operating the network in terms of message formats, etc.), such that c_{dom}^{comp}, which is the sum of the internal costs incurred to the agent in the complete network , will drop below c_{dom}, the costs for domestic trades in the previous network.

In the post-trading industry for example, domestic trades are in general rather efficient in terms of costs, only those agents which reshape their IT structure to fully adapt to the new network will reap the potential benefits from domestic trades indicated by the first term. Alternatively, agents having an inefficient domestic infrastructure for settling their domestic trades will also benefit from this term.

The second term is the most complex and reflects the benefits derived from cross-border trades. The term $\frac{r}{n(n-1)}$ reflects the density of the pre-existing network of bilateral links. The more dense it is (that is, the higher the number of links r), the less benefits are to be expected from latent profits now being able to be realised in the network. At the extreme case, the agents had already created all the links, *idem est* $r = n(n-1)$, and the benefits are only due to costs-reduction, since we end up with the term $c_{cross} - c_{cross}^{comp}$. For sparser and more incomplete pre-existing networks, other benefits stem from new opportunities to access other nodes. The term $\frac{r}{n(n-1)}c_{cross} - c_{cross}^{comp}$, if positive, reflects the decrease in aggregate costs for cross-border trade. Note that this term could be negative, that is, the *aggregate* costs of cross-border trade could actually increase with the adoption of the network. That is, the second term is always positive, indicating the benefits of the network for cross-border trade, benefits derived both from an increase of the cross-border trade being able to be realised and a reduction in unit costs for cross-border trade offered by the platform.

The last term is clearly expected to be negative, since in C_{fixed}^{comp} have to be included all the costs for the whole set of agents to adapt to the network. Nevertheless, these costs are one-off costs, hence temporary in nature. In that case C_{fixed}^{comp} would be best modelled by a decreasing function of time, $C_{fixed}^{comp} = f(t)$. The speed with which $f(t)$ decreases, and the long-run costs it converges to, is ultimately in the hands of the agents themselves. Of course, there are some incompressible costs, such that the operational costs charged by the complete network platform itself once it has amortized its initial investments. Nonetheless, by outsourcing their trade service to the operators of the complete network, economies of scale can be expected, which would make the limit of $f(t)$ lower than C_{fixed}. But this is conditional to the agent fully adapting to the complete network architecture and hence not duplicating the operational and back-office costs. In this article we do not model the relatively important questions or whether or not these one-off adaptation costs could be passed on to the agent final clients or not. For such a modelling in the context of the post-trading industry, the interested reader can refer to [40].

The more general case of evolving networks,albeit with homogeneity assumptions concerning the parameters and the probability distribution for realisation of profits, are treated in Annex 6.2 which provides similar results with similar interpretations.

Remark: There are several ways in which the model could be extended:

1) An expected effect of the introduction of the new complete network could be an increase of competition among agents, whereas the competition effect is not modelled at all here: we have not looked at individual agents' profits but only at their aggregate profits, answering the question: does it benefit the agent community *as a whole* to join the network? Nevertheless one can easily compute individual agent profits in the model, and see which one benefited the most, depending on its initial position in the network. Still, another scheme would need to be added to factor in competition among agents. We could

for example introduce competition by making at each steps different types of complementary signals on the nodes indicating possible trades, and the agents which have a signal complementary to a same other agent signal compete for the executing this trade (choose the winner by using a probability rule or by executing the least costly trade, or the trade with higher pay-off, or by choosing the ones that can lower their fees $g > c$ the most).

2) Another idea is to add a constraint in terms of the budget of the agents, reflecting that due to limited funding possibilities, a given agent might not be able to establish a link even though it would like to do so. This would introduce additional benefits derived from the network in the form of pooling of resources for the agent community.

3) For each agent there might be a majority of domestic versus cross-border trades (or alternatively the nodes could represent the clients of the agents and the costs and gains the clients' costs and gains). Introducing these two (or more) "types" of nodes would allow to access the impact of the post-complete phase in terms of benefits for the internationally-oriented agent as compared to the more domestic ones.

4) We have been working under the most favourable assumptions for the profit-maximisation of agents, that is, perfect information and rational agents. One may want to investigate further the effect of changing to the network assuming bounded rationality or imperfect information for the agents.

3 To join or not to join? Giving the agents the choice

We assumed a pre-existing evolving network and a new complete network proposed by the central planner or its manager and derived its marginal utility when *all* the agents join the network in the post-complete phase. Since participation in a network is usually voluntary, some agents may choose to stay outside the complete network. That is, they prefer to continue to benefit solely from the previously established links and those they can possibly build themselves afterwards, but will not contact any other agent through the newly available complete network. Using the results and modelling of the previous sections we will now show that it is easy to model a post-complete network phase in which some, but not all, of agents have joined the complete network. We then look at the benefits of the new network conditionally to the number of agents joining the network. Here too, the network effect is endogenous, contrary to most network models found in the literature where one has to assert a special form for the additional utility derived from the network (generally a proportional factor to the size of the network the agent can reach). With these benefits of a non-complete network derived, we will endogenize agent's decision to join or not the network. Agents decision will depend of the efficiency of both current and new network, as well as their believes about other agents' intentions. We will characterise the set of all Nash-equilibria as a function of our model parameters.

3.1 Assumptions

In any of the two previous models it is straightforward to derive each agent *individual* profits associated to a given network. If agents are rational investors with perfect foresight, their decision to join or not should be made by computing their expected marginal utility of joining the new network.

We make two important assumptions in what follows: first, that the agents that have joined the complete network do not price-discriminate towards those who have not, and thus continue to trade with them as before, and vice-versa. Hence, using the old links between a agent inside the network and another outside the network is still possible and comes with the same costs as before, as it uses the old, pre-existing network.

We also make two (unnecessary) assumptions in order to keep things as simple as possible: the first assumption is that the fixed operational costs of agents that have joined are higher than before because they still fully bear the fixed costs of operating the old network in order to be able to carry trade with those agents that have chosen to stay outside the network. More precisely we assume the fixed costs of those agents having joined the network can be written as the sum of the fixed costs of the complete network and of the old pre-complete network. The effect of this assumption is certainly an overstatement of the actual costs of the network: with less agents operating in the old network, there is certainly some room for the reduction of the operational costs of the agents having joined the network. Our assumptions, only made for convenience, amounts to assuming that the agents having joined basically bear the full costs of operating both the old network (now containing less nodes) and the new complete network (which has for size the number of nodes having joined). Relaxing this assumption, assuming, for example, a strict decrease in the operational costs of operating the old network for those agents that have joined the network, can easily be done, but then a precise form of the parameter C^A_{fixed} has to be assumed to analytically solve the model.[2]

The second simplifying assumption is that we will not distinguish between domestic and cross-border trade. Hence $P_{AB} = 1/n^2$ is the probability of an idea of a trade appearing on any given edge AB, for A, B in $V(G)$ (hence AB may be a loop), and the costs of using a loop is the same as of using an edge which is not a loop.

Concerning the other parameters c^t_{AB}, g^t_{AB}, and w^t_{AB} of the model we assume them constant other time and homogenous across the edges, in line with Sections 6.2.2 and 6.2.3.

3.2 Definition of the game

Now select (once for all) one of model 1 or 2, and consider the following simple game:

1) Each agent decides whether to join the complete network or not.

2) Each agent receives the associated payoff (the individual payoff as implied by the previous section concerning evoluting networks).

Although in the second step randomness is involved, this still results only in a simple static game of complete and perfect information, since the rational, non risk-adverse, and profit-maximizing agents will just act in step 1 such as to maximize their *expected* payoff in stage 2, and these expected payoffs are perfectly known and computable.

We will select, for example, the first model to illustrate. A similar analysis could of course be conducted by selecting the second model.

3.3 Analytical resolution

The expected gains from a given agent A in a general evoluting network G with no edge at period $t = 1$ is the sum of the expected gains from trades carried out through built edges of type AB with $B \in V(G)$, minus fixed operational costs:

$$E^1_A = \sum_{B \in V(G)} (E^1_{AB}(c_{AB}) - w_{AB}) 1_{(E^1_{AB}(c_{AB}) \geq w_{AB})} - C^A_{fixed}$$

with

$$E^1_{AB}(c_{AB}) := N(g_{AB} - c_{AB}) P_{AB}$$

By the simplifying assumptions of the model we get a general expression for the expected profits of an individual agent A within an evoluting network with n nodes:

$$E^{out}_A(n) := E^1_A = n(E^1_{AB}(c_{AB}) - w_{AB}) 1_{(E^1_{AB}(c_{AB}) \geq w_{AB})} - C^A_{fixed}$$

[2] Equation (C) (see later) may then be a polynomial of degree higher than 2, depending on the form of dependency of C^A_{fixed} on n_0 assumed.

To derive a similar expression for the complete network it is enough to use the previous formula with $w_{AB} = 0$ for all edge AB, as all edges between all vertices are provided by the complete network at no cost:

$$E_A^{comp}(n) = nE_{AB}^1(c_{AB}^{comp}) - C_{fixed}^{A,comp}$$

Moreover, the fixed costs of operating the complete network as a whole over the amortization period can be decomposed into two components: the portion of the fixed, incompressible, operational costs of the complete network platform that are passed on to the agents joining the new network (recovery fees) and, possibly, the margins set by the central planner or the manager of the new complete network; the sum of these two costs we denot by c and we assume equally shared by the n network participants joining the new network. Hence each of the n participants joining bears a cost of c/n. Moreover, each agent A bears fully its own individual adaptation costs c^A. Hence $C_{fixed}^{A,comp} = c/n + c^A$ and we have:

$$E_A^{comp}(n) = nE_{AB}^1(c_{AB}^{comp}) - c/n - c^A$$

Note in passing that, if we want to make the link with the aggregate constant C_{fixed}^{comp} of the previous sections, we would write:

$$C_{fixed}^{comp} = \sum_{A \in V(G)} C_{fixed}^{A,comp} = c + \sum_{A \in V(G)} c^A$$

The expected gains of joining the network for an individual agent holding the belief that $n_0 - 1$ agents other than himself will join the network is simply the sum of:

1) the gains $E_A^{full\,comp}(n_0)$ from settling transactions with the n_0 agents having joined the network, including itself,

2) the gains $E_A^{out}(n - n_0)$ from settling transactions with the $n - n_0$ agents still out of it.

Figure 3 illustrates the superposition of the two networks used to describe the post-complete hybrid environment: the gains of agent A having joined the complete network can be split into the gains from trade with other agents which also joined the new complete network (through the links inside the circle), which amount to $E_A^{full\,comp}(n_0)$, and the gains with agents staying outside the network, using the old network (through the links that cross the circle). Now notice that because of the assumption of spatial homogeneity across parameters, agent A profits arising from trades with agents outside the network are the same as the profits of any agent B outside the network that arise from trade with agents outside the network other than B itself, plus the profits arising from idea of trades of type (A, B). Again, by homogeneity of parameters the profits arising from idea of trade of type (A, B) are the same as those arising from idea of trade (B, B), hence $E_B^{out}(n - n_0)$ (which is also $E_A^{out}(n - n_0)$) is indeed the profits of agent A arising from settling transactions with the $n - n_0$ agents outside the network.

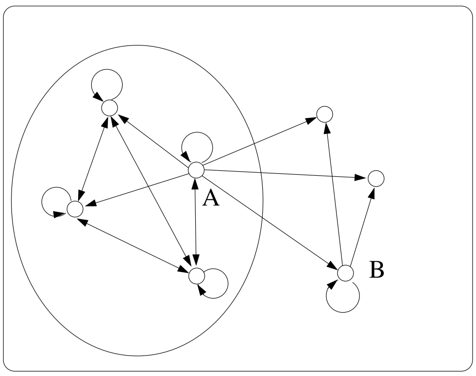


Figure 3:

These gains have to be compared with the expected gains from not joining the new complete network, also holding the same belief that $n_0 - 1$ agents will join the network. This quantity is simply $E_A^{out}(n)$ and does not depend on n_0.

Thus a given agent A holding expectations that $n_0 - 1$ agent different from himself will join the network would join himself the network if, and only if,

$$E_A^{comp}(n_0) + E_A^{out}(n - n_0) > E_A^{out}(n)$$

Note this also uses the previously stated assumption that the fixed operational costs of agents joining the network will be the sum of C_{fixed}^A and of $C_{fixed}^{A,comp}$ (left hand-side of the inequality). This is because C_{fixed}^A is accounted for in $E_A^{full\,comp}(n_0)$ while $C_{fixed}^{A,comp}$ is accounted for in $E_A^{out}(n_0)$. The previous inequality can be re-written as:

$$E_A^{comp}(n_0) - E_A^{out}(n_0) - C_{fixed}^A > 0$$

idem est:

$$n_0\{E_{AB}^1(c_{AB}^{comp}) - (E_{AB}^1(c_{AB}) - w_{AB})1_{(E_{AB}^1(c_{AB}) \geq w_{AB})}\} - c/n_0 - c^A > 0$$

Setting $z := E_{AB}^1(c_{AB}^{comp}) - (E_{AB}^1(c_{AB}) - w_{AB})1_{(E_{AB}^1(c_{AB}) \geq w_{AB})}$ we get

$$n_0 z - c/n_0 - c^A > 0 \quad (C)$$

Let $f(x) := xz - c/x - c^A$. Then $f'(x) := z + c/x^2$ is always non-negative so f is increasing in x. The positive real number n_l^A such that $f(n_l^A) = 0$ is

$$n_l^A = \frac{c^A + \sqrt{(c^A)^2 + 4zc}}{2z}$$

Hence if $n_0 \geq n_l^A$ then it is worth joining the complete network for agent A. Since every agent A can carry out this computation of n_l^A, which is independent from n_0, and assume the others do the same, at equilibrium n_0 is necessarily the number of agents A such that $n_l^A \leq n_0$.

Note that n_l^A is increasing both with c^A and with c: the higher the individual agent adaptation costs or the complete network platform costs are, the more volume n_l^A is required to make the complete network service competitive enough for agent A to decide to join. Note that z can be re-written as:

$$z := (NP_{AB}(c_{AB} - c_{AB}^{comp}) + w_{AB})1_{(E_{AB}^1(c_{AB}) \geq w_{AB})} + NP_{AB}(g_{AB} - c_{AB}^{comp})1_{(E_{AB}^1(c_{AB}) < w_{AB})}$$

Hence z is nothing else than an efficiency indicator of the network. Indeed, the first term quantifies the gains from the reduction of costs for those trades that were already realised in the old network, and can be further decomposed into two components: gains realised by reducing unit-costs for trade, $NP_{AB}(c_{AB} - c_{AB}^{comp})$, and gains from not having to build any further link for settling new transactions, w_{AB}. The second term represents the gains from the trade that were missed opportunities before the network due to the absence of links. Unsurprisingly, the first derivative of n_l^A with respect to z is negative, that is, the minimal number of agents joining the network required to make the platform attractive drops as the efficiency indicator z increases.

Remark: it can be shown that the case where the manager or central planner chooses not to pass on its costs c explicitely but includes them in the costs charged for using edges to realise latent profit is similar (Annex 6.3).

The theorem below characterises all the Nash equilibria of the game:

Theorem 1 *Nash equilibria are precisely the strategies for which, if n_0 denotes the number of participants joining the new complete network, then $n_0 = \#\{A \in V(G) \,/\, n_l^A \leq n_0\}$ and each participant A joins the new complete network if, and only if, $n_l^A \leq n_0$.*

Proof. Let a given strategy be a Nash equilibrium. Assume an agent A has chosen to join the network if and that $n_l^A > n_0$. Then by the above A would be better off not joining the network, a contradiction with the definition of a Nash equilibrium. Conversely, the same reasoning shows that these conditions are sufficient for the profile to be a Nash equilibrium: assume the number of agents having joined the network is n_0 and that those agents having joined are precisely those agents A such that $n_l^A \leq n_0$. Then there is no incentive for any agent to deviate unilaterally from playing its current strategy. ■

There can clearly be different Nash equilibria in our model. For example, since $n_l^A > 0$ by definition, the situation in which no agent join the network is always an equilibrium, since for $n_0 = 0$ we do have that the set $\{A \in V(G) \,/\, n_l^A \leq n_0\}$ is empty. At the other extreme, if we assume $\max_{A \in V(G)}(n_l^A) \leq n_0$, the situation where everyone join the network is also a Nash equilibrium. In-between there may be a large number of Nash equilibria, or none, depending on the position of the subset of real numbers $\{n_l^A,\ A \in V(G)\}$ with respect to the set of natural numbers.[3]

3.4 A simple example

Figure 4 illustrates the three different Nash equilibria for a game with three agents A, B, and C, and $n_l^A < 1$, $2 < n_l^B \leq n_l^C < 3$: the values of n_0 corresponding to each of these equilibria are circled. Note than the integer 2 does not correspond to any possible value of n_0 in a Nash equilibrium, since $\#\{A \in V(G) \,/\, n_l^A \leq 2\} = 1 \neq 2$.

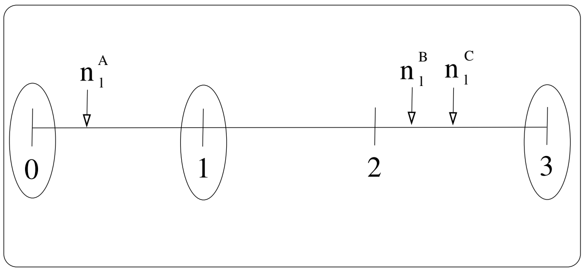


Figure 4: Three distinct Nash equilibria

Note we implicitly assume that the agents all hold a *common* belief n_0 about the number of agents joining the network. Since many possible n_0 can yield a Nash equilibrium, this is a non-trivial assumption: it would be perfectly rational for different agents to hold *different* beliefs about n_0, as soon as these beliefs correspond indeed to some Nash equilibrium.

As an example, consider the extreme case where $n - 1 < n_l^A < n$ for all agent A. Assume all agents hold the same belief. Then depending on this belief either nobody joins (if $n_l^A > n_0$), or everybody joins (if $n_l^A < n_0$). Now if the beliefs are heterogeneous, using that $n - 1 < n_l^A$ we can see that if at least one agent believes that the number of agents joining will be less than n, that is, if only one agent assume that another agent will not join, then it will not join itself, and that this will cause agents deciding to join to select a non-profit maximizing solution for themselves.

3.5 Discussion and implications

The previous simple example, as well as considerations on the dynamics of the repeated game associated with the previous one stage game described (Annex 6.4), both illustrate the importance of a given agent:

1) knowledge of its own adaptation costs and of the adaptation costs of others (since c^A determines n_l^A),

2) belief about other agents' beliefs.

[3] It could thus be said that, for most parameter constellations, the concept of Nash equilibrium alone does not allow to "solve " the model, as there may be more than one Nash equilibrium. Instead, the concept provides us with a *variety* of different plausible outcomes.

Indeed both factors will contribute to shape a agent's decision to join the network or not.

Now one could ask what would happen if the game was not of perfect information concerning the individual agents' adaptation costs. After all, it is not even certain that a given agent knows its own adaptation costs, least the adaptation costs of others. If there is uncertainty about the other agents adaptation costs, then the belief n_0 of the number of agents joining might actually not be true. Probably, there will not be heterogeneous beliefs anymore[4]. Similarly, even in a game of perfect information, if there is no consensus about n_0, the number of agents assumed to be joining the network, then to decide to join each individual agent A will compare its (individual) n_l^A to its (individual) belief n_0^A. The resulting solution cannot be guaranteed to be even a Nash equilibrium, even less one that maximizes social welfare, understood here as the sum of agents profits. The policy implications are that any reduction of uncertainty about the individual agents' adaptation costs c^A, the complete network operational costs c and the number n_0 of agents which plan to join the platform improves aggregate welfare by diminishing the noise and the possibility of mis-judgement of the effectiveness of the network by the agents. Reducing uncertainty can be obtained by continuous discussion among agents, allowing for a common belief about n_0 to emerge, and by disclosure from the people in charge of establishing and operating the complete network (that is, the manager or centralplanner) about its operational costs c, and of its technical specificities, such that agents could better guess their own adaptation costs c^A, knowing what they will need to change in their own technical system to become compatible with the complete network platform.

4 Extension of the model to longer intermediary chains

4.1 Introduction

4.1.1 Motivation

Our previous network model could be used unchanged to model the benefits of the network on the set of intermediaries by re-interpreting the nodes as final investors and the links as the cheapest chains of intermediaries allowing them to realise a latent profit. The cost of the link would thus be the overall costs of the chain of intermediaries involved. Nevertheless, this modelling suffers from many flaws:

- it can only reflect the overall picture, as it fails to model the incentives of the different intermediaries to build links *between themselves* to realise latent profits, as a single edge (A, B) represents a whole chain of intermediaries from A to B;

- if we interpret the costs of using an edge of the network as the sum of the costs of the intermediaries involved, this implies intermediaries do not make any profit from their intermediary function, a rather unrealistic assumption. The way to interpret costs in this new setting would be the costs *to the end investors*, that is, the cost component would include the margin of the intermediaries involved. Nevertheless, this constitutes a problem when we would like to allow the final investor (as modelled in the network) to himself provide intermediary services to other players, since those intermediation profits cannot be captured by the model;

- there is no way to get any precise idea of the dynamics of the costs involved, because nothing is known about the length of the custodian chains and the different intermediation services used to carry out

[4] Using the concept of *Bayesian* Nash equilibrium, where the beliefs of market participants are elevated to the role of strategies, does not seem to bring additionnal value to our discussion. Indeed, assuming for example believes of agent i for the adaptation costs of agent j is absolutely continuous with respect to the Lebesgue measure and has density f_j^i, we can compute, in a Bayesian equilibrium, the expected number of agents joining the network, from the point of view of agent i. It is just $E_i(n_0) = \#\{j \in V(G) \,/\, E_j(n_0) \geq E_i(n_l^j)\}$. Indeed, by the same reasonning as previously, the strategy for any agent j is to join if, and only if, $E_j(n_0) \geq n_l^j$. In an equilibrium agent i knows agent j beliefs and is thus able to compute $E_j(n_0)$. Because it does not know, contrary to agent j, agent j 's adaptation costs, it computes $E_i(n_l^j)$ using its own belief f_j^i about its competitor costs. Consistency of beliefs also implies that $E_i(n_0)$ solves

$E_i(n_0) = \#\{j \in V(G) \,/\, E_i(n_l^j) \leq E_i(n_0)\}$. Hence, believes that does not make the above equality true are non-rational and do not result in a Bayesian Nash equilibrium. Conversely it is easy to check that any strategy profile satisfying the above is a Bayesian Nash equilibrium. Hence, compared to the simple Nash equilibrium concept, the Bayesian Nash equilibrium only impose a rationality restriction on the beliefs.

a trade. Furthermore, solving the model analytically as was done before assumes all cheapest custodian chains have the same costs, whatever their length, a rather unrealistic assumption.

We develop an alternative model to tackle this issue by:

1) writing down a static version of our first model where the N periods are condensed into a single-period decision game.
2) altering the rules of the payoff to allow each of the two nodes involved to derive profits from any transaction it participates to settle.
3) generalising these payoffs rules to allow more than two intermediaries, and each of the intermediary to derive profits from helping to realise a latent profit.

We then discuss Nash equilibrium, and a closely related concept of stability, as well as efficiency, of this new model.

4.1.2 A static model of the first model

Assume all parameters are constant across time and space (an hypothesis we will refer to as time and space homogeneity). The payoff for agent A to build edge (A, B) is $E^1_{AB} - w^1_{AB} = E^1_{AB} - w_{AB}$. Since in the first model $E^1_{AB} \geq E^t_{AB}$ for all t in $1...N$, and agents do not have any budget constraints, the optimal choice is to build all the links that will be built at the first period of the model. Hence our previous N-period dynamic model collapses into the (simplistic) one period game:

1) Each player (simultaneously) choose to establish a set of links.
2) They earn the associated payoffs as implied by the first model.

Players play as to maximize their *expected* payoffs.

Remarks:
- In 1) the decision rules are rather simple: player A build edge AB if, and only if, $E^1_{AB} \geq w_{AB}$. In 2), it earns $\pi_A = \sum\limits_{B \in V(G)-\{A\}} (E^1_{AB} - w_{AB}) 1_{(E^1_{AB} \geq w_{AB})}$.
- Since we assume all agents of the model are able to settle "domestic" trades, there is no reason to include the profits derived from these trades in the payoff function associated to a given network. This is the reason why domestic trades do not enter the definition of the payoff function π_A.

4.1.3 Making intermediation profitable

Sharing benefits derived from trade in the previous static two intermediary model Now notice that in the previous model player B earns nothing when a trade is settled using edge AB, whereas B is, as A, an intermediary for the realisation of the idea of trade (A, B). Since we would like that all intermediaries participating in the trade of a trade (A, B), including the "final" intermediary B, derive some profit from their intermediation services, we modify the payoff rules such that both A and B equally shares the profits from settling the trade (A, B). Hence, realising the idea of trade (A, B) using the edge AB in the network incurs a profit of $E^1_{AB}/2$ both for A and for B. Of course, this also modifies player A best-response function which becomes: build a link if, and only if, $E^1_{AB}/2 \geq w_{AB}$.[5]

[5] Note also that holding the whole set of parameters constant, this new repartition of the (same) profits derived from a given transaction decreases the number of links built and the aggregate welfare. This comes from the asymmetry between player A, which bears the full costs w_{AB} of establishing the link, with player B, which bears none but still share the profits. Could this asymmetry be unrealistic considering the fact that player A and B could bilaterally negotiate to share the burden w_{AB} of building a new link? Asymmetry of information between player A, which knows its clients (and knows they want to invest in B domestic market), and player B, which only knows player A, could certainly help to explain this asymmetry: only player A is able to evaluate E^1_{AB} properly, because it knows the demand for assets of the domestic market of player B steaming from his own domestic market. player B, not able to evaluate this demand itself, might want to avoid taking the risk that player A was wrong. More fundamentally, sharing the costs of establishing a link could be assumed, and would give rise to a different game, probably also very interesting by its own.

Contrary to the costs w_{AB} of *establishing* the link AB, the costs of *using* the link AB is equally shared by A and B, because it is already included in $E^1_{AB} = NP_{AB}(g_{AB} - c_{AB})$, the profits from settling the trade (A, B) they will share. We define

$$g'_{AB} := NP_{AB}g_{AB}$$

and

$$c'_{AB} := NP_{AB}c_{AB}$$

We thus have $E^1_{AB} = NP_{AB}g_{AB} - NP_{AB}c_{AB} = g'_{AB} - c'_{AB}$ for the total profit, which will be equally shared between both A and B. Hence the derived profit for A and B is:

$$\frac{g'_{AB} - c'_{AB}}{2}$$

In allowing more than the two final intermediaries to participate in a given trade, we can choose to allow for three intermediaries, or more, or any number of intermediaries, insofar the trade is still profitable. We will consider now the case of three intermediaries and of an arbitrary length for the chain of intermediaries.

Allowing for a maximum of three intermediaries... or more We now want to allow for up to three intermediaries. Hence, we change the rules of the game above to allow an idea of trade (A, B) to be realised if, and only if, the edge (A, B) is in the network, or if there exists a node C such that both edges AC and CB are in the network. If both situation occurs, then the trade is settled using the chain of intermediaries (or path) of least cost, that is, the chain AB if $c'_{AB} < c'_{AC} + c'_{CA}$, and the chain ACB if $c'_{AB} > c'_{AC} + c'_{CA}$, where $c'_{AB} := NP_{AB}c_{AB}$ is the cost of using a given edge (A, B) in this new game. If both chains incur the same costs, *idem est* if $c'_{AB} = c'_{AC} + c'_{CA}$, then one of the two chains AB or ACB is selected randomly to settle the trade, both chain having the same probability to be chosen. The gain is split equally between each intermediary of the chain selected. As we denote by g'_{AB} the profit derived from settling transaction (A, B) in the static game, that is, $g'_{AB} = NP_{AB}g_{AB}$, and a chain of intermediary ABC is used, the profit of each of the nodes A, B and C for acting as intermediaries for settling the transaction (A, B) is

$$\frac{g'_{AB} - (c'_{AB} + c'_{CA})}{3}$$

If the same transaction was settled using a direct link between A and B, the derived profit for A and B is:

$$\frac{g'_{AB} - c'_{AB}}{2}$$

This is consistent with the rule given in Section 4.1.3, where only two intermediaries are allowed. Nevertheless solving the game is now far from being a simple one-parameter profit-maximisation problem, since into the decision-making process of the nodes enter strategic considerations that take into account the overall structure of the network, which in turn depends on the other players' actions. To give a very simple example, in both the three intermediaries and the arbitrary number of intermediaries model it might be that althought $E^1_{AB}/2 > w_{AB}$, a node A will choose not to build the edge AB because it has some intermediary C that already allows him to get enough profits from trades of type (A, B). It might also be the case that it wants to connect directly to node B because node B is connected to many other nodes and thus increase its intermediation profits when compared to settling through C. Node A and C might also be in competition as intermediaries, particularly when their set of neighbours overlap.

We use in the next section exactly the same idea to define the rules of a payoff where any trade (A, B) can be realised insofar there exists a chain of intermediaries from A to B in the network not exceeding a given fixed length and the sum of their costs do not exceed the latent profit from realising the trade (A, B). Some theoretical justification of this equal profit splitting allocation rule can be found in [49].

Indeed, in the author model of stochastic bargaining involving different chains of intermediaries, when the discount factor of the bargaining model tends to 0, that is, when players do not have sufficient time to bargain over the splitting of profits without missing the trade opportunity – and thus become very impatient – the expected payoff of each intermediary become equal.[6] We are taking this view here; the opposite view would be to assume agents have plenty of time to bargain before moving on to carry out the trade; in such a case, according to [49], intermediaries which are not "essential" – meaning there exits, in the network, another chain of intermediaries to which they do not belong – would earn a profit of zero.

4.2 Studying the static game $\mathcal{G}^n_{k_0}(w, c, g)$

We formalise the previous ideas and define the following game on networks.

Notation 2 *For simplicity of notations, we will from now on write g_{AB} for g'_{AB}, w_{AB} for w'_{AB} and c_{AB} for c'_{AB}. When there is space homogeneity of these parameters, we also drop the index, writing, for example, c for c_{AB}.*

Definition 3 *A path is a network P whose nodes can be labelled $V(P) = \{u_0, u_1, u_2, ..., u_k\}$ such that its set of edges is precisely $E(P) = \{(u_0, u_1), (u_1, u_2), ..., (u_{k-1}, u_k)\}$. We often write $P = u_0, u_1, u_2, ..., u_k$ in this case, and say that P is a u_0u_k-path, or a path from u_0 to u_k. The length of P is k, while its cardinality is $k+1$.*

Note that our notion of path thus have a direction: a path from u to v is not a path from v to u, at least when $u \neq v$.

Definition 4 *Let G and H be two networks. Then H is a subnetwork of G if, and only if, $V(H) \subseteq V(G)$ and $E(H) \subseteq E(G)$.*

Definition 5 *A path of a network G is a path that is a subnetwork of G.*

Let n and k_0 be two positive integers, N a set of node of cardinality n and w, c, and g real-valued functions defined on the cartesian product $N^2 = N \times N$. Players are the nodes in N and decide to link, or not, to other nodes. Hence the action set of a given player u is the set of edges (u, v) for $v \in N$, $v \neq u$. If node u establishes the link (u, v) to node v it has to bear a cost of $w(u, v)$. Some paths in the resulting network are used to carry potential trades that benefit all the players along those paths. More precisely, for each ordered pair (u, v) there is the potential for a trade of value $g(u, v)$ to be realised. For the trade

[6] It is of course not necessary to make intermediation profitable to all intermediaries to extend the model to more than two intermediaries. Defining the profit rule such that the first intermediary on a chain of intermediaries derives all the benefit from a given trade gives rise to a simpler, yet also interesting, problem. In such a setting, where node A earns the whole profit derived from settling the trade of $g - kw$ when at the beginning of an intermediation chain of length k, one can easily prove that:

1) if $w \geq g-c$ then the empty network is a NE, and that it is the only NE if we assume further that $w \geq g-c+(g-2c)(n-2)$
2) the complete network is the only strict NE if, and only if, $c > w$.
3) a directed cycle is a NE if, and only if, $g - c + g - 2c + ... + g - (n-1)c > w$.
3) the star is a NE if, and only if, $c \leq w \leq g - c$
4) If $-w + g - c > 0$ then any NE is strongly connected, with diameter $\Delta < w/c + 1$
5) In a NE having more than two components, all vertices earn a 0 payoff.

Point 5) involves a very easy argument. Indeed, assumes G is a NE with at least two different components, with A and B two nodes belonging to different components of G and with B earning a non-zero payoff $\pi_B > 0$. Because A is not in the same component of B it does not derive any profits from trades with the vertices B is trading with. A can thus strictly increase its profit by π_B by simply building an edge to each of the out-neighboors of B, a contradiction with NE. Notice this argument is not valid anymore if all intermediaries share profits, since the higher payoff of B maybe due to its strategic intermediation role between vertices of its own component, and A cannot replicate B role in that matter as it cannot replicate the edges CA for C in-neighboor of B. It is this more complex environment that we investigate in the next section.

to be realised there must be a path of length less than k_0 in the network. A cost $c(u,v)$ is associated to each edge (u,v) of the network. The cost of a path is simply the sum of the costs of all its edges. For each pair $(u,v) \in N^2$ a uv-path of minimum cost can be selected for the trade if it is an *admissible* path, that is, a path of length not more than k_0 and whose costs do not exceed $g(u,v)$[7]. If more than one such admissible path of minimum cost exist then the path used to perform the trade is selected randomly among all admissible paths of minimum cost, the uniform probability distribution being used. Each node of the selected admissible uv-path of minimum cost $u = u_1$, u_2,..., $u_k = v$ receives an equal share of the aggregate net profit $g_{u_1u_k} - (c_{u_1u_2} + c_{u_2u_3} + ... + c_{u_{k-1}u_k})$ derived from the trade, that is:

$$\frac{g_{u_1u_k} - (c_{u_1u_2} + c_{u_2u_3} + ... + c_{u_{k-1}u_k})}{k}$$

The player u_i are called the *intermediaries* of the trade. The two players u_0 and u_k are its *final* intermediaries, while players u_2,..., u_{k-1} are its *strict* intermediaries. Players build edges such as to maximise their expected payoff function.

In what follow we will assume for simplicity that w, c, and g are constant functions. Hence the above profit simply becomes

$$\frac{g - (k-1)c}{k} =: \pi_k$$

To write down players' payoffs explicitly under this hypothesis it is useful to define a minimum path, as minimum path correspond to path of minimum costs:

Definition 6 *A minimum uv-path is a path from u to v of minimum cardinality. The set of all minimum uv-paths is denoted by $\mathcal{P}_{uv}$, and the cardinality of any minimum uv-path by n_{uv}. For any positive integer k, the subset $\mathcal{P}_{uv}^{\leq k}$ of $\mathcal{P}_{uv}$ is the set of minimum uv-path of cardinality less or equal than k. We also set $\mathcal{P}_{uv}^{\leq\infty} := \mathcal{P}_{uv}$ for convenience.*

Definition 7 *A minimum path is a path such that there exists two nodes u and v such that it is a minimum AB-path. We denote by $\mathcal{P}$ the set of all minimum paths of the network. Hence, $\mathcal{P} = \bigcup_{\{u,v\}\subseteq V(G)} \mathcal{P}_{uv}$.*

Since the cost function is constant, the set $\mathcal{P}_{uv}$ of minimum uv-path is precisely the set of paths of minimum costs, from u to v and the set $\mathcal{P}_{uv}^{\leq k_0}$ of minimum uv-path of cardinality less or equal than k_0 is precisely the set of admissible AB-paths. The full expression of the payoff function for agent i of the network in the game $\mathcal{G}_{k_0}^n(w,c,g)$ can then be expressed as:

$$\pi_i = \sum_{\substack{u,v \in V(G)/ \\ \#\mathcal{P}_{uv}^{\leq k_0} > 0}} \left(\frac{g - (n_{uv}-1)c}{n_{uv}}\right) \frac{\#\{P \in \mathcal{P}_{uv}^{\leq k_0} \quad | \quad i \in P\}}{\#\mathcal{P}_{uv}^{\leq k_0}} 1_{(g > (n_{uv}-1)c)}$$

We will always denote by k_0 the maimum number of intermediaries allowed in the rule of the game, and set $k_0 = +\infty$ if chains of intermediaries of arbitrary length are allowed. Note that for any given game with n players, allowing an infinite number of intermediaries is equivalent to allowing at most n intermediaries. That is, for all n, we have: $\mathcal{G}_{\infty}^n(w,c,g) = \mathcal{G}_n^n(w,c,g)$. Also note that in the formula giving agent u's profits we always have $n_{uv} \in \{1, 2, ..., k_0\}$ for all pairs of nodes (A,B) which appear in the sum. Finally, because paths of cardinality k such that $g < (k-1)c$ are not admissible, we have that for all n, $\mathcal{G}_{k_0}^n(w,c,g) = \mathcal{G}_{\min(k_0, \frac{g}{c}+1)}^n(w,c,g)$. Hence we can assume without loss of generality that $k_0 \leq \frac{g}{c} + 1$. This allow to drop the factor $1_{(g > (n_{AB}-1)c)}$ in the expression of π_i.

[7] Deleting this assumption would result in a different still interesting problem, where the nodes of the network are actually *obliged* to perform the transaction, even in case they take a net loss.

Definition 8 *A Nash equilibrium (resp. a strict Nash equilibrium) is a strategy profile G in which any agent deviating unilaterally would not be better off (resp. would be worse off).*

Because this payoff function takes into account the overall structure of the network by making the set of all minimum paths containing u a pre-requisite for computing u's profits, it can be very difficult to characterise all nodes' strategies that result in a Nash equilibrium. Indeed, because from a given strategy a unilateral deviation from a given player consists in building or deleting edges starting at u at the same time, thus reshaping the set of all minimum paths in a way difficult to predict, we will often restrict ourselves to the more general notion of stability:

Definition 9 *A network G is stable (resp. strictly stable), if, and only if, adding or deleting any single edge (u, v) would not increase agent u's profits (resp. strictly increase agent u's profits).*

This would translate the idea that given an initial network, each node in turn only considers one edge at the time, computing whether it would be better off by severing it if this edge is in the network or by creating it if it is not in the network. In a stable network no node would add or delete an edge.

Proposition 10 *If G is a Nash equilibrium (resp. strict Nash equilibrium) then it is stable (resp. strictly stable).*

The proof is trivial and follows from definitions. Hence characterising the set of all stable networks or establishing some necessary properties of these networks certainly establishes the same properties for the more restricted class of Nash equilibrium networks. Note that the converse need not be true, although stable graphs which are not Nash equilibria appear to be somehow difficult to construct.

Remark 11 *The question of defining a family of stable graphs which are not Nash equilibria is left open.*

Definition 12 *The distance between a given pair of nodes (u, v), denoted by $d(u, v)$, is the length of a minimum uv-path. If no such path exists it is set equal to $+\infty$.*

Definition 13 *A network is strongly connected if, and only if, for any pair (u, v) of its nodes, there exists a path going from node u to node v.*

Definition 14 *The diameter Δ of a network is defined by $\Delta = \max\{d(u, v), u \neq v\}$.*

The diameter of a network is thus the maximum distance between any two nodes of the network. Note that the diameter of a network is finite if, and only if, the network is strongly connected.

Proposition 15 *If $w < \frac{g-c}{2}$ then any stable network is strongly connected and its diameter satisfies $\Delta \leq \min(k_0 - 1, \frac{\frac{g+c}{2}+w}{\frac{g+c}{2}-w})$.*

The proof is provided in Annex 6.5.

Notice in passing that another (weaker) necessary condition for the minimum path to be able to carry out trade is that $g - \Delta c > 0$ ie $\Delta \leq g/c$. Hence the above results also contains the somehow simpler condition that $\Delta \leq g/c$.

Remark 16 *The question of the optimality of the bound is left open.*

Definition 17 *Let G be a network and A be a node of G. We denote by $N^+(A)$ the set of all nodes towards which A maintains a link, and by $N^-(A)$ the set of all nodes which maintain a link towards A. The cardinality of $N^+(A)$ is called the out-degree of A, and is denoted $d^+(A)$, while the cardinality of $N^-(A)$ is called the in-degree of A, and is denoted $d^-(A)$.*

Definition 18 *Let G be a network. The maximum out-degree of G is the maximum of the set of the out-degrees of the nodes of G, that is: $d^+(G) = \max\{d^+(A), A \in V(G)\}$. It is denoted $d^+(G)$. Similarly, the maximum in-degree of G is the maximum of the set of the in-degrees of the nodes of G, and is denoted $d^-(G)$. Hence $d^-(G) = \max\{d^-(A), A \in V(G)\}$.*

4.3 Extreme cases: no links, or all the links

4.3.1 The empty network

Proposition 19 *If $w \geq \frac{g-c}{2}$ then the empty network is a Nash equilibrium.*

Proof. If any player deviates unilaterally from the strategy profile yielding the empty network, it means it has built some number k of edges. But building such edges incurs a cost of kw, and a profit of only $k\frac{g-c}{2}$. Since $k\frac{g-c}{2} - kw \leq 0$, this player is not better off. ■

Clearly, if a strict inequality hold, the empty network becomes a strict Nash equilibrium.

Note that if $\frac{g-c}{2} \leq w < g - c$ deviating would be socially strictly more efficient. Indeed, the aggregate profit would be of $g - c - w > 0$.

Remark 20 *This proposition does not imply that the empty network is the only Nash equilibrium. For example, assume $g > 2c$ and $\frac{g-c}{2} \leq w \leq \frac{g-c}{2} + \frac{g-2c}{3}$. Then the circle of cardinality three is a strict Nash equilibrium, as well as the empty network.*

Nevertheless we have the following result:

Proposition 21 *If $w > \frac{g-c}{2} + 2(n-2)\frac{g-2c}{3}1_{(k_0>2)} + (n-2)(n-3)\frac{g-3c}{4}1_{(k_0>3)} =: l_0$, then the empty network is the only strict stable network (and the only strict Nash equilibrium).*

The proof is provided in Annex 6.5.

4.3.2 The complete network

Proposition 22 *If $w < \frac{g+c}{6}$, then the complete network is the unique strict stable network (and it is a strict Nash equilibrium).*

Proof. Suppose $w < \frac{g+c}{6}$ and let G be a stable network. First, since $\frac{g+c}{6} < \frac{g-c}{2}$ we have that $w < \frac{g-c}{2}$ and thus by Proposition 15 the network is strongly connected. We then prove a property called transitivity: if (A, B) is in the network and (B, C) is in the network then (A, C) is in the network. Together with strong connectivity it can easily be checked that it implies the network G is the complete network.

Let (A, B) and (B, C) be two edges in the network. Assume by contradiction that the edge (A, C) is not in the network. Then the additional profit for A to build edge (A, C) is equal to the sum of the additional gains from trades of type (A, C) and of the intermediation fees from all new minimum paths using AC, minus the cost w of building AC. Since the additional gains from being able to settle trades with C directly (without using B as an intermediary) are $\frac{g-c}{2} - \frac{g-2c}{3} > w$, the node A is willing to create the edge (A, C), a contradiction with stability. Hence the edge (A, C) was already there. This concludes the proof. ■

4.4 Intermediate cases

The following results single out the complete star as being a remarkable structure, since it is always a NE when the costs of building new edges are neither too high, nor too low. Again, there may exists other NE, like the triangle mentioned in Remark 20.

4.4.1 Characterisation of stable and Nash equilibrium complete stars

Proposition 23 *Assume $g > 2c$. A complete star of size n is a stable network if, and only if,*

$$\frac{1}{6}(g+c) \leq w \leq \frac{g-c}{2} + (n-2)\left(\frac{g-2c}{3}\right)$$

Proof. a) No leaf is strictly better off creating a new link if, and only if

$$\frac{g-c}{2} - w - \frac{g-2c}{3} \leq 0$$

b) Nor to delete one, if, and only if,

$$w \leq \frac{g-c}{2} + (n-2)\left(\frac{g-2c}{3}\right)$$

c) The center does not want to delete a link, if, and only if,

$$w \leq \frac{g-c}{2} + (n-2)\left(\frac{g-2c}{3}\right)$$

which happens to be the same condition as for a leaf not to delete a link. ■

The inequality, if strict, characterises stars that are strict Nash equilibrium.

Interestingly, the condition for a complete star to be a stable network (resp. strict stable network) is the same as for being a Nash equilibrium (resp. strict NE):

Proposition 24 *Assume $g > 2c$. Then stable complete stars are also Nash equilibria*

The proof is provided in Annex 6.5.

Notice that, holding all parameters except n constant and making n tend towards infinity, there exists a threshold from which the complete star becomes a Nash equilibrium. This is in line with the intuition that when the number of agents in the economy increases, it becomes more useful to have a single intermediary (and pay to it fees for its intermediation services) than to establish connections towards each individual agent.

Remark 25 *This does not imply that a subnetwork or a induced subnetwork of a network which is a star will be stable if it satisfies this condition, since strategic considerations involving nodes outside the star may dominate. For example, if a star is a subnetwork of a bigger network and some leaf A of the star is linked to many vertices outside the star, another leaf B of the star may want to establish a direct link to A to derives higher profits from trades with those out-neignboors of A which are not in the star (avoiding the center of the star as an intermediary for such trades). Hence, the concept of stability of a network does not translate easily to its subnetworks.*

We now turn our attention specifically to the case where only a maximum of three intermediaries is allowed. Although such an assumption is clearly restrictive, this model will still capture the dependency of a node's strategic choice relatively to the overall network topology. Hence, the three intermediaries cases can allow us to understand the nature of strategic interactions better.

As previously shown, for some constellations of parameters the complete star, the complete network or the empty network are stable networks or Nash equilibria. Nevertheless, to characterise all stable networks or Nash equilibria for any constellation of parameters is still an open question even in the three intermediaries case.

4.4.2 Characterisation of stable regular uninetworks

We saw that when the cost of establishing an edge w is neither too high nor too low, the complete star is always a stable network and a NE. The complete star exhibits two particular features: the first is that any two nodes are linked by a *unique* path, implying there is no alternative paths for any given trade. The second is the asymmetry of this network: the center has to sustain the costs of links towards all other vertices of the star, and derives benefits from all the trades of the network, whereas the leaves have only to pay for a single link, but are the strict intermediaries of no trade. A natural question is thus to ask if a more "balanced" structure – by which we mean more symmetry in the underlying network – would necessarily provide alternative paths for dealing with a same given trade. This section answers this question by the affirmative: a more regular structure implies more alternative paths for carrying out a given trade. Very symmetric and stable structure cannot exhibit monopolistic properties with respect to the choice one has to execute a trade, there can be no monopoly from a single node on the trade roads.

When the parameters are homogenous, it makes sense to look at *symmetric equilibrium*, that is, equilibrium where all players play the same strategy. This is because looking at symmetric equilibrium is consistent with not favouring one agent more than another in predicting the network that will result from the game. More precisely, a symmetric equilibrium is a vertex-transitive graph, that is, a graph G for which, for any player i, j there exists a graph automorphism σ of G such that i to j. Hence, both i and j are indeed playing exactly the same action with respect of the other nodes once a proper relabelling of the names of the players has been applied. These are thus very "balanced" networks. In particular, these networks are *d-regular*, that is all their nodes have out-degree and in-degree d. We say a network is *regular* if there exist a natural number d such that it is d-regular. In what follows we will actually study the more general case of regular networks (some regular networks do not arise from symmetric strategies, but from other, less stringent, strategies).

Given a game $\mathcal{G}^n_{k_0}$, that is, a maximum number k_0 of intermediaries allowed, a *uninetwork* is a network such that each trade (A, B) has precisely *one* corresponding admissible path. For example, in the case of $\mathcal{G}^n_3$, that is, when a maximum of three intermediaries is allowed, a uninetwork is such that for any trade (A, B) either A is directly linked to B or there exists a unique intermediary C such that (A, C) and (C, B) are in the network. Consequently the structures depicted in Figure 5 cannot be subnetworks of any uninetwork of a game $\mathcal{G}^n_3$, and, by extension, of any uninetwork of $\mathcal{G}^n_k$ for $k \geq 3$.

In our model, we do not need a given node A to possess a closed path towards itself to carry out a domestic transaction. Nevertheless, to use results from algebraic graph theory, we will make here this requirement, that is, we ask that in a uninetwork there is precisely one single admissible path from any node A to any node B with possibly $B = A$. Admittedly, this imposes more structure than should have been the case, but this additionalrequirement enable us to use the classic algebraicresult that the (i, j)th entry of the kth power of the adjacency matrix of a graph gives the exact number of walks from node v_i to node v_j. We use one of Gimbert [20]'s main theorem to derive a necessary condition for a regular uninetwork to be stable. Because the condition cannot be fulfilled for any parameter constellations we then deduce there exists no such object. For the games $\mathcal{G}^n_k$ with $k \geq 4$, the inexistence of stable regular uninetwork is straightforwardly implied by a result from the literature that there exists no regular uninetwork with a unique path of length at most l for $l \geq 3$ between any two vertices. This

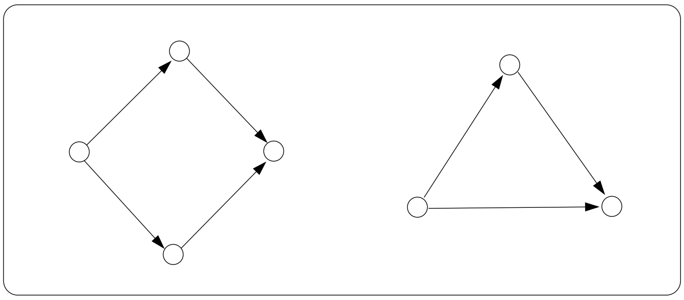


Figure 5: Forbidden subnetworks in uninetwork $\mathcal{G}_3^n$

partial result of unstability of regular uninetwork favours the view that more balanced networks do indeed provides nodes with alternative means of carrying out the same trade, as opposed to the most centralised network, the complete star.

Definition 26 *Given a network G the line-network of G is the network LG with set of nodes $V(LG) = E(G)$, the set of edges of G, and where a node (A, B) of $V(LG)$ is linked to a node (C, D) of $V(LG)$ if, and only if, $B = C$ in G.*

Theorem 27 *[Gimbert] Let $d \geq 2$. There exists a unique d-regular network for which there exists a unique path of length 1 or 2 between any two distinct nodes and a unique closed path of length 2 from any given node to itself: it is the line network LK_{d+1} of the complete network K_{d+1}.*

Figure 6 illustrates such a network for $d = 2$:

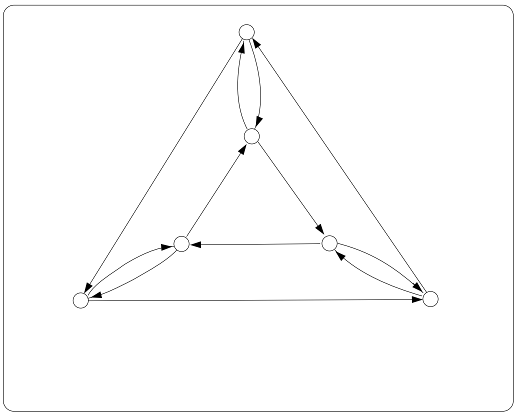


Figure 6: The unique regular uninetwork of degree 2

Proposition 28 *Assume $g > 2c$ Then a d-regular network is stable if, and only if,*

$$2d\left(\frac{g-2c}{3}\right) + \left(\frac{g+c}{6}\right) \leq w \leq \frac{g-c}{2} + 2d\left(\frac{g-2c}{3}\right)$$

The proof is provided in Annex 6.5.

4.5 When linking up networks does not re-organize them: some examples.

Real-life networks are seldom completely built from scratch. Rather, they usually arise from linking up pre-existing networks. In a market-environment with no central social planner, linking up is left to the incentives of individuals. Hence there is always the possibility that the linking, left to individualistic profit-maximising agents, would not result in an efficient network for the whole, even though taken individually each pre-existing network (which could be thought as, for example, national markets before the opening of their borders) was efficient. That is, the incentives of the agents in linking these different historical networks may not lead to an efficient network of the whole.

The notion of incentives is, in our framework, captured by the notion of stability and Nash-equilibrium. Indeed, by definition in a strict Nash equilibrium no player would add and/or delete any set of links, since it would result in a lower payoff for him. In a strict stable network, no player would add a single new link nor delete a single existing link for the same reason. Now the problem of knowing if letting individual agents linking separated networks leads to an efficient network can be easily formalised in the following way: a certain number of efficient networks H_1, H_2,..., H_k are suddenly considered as the different disjoint parts of a (bigger) aggregate network G. Which structure will arise? Will this structure be efficient? Are there many different structures that can arise from such an operation or is the outcome predictable?

Definition 29 formalizes this idea:

Definition 29 *A network G has a partition $\{H_i,\ i \in I\}$ of subnetworks if, and only if, each H_i is a subnetwork of G, and each node of G belongs to a unique H_i. In other words:*

$$E(H_i) \subseteq E(G) \text{ for all } i \in I \text{ and } V(G) = \coprod_{i\in I} V(H_i)$$

We see that asking G to be stable, or to be a Nash equilibrium, is actually simply rephrasing the previous problem in the particular case where none of the pre-existing networks H_1, H_2,..., H_k has been modified - and thus all are still present as such in the resulting aggregate network. That is, "domestic" structures were preserved and linked in a way compatible with the individual incentives of each node. How does this linking occur? Does it lead to an efficient network for the whole without the need for coordination of the economic agents involved?

A simple example where the answer is affirmative is for parameter constellations such that $w < \frac{g+c}{6}$. Indeed, by Proposition 22 we know that the unique Nash equilibrium network in that case is the complete network, and it is easily checked that it is also efficient. Now it can easily be show that taking the union of any number of complete networks will result in a global network which is also complete, hence also efficient for such constellations of parameters: since $w < \frac{g-c}{2}$ the union, at equilibrium, will be strongly connected. Because $w < \frac{g+c}{6} = \frac{g-c}{2} - \frac{g-2c}{3}$ one can prove that transitivity holds at equilibrium, and conclude as in the proof of Proposition 22.

Theorem 31 and Theorem 36 give an insight in the case where the pre-existing markets were complete stars, and assuming the star structure on each (national) pre-existing network is maintained in the equilibrium. These theorems provide a condition under which the supremacy of the centers of a star is in fact re-enforced by aggregation, since each foreign leaf wants to send a link to the center of each stars, and to their center uniquely. This proves that, for some constellation of parameters, aggregation of different markets can actually increase the power of the historical "oligopoly" (formed by the centers of the stars) instead of creating new, alternative intermediation chains.

We then prove that, for a wide range of parameters, these resulting networks are Pareto-dominated by the complete star. In particular, they cannot be efficient. More precisely, we will prove that the leaves would always be better off in a complete star, while for the center s_i of star S_i this depends on a condition that interlinks the size of the network, the size of S_i and the number of other stars.

Definition 30 *Define a multistar as a network G having a partition $\{S_i,\, i \in I\}$ of complete stars S_i of center s_i such that each of the node $v \notin \{s_i,\, i \in I\}$ is linked to precisely all the $\{s_i,\, i \in I\}$ and each of the center s_i is linked to precisely all the nodes in $(S_i \setminus \{s_i\}) \cup \{s_j,\, j \in I \backslash \{i\}\}$. We then say that G is a multistar with partition $\{S_i,\, i \in I\}$ of complete stars S_i. The leaves of the multistar G are defined as the leaves of the S_i, $i \in I$.*

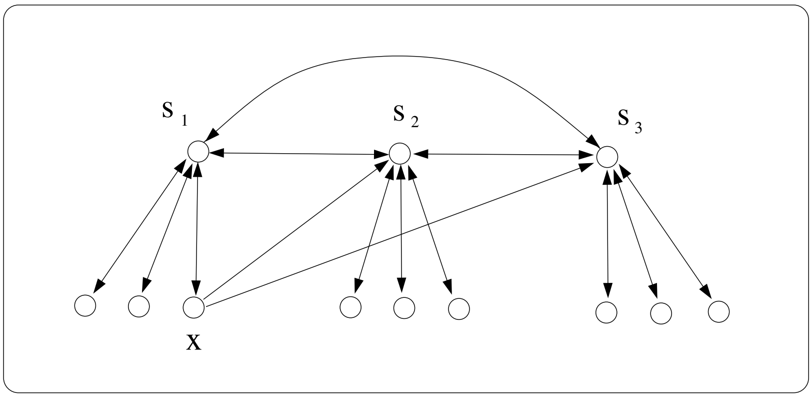


Figure 7: Representation of a multistar. Only all the links of a single leaf, the node x, are represented in the picture for clarity.

Theorem 31 *Assume $g > 2c$ and $w > \frac{g+c}{6}$. Let G be a strict Nash equilibrium. Then there is a threshold $d_0 \leq \frac{3}{g-2c}(w - \frac{g+c}{6}) + 1$ such that, if G has a partition $\{S_i,\, i \in I\}$ of complete stars, where s_i denotes the center of star S_i, and such that, for all $i \in I$, we have:*

$$|(N^+(s_i) \cap S_i) \backslash \bigcup_{w \notin S_i} N^+(w)| > d_0 \quad (C)$$

then G is a multistar with partition $\{S_i,\, i \in I\}$ of complete stars.

A similar result to Theorem 31 can be derived if we allow more than three intermediaries (Theorem 36, Annex 6.5). The proof is of Theorem 31 is provided in Annex 6.5. Conversely, it can be checked that this unique possible structure is indeed a NE when we had the now well-known necessary condition that $w \leq \frac{g-c}{2} + (n-2)\frac{g-2c}{3}$, together with a sufficiently high degree for each center of the stars.

Proposition 32 *Let G be a multistar with partition $\{S_i,\, i \in I\}$ of complete stars S_i. Assume moreover than $d^+(s_i) - |I| + 1 > d_0$ for all $i \in I$, and that $\frac{g+c}{6} < w < \frac{g-c}{2} + (n-2)\frac{g-2c}{3}$. Then G is a strict Nash equilibrium.*

The proof is provided in Annex 6.5.

4.5.1 Efficiency comparison between stars and multistars

The previous section showed how multistar can endogenously arise from the linking up of stars, with each star having its previous links preserved, as formalised in the definition of the partition of the network we provided (Definition 29). Assume G is a multistar with partition $\{S_i,\, i \in I\}$ of complete stars S_i of center s_i. When $d^+(s_i) - |I| + 1 > d_o$ for all i, and $\frac{g+c}{6} < w < \frac{g-c}{2} + (n-2)\frac{g-2c}{3}$, we know by Proposition 23 and Proposition 32 that both the complete star and the multistar are strict Nash equilibrium. The question then arises to know which of the two networks is the most efficient from an aggregate perspective (social welfare). This has clear policy implications, if we remember that the multistar arises naturally as an equilibrium when the re-organisation of the network obtained from various disjoint stars is left to the individual, profit-maximising agents of each of the stars. We characterise parameter constellations for which, from an aggregate utility perspective, the star is always a more efficient network than a multistar.

Theorem 33 *Let* $h(c, w) = 5n/4 - 1 - (n-2)w/c$ *. Let* i *be the closest positive number to* $h(c, w)$*. Then among all the multistars, the ones with presicely* i *centers are the most efficient networks.*

The proof uses simple combinatorics arguments and the maximisation of the aggregate cost function. It is relegated to Annex 6.5.

Corollary 34 *If* $h(c, w) < 1.5$ *then the star is more efficient than any other multistar.*

To compare stars and multistars, we can also use the notion of Pareto-domination for unlabelled graphs. We will convene that a (unlabelled) network G Pareto-dominates another (unlabelled) network H if there is some way to relabel the vertices of G such that each node in the relabelled network is better off than in H.[8] A characterisation for stars to Pareto-dominate multistar in that sense is given in Theorem 39 in Annex 6.5.

Notice that the arising inefficiencies of multistars do not come from the fixed investments spent for building links - as would undoubtedly also be the case in real life - but lie uniquely in the new individual incentives for reshaping the network that take into account the pre-existing structures and thus are not enough to lead to an efficient resulting network. This finding has obvious policy implications, such as the potential benefits of the intervention of a "social planner" for re-arranging by himself the network or for acting as a catalyst for coordination among the different players, or for introducing incentives exogenous to the network that could help fostering a more efficient aggregate network.

5 Conclusion

We defined rules for network formation to model a variety of situations, including cross-agent links or, alternatively, banks holding an account with each other's for the purpose of securities settlement. These evolving network models are fairly general and can be used in any setting where building a link incurs immediate costs to the builder of the link but also potential future rewards. We applied our models, assuming a hypothetical pattern of trade, both in the case of an evoluting network, where links are created when the expected benefits derived from them are higher than the costs of establishing them, and in the case of a complete network, where all links are present. This allowed us to compute the difference in profits in the two situations. This difference is a closed-form formula which makes analytically explicit the different types of benefits stemming from establishing a complete network: the possible reduction of the costs of domestic trades, the reduction of the costs of cross-border trades, the additional profits from better coverage, and the difference between the current fixed costs due to economies of scale.

We then modelled the decision to join or not of the agents based on these results. We found that the decision to join of a given agent is highly influenced by other agents' costs and assumed decisions. In particular, in some situations that we characterise, some level of synchronisation between agents who plans to join can be useful for them to reach more efficient Nash equilibria. If negative believes and lack of coordination are too strong, agents may be stuck in a Pareto-inferior Nash equilibrium. The policy

[8] It is easy to see how the notion of Pareto-domination for an unlabelled graphs may differ from the more common game theoretic Pareto-domination. For example consider two identical complete stars of same size G and H, and assume their center earns more than their leaves. Suppose also that node 1 is the name of the center of G but the name of a leaf in H. Then H does not Pareto-dominate G as a labelled graph, since the node 1 is worse off in the network H than in the network G. But H does Pareto-dominate G as an unlabelled graph: there is a way to re-labell its vertices so that to obtain a labelled graph which Pareto-dominates H: any re-labelling where the center of G has the same name as the center of H will of course do.

implication is that incentives for agents to synchronize their moves into the network are useful in certain situations. The desirability of such synchronisation depends on the dispersion of the agents' individual adaptation costs along the real line, and we gave an explicit condition of when synchronisation for moving to the network is useful.

We concluded the article by a purely theoretical section analyzing different general network structures, and deriving the conditions for such networks to be stable – that is, no individual in the network would be better off by adding or deleting a link it has some power on. In particular we derived two theorems in which the overall network, obtained by linking up stars, is not efficient nor Pareto-efficient, but is stable. This emphasizes the theoretical need, in some case, for a "social planner" to re-arrange the network or at least act as a catalyst in the new network formation, in particular in complex networks obtained by linking pre-existing networks, even when these smaller networks were, individually taken, efficient. We signal that further work is needed to characterise, in this context, the whole set of efficient networks, as well as of stable networks and Nash equilibrium network, but that it appears a difficult tasks, in particular when long chains of intermediaries are allowed, because of the strategic considerations at play, which are function of a greater portion of the topology of the network.

References

[1] Bala V., Goyal S. 2000. A noncooperative model of network formation. *Econometrica* **68**: no 5, 1181-1229.

[2] Belleflamme P., Bloch F. 2004. Market sharing agreements and collusive networks. *International Economic Review* **45**: no 2, 387-411.

[3] Calvo-Armengol A., Jackson M. O. 2004. The effect of social networks on employment and inequality. *American Economic Review* **94**: 426-454.

[4] Van Cayseele P., Wuyts C. 2007. Cost efficiency in the European securities settlement and depository industry. *Journal of Banking & Finance* **31**: no 10, 3058-3079.

[5] Cales M.-N., Chabert D., Hichri W., Marchand N. 2011. The reform of the European securities settlement systems: towards an integrated financial market. Groupe d'analyse et de théorie économique de Lyon-St Etienne.

[6] Cherbonnier F., Rochet J.-C. 2010. Vertical integration and regulation in the securities settlement industry. Working Paper.

[7] Clearstream. 2002. Cross-border equity trading, clearing and settlement in Europe. White Paper.

[8] Clearstream. 2005. The European post-trade market – an introduction. `http://www1.deutsche-boerse.com/INTERNET/EXCHANGE/zpd.nsf/Web\%2BPublikationen/HAMN-698JWV/\%24FILE/WP\_European\_Post-Trade\_Market.pdf.http://www1.deutsche-boerse.com/INTERNET/EXCHANGE/zpd.nsf/PublikationenID/AKLS-58TMV6/\%24FILE/White-Paper\_online\_d.pdf`.

[9] Hirata de Carvalho C. 2004. Cross-border securities clearing and settlement infrastructure in the European Union as a prerequisite to financial markets integration: challenges and perspectives. *HWWA Discussion Paper* **287**.

[10] Dutta B., Ghosal S., Ray D. 2004. Farsighted network formation. *Journal of Economic Theory* **122**: 143-164.

[11] ECB. 2007. *Target2 Securities economic feasibility study*. `https://www.ecb.int/press/pr/date/2007/html/pr070308_2.en.html`.

[12] ECB. 2008. *Target2 Securities economic impact assessment*. https://www.ecb.int/paym/t2s/about/keydocs/html/index.en.html.

[13] ECB. 2009. *Eligibility criteria for CSDs in Target2 Securities*. http://www.ecb.int/paym/t2s/pdf/csd/eligibility_criteria.pdf.

[14] ECB. 2010. *The Payment System. Payments, Securities and Derivatives, and the Role of the Eurosystem*. Editor: Tom Kokkola.

[15] Economides N. 1996. The economics of networks, *International Journal of Industrial Organization* **14**: no 2, 675-699.

[16] Economides N., Salop S. C. 1992. Competition and integration among complements, and network market structure. *The Journal of Industrial Economics* **10**: no 1, 105-123.

[17] Eliashberg J., Steinberg R. 1991. Competitive strategies for two firms with asymmetric production cost structures. *Management Science* **37**: no. 11, 1452-1473.

[18] FESE. 2008. *Share Ownership Structure in Europe*. http://www.fese.be/_lib/files/Share_Ownership_Survey_2007_Final.pdf.

[19] Fine C. H., Freund R. M. 1990. Optimal investment in product-flexible manufacturing capacity. *Management Science* **36**: no. 4, 449-466.

[20] Gimbert J. 1999. On digraphs with unique walks of closed lengths between vertices. *The Australasian Journal of Combinatorics* **20**: 77-90.

[21] Giovannini A. 2001. Cross-border clearing and settlements arrangement in the European Union, http://ec.europa.eu/internal_market/financial-markets/docs/clearing/first_giovannini_report_en.pdf.

[22] Giovannini A. 2003. Second report on clearing and settlement arrangements in the European Union, http://ec.europa.eu/internal_market/financial-markets/docs/clearing/second_giovannini_report_en.pdf.

[23] Goyal S., Joshi S. 2003, Networks of collaboration in oligopoly. *Games and Economic Behavior* **43**: 57-85.

[24] Goyal S., Vega-Redondo F. 2003. Network formation and social coordination. *Working Papers* **481**, Queen Mary, University of London, School of Economics and Finance.

[25] Holthausen C., Tapking J. 2004. Raising rival's costs in the securities settlement industry. *Journal of Financial Intermediation* **16**: 91-116.

[26] Jackson M. O. 1996. A strategic model of social and economic networks. *Journal of Economic Theory* **71**: 44-74.

[27] Jackson M. O. 2002. The evolution of social and economic networks. *Journal of Economic Theory* **106**: 265-295.

[28] Jackson M. O., Wolinsky A. 1996. A strategic model of economic and social networks. *Journal of Economic Theory* **71**: (1) 44-74.

[29] Jackson M. O., Watts A. 2001. The existence of pair-wise stable networks. *Seoul Journal of Economics* **14**: 299-321.

[30] Kauko K. 2002. Links between securities settlement systems: an oligopoly theoretic approach. *Bank of Finland Discussion Papers* **27**.

[31] Kauko K. 2007. Interlinking securities settlement systems: a strategic commitment? *Journal of Banking and Finance* **31**: 2962-2977.

[32] Kemppainen K. 2008. Integrating european retail payment systems: some economics of SEPA. *Bank of Finland Research Discussion Paper* **22**.

[33] Knieps G. 2006. Competition in the post-trade markets: a network economic analysis of the securities business. *Journal of Industry, Competition and Trade* **6**: no 1, 45-60.

[34] Koeppl, T. V., Monnet C. 2007. Guess what: it's the settlements! Vertical integration as a barrier to efficient exchange consolidation. *Journal of Banking & Finance* **31**: 3013-3033.

[35] Lannoo K. 2001. EU securities market regulation: adapting to the needs of a single capital market. *CEPS Research Report*, Brussels.

[36] Lannoo K., Levin M. 2001. The securities settlement industries in the EU: structure, costs and the way forward. *CEPS Research Report*, Brussels.

[37] Matutes C., Padilla A. J. 1994. Shared ATM networks and banking competition. *European Economic Review* **38**: 1113-1138.

[38] Maskin E., Tirole J. 1988. A theory of dynamic oligopoly, II: price competition, kinkeddemand curves, and edgeworth cycles. *Econometrica* **56**: no. 3, 571-599.

[39] McAndrews J. J., Rob R. 1996. Shared ownership and pricing in a network switch. *International Journal of Industrial Organization* **14**: 727-745.

[40] Mercier, F., Sauer, S. 2014. Optimal CSD reshaping towards T2S, *ECB Working Paper Series* **1547** (2013) and *Journal of Financial Market Infrastructures* **2**: 3-51.

[41] Milne A. 2007. The industrial organisation of post-trade clearing and settlement. *Journal of Banking and Finance* **31**: 2945-2961.

[42] Oxera. 2007. Methodology for monitoring prices, costs and volumes of trading and post-trading services, report prepared for the European Commission Directorate General Internal Market and Services

[43] Oxera. 2009. Monitoring prices, costs and volumes of trading and post-trading services, report prepared for the European Commission Directorate General Internal Market and Services

[44] Oxera. 2011. Monitoring prices, costs and volumes of trading and post-trading services, report prepared for the European Commission Directorate General Internal Market and Services, http://ec.europa.eu/internal_market/financial-markets/docs/clearing/2011_oxera_study_en.pdf.

[45] Park Y.-S., Ahn B.-H. 1999. Joint ownership and interconnection pricing in network industries. *International Review of Economics and Finance* **8**: 183-198.

[46] Pirrong C. 2007. The Industrial Organization of Execution, Clearing and Settlement in Financial Markets. *Working Paper*, `http://www.cba.uh.edu/spirrong/`.

[47] Rochet. 2006. The Welfare Effects of Vertical Integration in the Securities Clearing and Settlement Industry, Mimeo, IDEA, Toulouse University.

[48] Schmiedel H., Malkamaki M., Tarkka J. 2006. Economics of sale and technological developments in securities depository and settlement systems, *Journal of Banking & Finance* **30**: 1783–1806.

[49] Siedlarek J. 2011. Intermediation in Networks, Nota di Lavoro. *FEEM Working Paper* **42**.

[50] Spence M. 1984. Cost reduction, competition, and industry performance. *Econometrica* **52**: no 1, 101-122.

[51] Tapking J. 2007. Pricing of settlement link services and mergers of central securities depositories. *ECB Working Paper Series* **710**.

[52] Tapking J., Yang J. 2004. Horizontal and vertical integration in securities trading and settlement. *ECB Working Paper Series* **387**.

[53] Tirole J. 1988. *The theory of industrial organisation*, the MIT press, Cambridge.

[54] Vickers J. 1995. Competition and regulation in vertically related markets. *Review of Economic Studies* **62**: 1-17.

[55] Valletti T. M., Cambini C. 2005. Investments and network competition. *The RAND Journal of Economics* **36**: no 2, 446-467.

[56] Volgelsang I. 2003. Price regulation of access to telecommunications networks. *Journal of Economic Literature* **41**: 830-862.

[57] Wu S. D. 2005. Managing capacity in the high-tech industry: a review of literature. *The Engineering Economist* **50**: 125-158.

6 Annexes

6.1 Annex 1: derivation of profit expectations for a r-regular, non-evoluting network

Taking expectation on both sides of the previous expression of profits yields:

$$\begin{aligned}
E(profits) &= \int_{\omega\in\Omega} \sum_{t=1}^{N} (g^t_{e_t(\omega)} - c^t_{e_t(\omega)}) 1_{(e_t \in E(G))}(\omega) dP(\omega) - C_{fixed} \\
&= \sum_{t=1}^{N} \int_{\omega\in\Omega} (g^t_{e_t(\omega)} - c^t_{e_t(\omega)}) 1_{(e_t \in E(G) \text{ and } e_t \text{ is a loop})}(\omega) dP(\omega) \\
&\quad + \sum_{t=1}^{N} \int_{\omega\in\Omega} (g^t_{e_t(\omega)} - c^t_{e_t(\omega)}) 1_{(e_t \in E(G) \text{ and } e_t \text{ is not a loop})}(\omega) dP(\omega) - C_{fixed} \\
&= \sum_{t=1}^{N} (g^t_{dom} - c^t_{dom}) \int_{\omega\in\Omega} 1_{(e_t \in E(G) \text{ and } e_t \text{ is a loop})}(\omega) dP(\omega) \\
&\quad + \sum_{t=1}^{N} (g^t_{cross} - c^t_{cross}) \int_{\omega\in\Omega} 1_{(e_t \in E(G) \text{ and } e_t \text{ is not a loop})}(\omega) dP(\omega) - C_{fixed}
\end{aligned}$$

The integral in the first sum is equal to $P(e_t \in E(G) \text{ and } e_t \text{ is a loop}) = P(e_t \text{ is a loop}) P(e_t \in E(G) \mid e_t \text{ is a loop}) = \alpha.1$ and the second to $P(e_t \in E(G) \text{ and } e_t \text{ is not a loop}) = P(e_t \text{ is not a loop}) P(e_t \in E(G) \mid e_t \text{ is not a loop}) = (1-\alpha)\frac{r}{n(n-1)}$, hence we get:

$$E(profits) = \alpha \sum_{t=1}^{N} (g^t_{dom} - c^t_{dom}) + \frac{r(1-\alpha)}{n(n-1)} \sum_{t=1}^{N} (g^t_{cross} - c^t_{cross}) - C_{fixed}$$

Supposing for simplicity gains and costs do not vary over time, we get:

$$E(profits) = \alpha N (g_{dom} - c_{dom}) + N(1-\alpha)\frac{r}{n(n-1)} (g_{cross} - c_{cross}) - C_{fixed}$$

6.2 Annex 2: Closed form formula for evolving networks

6.2.1 Analysis of the second model

General formula for the profits Without loss of generality, we can assume $w_{AB}^t = 0$ for all edges AB already there before period 1. From the definition of the first model, it is easy to see that the profits associated to a given sequence of ideas of trade $(e_t)_{t=1...T}$ are:

$$profits = \sum_{t=1}^{N} \max(0, g_{e_t}^t - c_{e_t}^t) 1_{\mathcal{E}_{e_t}^t} - \sum_{t=1}^{N} \sum_{A,B \in E(G)} w_{AB}^t 1_{(E_{AB}^t \geq w_{AB}^t) \cap \overline{\mathcal{E}_{AB}^{t-1}}} - C_{fixed} \qquad (P_1)$$

where we denote by $\mathcal{E}_{AB}^t$ the event that AB is an edge before step 3) of period t.

We assume in the following that g_{AB}^t and c_{AB}^t do not depend on the edge AB, so that we can write:

$$E(profits) = \sum_{t=1}^{N} \max(0, g_{AB}^t - c_{AB}^t) P(\mathcal{E}_{e_t}^t) - \sum_{t=1}^{N} \sum_{A,B \in E(G)} w_{AB}^t P((E_{AB}^t \geq w_{AB}^t) \cap \overline{\mathcal{E}_{AB}^{t-1}}) - C_{fixed} \qquad (P_1')$$

The expected profits at time t derived from a given edge AB can be expressed as:

$$E_{AB}^t = \sum_{k=t}^{N} \max(0, g_{AB}^k - c_{AB}^k) P_{AB}^k \qquad (Q_1)$$

The event $(E_{e_t}^t > w_{e_t}^t) = \{\omega \in \Omega \quad E_{e_t(\omega)}^t \geq w_{e_t(\omega)}^t\}$ can be different from the whole universe and the empty set because it depends on the edge e_t which is a random edge, and thus on the probability distribution for selecting an edge. Hence the probability of this event will depend of the probability measure defined on Ω, usually $P = \bigotimes_{t=1}^{N} P^t$. In the sections illustrating evolving network we actually avoid this difficulty by making strong assumptions of uniformity across the parameters of the different edges, namely, the same assumptions of homogeneity among cross-border and domestic links as in Section 2.2.3.

6.2.2 Analytical resolution: a closed-form expression of the network benefits

Assume the same set of assumptions on the parameters of the model as in Section 2.2.3. Suppose AB is not a loop and let us re-write the event $(E_{AB}^t \geq w_{AB}^t)$ in this context; using equality (Q_1) we can write:

$$E_{AB}^t \geq w_{AB}^t \Leftrightarrow \sum_{k=t}^{N} (g^k - c^k) P_{AB} \geq w_{AB}^t \Leftrightarrow (N - t + 1)(g - c) P_{AB} \geq w \Leftrightarrow t \leq q$$

where $q = N + 1 - \lfloor \frac{w}{(g-c)P_{AB}} \rfloor$.

Thanks to our simplifying assumptions we can re-write the expectation of the profits as:

$$\begin{aligned} E(profits) \quad &= \quad \alpha N (g_{dom} - c_{dom}) + \sum_{t=1}^{N} (g_{cross} - c_{cross}) P(\mathcal{E}_{e_t}^t \cap (e_t \text{ is not a loop})) \\ - \sum_{t=1}^{N} n(n-1) w P((E_{AB}^t \quad &\geq \quad w_{AB}^t) \cap \overline{\mathcal{E}_{AB}^{t-1}} \cap (AB \text{ is not a loop})) - C_{fixed} \end{aligned}$$

Now

$$\begin{aligned} P(\mathcal{E}_{e_t}^t \cap (e_t \text{ is not a loop})) \quad &= \quad P(\mathcal{E}_{e_t}^t | (e_t \text{ is not a loop})) P(e_t \text{ is not a loop})) \\ &= \quad F_t (1 - \alpha) \end{aligned}$$

where

$$\begin{aligned} F_t &= P(\mathcal{E}_{e_t}^t|(e_t \text{ is not a loop})) = P(\bigcup_{i=1}^{t}(E_{e_t}^i \geq w_{e_t}^i|(e_t \text{ is not a loop}))) \\ &= \sum_{k=1}^{t}(-1)^{k-1}\sum_{\{i_1,\dots,i_k\}\subseteq\{1,\dots t\}} P(\bigcap_{j=1}^{k}(E_{e_t}^{i_j} \geq w_{e_t}^{i_j}|(e_t \text{ is not a loop}))) \end{aligned}$$

Note that for all i, $(E_{e_t}^i \geq w_{e_t}^i|(e_t$ is not a loop)) is either Ω or $\varnothing$ (depending if $i \leq q$ or not), hence this family of events is independent of any other event. In particular, for any i and j, the event $(E_{e_t}^i \geq w_{e_t}^i|(e_t$ is not a loop)) is independent from $(E_{e_t}^j \geq w_{e_t}^j|(e_t$ is not a loop)). Hence we have

$$P(\bigcap_{j=1}^{k} E_{e_t}^{i_j} \geq w_{e_t}^{i_j}|(e_t \text{ is not a loop})) = \prod_{j=1}^{k} P(E_{e_t}^{i_j} \geq w_{e_t}^{i_j}|(e_t \text{ is not a loop})) = \prod_{j=1}^{k} 1_{(i_j \leq q)}$$

Hence

$$F_t = \sum_{k=1}^{t}(-1)^{k-1}\sum_{\{i_1,\dots,i_k\}\subseteq\{1,\dots t\}}\prod_{j=1}^{k} 1_{(i_j \leq q)}$$

Erasing the (eventual) terms which are 0 in the second sum gives:

$$\begin{aligned} F_t &= \sum_{k=1}^{t}(-1)^{k-1}\sum_{\{i_1,\dots,i_k\}\subseteq\{1,\dots \min(t,q)\}} 1 \\ &= \sum_{k=1}^{t}(-1)^{k-1}\binom{\min(t,q)}{k} \end{aligned}$$

Remark: In particular if $t \geq q$ then $F_t = 1$, using the binomial formula on $0 = (1-1)^q$.

We now turn to the computation of $P((E_{AB}^t \geq w_{AB}^t) \cap \overline{\mathcal{E}_{AB}^{t-1}} \cap (AB$ is not a loop)). Our hypothesis on the probability distributions at each period and on the costs and gains of edges implies the two events $(E_{AB}^t \geq w_{AB}^t|(AB$ is not a loop)) and $\overline{\mathcal{E}_{AB}^{t-1}}|(AB$ is not a loop) are independent. Note that under other assumptions this need hardly be the case since a higher gains relative to costs for a particular edge can both increase the probability that $E_{AB}^t > w_{AB}^t$ and decrease the probability of $\overline{\mathcal{E}_{t-1}^{AB}}$.

$$\begin{aligned} P((E_{AB}^t &\geq w_{AB}^t) \cap \overline{\mathcal{E}_{AB}^{t-1}}) \cap (AB \text{ is not a loop}) \\ &= P((E_{AB}^t \geq w_{AB}^t) \cap \overline{\mathcal{E}_{AB}^{t-1}}|(AB \text{ is not a loop}))(1-\alpha) \\ &= P(E_{AB}^t \geq w_{AB}^t|(AB \text{ is not a loop}))P(\overline{\mathcal{E}_{AB}^{t-1}}|(AB \text{ is not a loop}))(1-\alpha) \\ &= 1_{(t-1\leq q)}(1-F_{t-1})(1-\alpha) \end{aligned}$$

since our assumptions also imply $P(\overline{\mathcal{E}_{AB}^{t-1}}|(AB$ is not a loop)) $= P(\overline{\mathcal{E}_{e_t}^{t-1}}|(e_t$ is not a loop)) for all A, B.

Now plugging these values back into our formula (P) we get:

$$\begin{aligned} E(profits) &= \alpha N(g_{dom} - c_{dom}) + (g_{cross} - c_{cross})(1-\alpha)\sum_{t=1}^{N} F_t \\ &\quad - n(n-1)w(1-\alpha)\sum_{t=1}^{N} 1_{(t-1\leq q)}(1-F_{t-1}) - C_{fixed} \end{aligned}$$

Here again to find the expected benefits of the network we subtract the formula applied to both networks, after a proper indexing of the parameters which differ across the two networks, that is, the costs, thus holding the transaction flows constant for the comparison. For the complete network $F_t = 1$ and $w = 0$ hence:

$$E(profits \text{ with the network}) = \alpha N(g_{dom} - c_{dom}^{comp}) + (g_{cross} - c_{cross}^{comp})(1-\alpha)N - C_{fixed}^{comp}$$

Subtracting (P') from this formula yields the expected benefits from adopting the network:

$$\alpha N(c_{dom} - c_{dom}^{comp}) + \{(g_{cross} - c_{cross}^{comp})(1-\alpha)N - (g_{cross} - c_{cross})(1-\alpha)\sum_{t=1}^{N} F_t$$
$$-n(n-1)w(1-\alpha)\sum_{t=1}^{N} 1_{(t-1\leq q)}(1 - F_{t-1})\} + C_{fixed} - C_{fixed}^{comp}$$

which can be further re-arranged as:

$$\alpha N(c_{dom} - c_{dom}^{comp}) + (1-\alpha)\{g_{cross}(N - \sum_{t=1}^{N} F_t) + (c_{cross}\sum_{t=1}^{N} F_t - c_{cross}^{comp}N)$$
$$-n(n-1)w\sum_{t=1}^{N} 1_{(t-1\leq q)}(1 - F_{t-1})\} + C_{fixed} - C_{fixed}^{comp}$$

Discussion on the analytical split-up of the network benefits We see that in the case of an evoluting network, the break-up of the network benefits is similar to the one discussed in the case of a non-evoluting network in Section 2.2.3. For example, the first term represents the benefits from domestic trades and is the same than in the non-evoluting model. This is because since all agents were assumed to be able to settle domestic trades, giving them the ability to build new links does not influence the benefits linked to the trade of those trades.

Now, the quantity $\sum_{t=1}^{N} F_t$ is an indicator of density of the network and play a role similar to the term $N\frac{r}{n-n-1)}$ of Section 2.2.3. One way to see this is to consider what happens when w tends to 0. Then it becomes more and more profitable to build edges, hence F_t, which is the probability that the idea of a trade selected at period t involves an edge already built, tend to 1, and $\sum_{t=1}^{N} F_t$ tend to N. The additional profits $g_{cross}(N - \sum_{t=1}^{N} F_t)$ derived from trade that could not be realised before thus tend to 0.

There is another interesting way to interpret this density indicator. First, note that all these results are derived with perfect information from part of the agents, which are assumed to know perfectly what is the probability distribution of the *ideas* of the trade wanted to be realised. Now assume for one instant it is not the case anymore, for example because this probability distribution has changed and the agents are not immediately aware of this fact, or are still learning about the new probability distribution of the ideas of the trade. Then F_t, which is the probability that the right edge be at the right place at the right time, will necessarily be lower than in our perfect information case. Hence the benefits term $g_{cross}(N - \sum_{t=1}^{N} F_t)$ from not missing any trade trade opportunities with the network will be even higher than in the perfect information case. Hence the indicator is factoring in more benefits when the idea of trade pattern is not know or ever changing – which is more likely to be the case in the real world.

The term $c_{cross}\sum_{t=1}^{N} F_t - c_{cross}^{comp}N$, if positive, reflects the decrease in aggregate costs for cross-border trade. Note that here also this term could be negative, that is, the *aggregate* costs of cross-border trade could actually increase with the adoption of the network, in case of a large increase in the volume of cross-border trades. In any case, that would not be a loss for agents as the overall term $(1-\alpha)\{g_{cross}(N - \sum_{t=1}^{N} F_t) + (c_{cross}\sum_{t=1}^{N} F_t - c_{cross}^{comp}N)$ would be positive.

The term $n(n-1)w\sum_{t=1}^{N} 1_{(t-1\leq q)}(1 - F_{t-1})$, which has a negative sign in the analytical expressions, indicate the cost-reduction in the building of new links due to the network.

The last term is the same and has the same interpretation as in Section 2.2.3.

To conclude, the discussion of the closed-form formula for the benefits of the network versus an evoluting network with perfect sighted agents provides a robustness test to the non-evolutionary formula, since most terms have similar representatives in the more complex formula.

6.2.3 Analysis of the second model

The analytical resolution of the second model, where agents are given the ability to *react* to the ideas of trade by building links in *response to* a given idea of trade wanted to be realised, is more complicated than the previous model. Nevertheless, we provide the closed-form formulas as a robustness check for the profit break down already commented previously in Section 2.2.3 and 6.2.2. Hence, these technical derivations could be skipped without impacting the understanding of the rest of the article.

General formula for the profits It is clear that, from the definition of our model, the profits associated to the realised sequence of idea of trade $(e_t)_t$ is:

$$profits = \sum_{t=1}^{N} \max(0, g^t_{e_t} - c^t_{e_t}) 1_{\mathcal{Q}^t_{e_t}} - \sum_{t=1}^{N} w^t_{e_t} 1_{(E^t_{e_t} > w^t_{e_t}) \cap \overline{\mathcal{Q}^{t-1}_{e_t}}} - C_{fixed} \qquad (P_2)$$

where we denote, for a given edge AB by $\mathcal{Q}^t_{AB}$ the event that AB is an edge after step 4) of period t (that is, after possibly being created at step 3) of period t).

In the next line we assume that g^t_{AB}, w^t_{AB} and c^t_{AB} do not depend on the edge AB, so that $g^t_{e_t}$, $w^t_{e_t}$ and $c^t_{e_t}$ are not random variable but deterministic functions and we can write:

$$E(profits) = \sum_{t=1}^{N} \max(0, g^t - c^t) P(\mathcal{Q}^t_{e_t}) - \sum_{t=1}^{N} w^t_{e_t} P((E^t_{e_t} > w^t_{e_t}) \cap \mathcal{Q}^{t-1}_{e_t}) - C_{fixed} \qquad (P_2')$$

Of course, the expected profits, at time t, derived from a given edge AB, can be expressed as:

$$E^t_{AB} = \max(0, g^t_{AB} - c^t_{AB}) + \sum_{k=t+1}^{N} \max(0, g^k_{AB} - c^k_{AB}) P^k_{AB} \qquad (Q_2)$$

Analytical resolution: a closed-form expression of the network benefits We make again the same assumptions as in Section 2.2.3 and Section 6.2.1 concerning the probability of an idea of a trade being cross-border, and about the costs and gains associated with domestic and cross-border trades.

From (Q_2) we get, assuming AB is not a loop:

$$E^t_{AB} \geq w^t_{AB} \Leftrightarrow (g - c)(1 + (N - t)P_{AB}) \geq w \Leftrightarrow t \leq q$$

where $q = N + \lfloor \frac{1 - w/(g-c)}{P_{AB}} \rfloor$.

From (P_2') we can thus deduce:

$$\begin{aligned} E(profits) &= \alpha N(g_{dom} - c_{dom}) + \sum_{t=1}^{N} (g_{cross} - c_{cross}) P(\mathcal{Q}^t_{e_t} \cap (e_t \text{ is not a loop})) \\ - \sum_{t=1}^{N} wP((E^t_{e_t} &\geq w^t_{e_t}) \cap \overline{\mathcal{Q}^{t-1}_{e_t}} \cap (e_t \text{ is not a loop})) - C_{fixed} \end{aligned}$$

Now

$$\begin{aligned} P(\mathcal{Q}_{e_t}^t \cap (e_t \text{ is not a loop})) &= P(\mathcal{Q}_{e_t}^t | (e_t \text{ is not a loop})) P(e_t \text{ is not a loop}) \\ &= P(\bigcup_{i=1}^t (A_t^i | (e_t \text{ is not a loop})))(1-\alpha) \end{aligned}$$

where $A_t^i = (E_{e_t}^i \geq w_{e_t}^i) \cap (e_t \text{ was selected at period } i) = (E_{e_t}^i \geq w_{e_t}^i) \cap (e_i = e_t)$. Using the exclusion-inclusion formula again we get

$$\begin{aligned} P(\bigcup_{i=1}^t (A_t^i | (e_t \text{ is not a loop}))) &= \sum_{k=1}^t (-1)^{k-1} \sum_{\{i_1,\ldots,i_k\} \subseteq \{1,\ldots t\}} \prod_{j=1}^k P(A_t^{i_j} | (e_t \text{ is not a loop})) \\ &= \sum_{k=1}^t (-1)^{k-1} \sum_{\{i_1,\ldots,i_k\} \subseteq \{1,\ldots t-1\}} \prod_{j=1}^k P(A_t^{i_j} | (e_t \text{ is not a loop})) \\ &+ \sum_{k=1}^t (-1)^{k-1} \sum_{\{i_1,\ldots,i_{k-1}\} \subseteq \{1,\ldots t-1\}} P(A_t^t) \prod_{j=1}^{k-1} P(A_t^{i_j} | (e_t \text{ is not a loop})) \end{aligned}$$

with $P(A_t^t) = P(E_{e_t}^i \geq w_{e_t}^i) P(e_t = e_t) = P(E_{e_t}^i \geq w_{e_t}^i) = 1_{(t \leq q)}$

If $\{i_1, \ldots, i_k\} \subseteq \{1, \ldots t-1\}$ then

$$\begin{aligned} \prod_{j=1}^k P(A_t^{i_j} | (e_t \text{ is not a loop})) &= \prod_{j=1}^k P((E_{e_t}^i \geq w_{e_t}^i) \cap (e_i = e_t)) \\ &= \prod_{j=1}^k P((E_{e_t}^i \geq w_{e_t}^i)).P(e_i = e_t) \\ &= \prod_{j=1}^k 1_{(i_j \leq q)} P_{AB} = P_{AB}^k \prod_{j=1}^k 1_{(i_j \leq q)} \end{aligned}$$

If $\{i_1, \ldots, i_{k-1}\} \subseteq \{1, \ldots t-1\}$ then, similarly:

$$\prod_{j=1}^{k-1} P(A_t^{i_j} | (e_t \text{ is not a loop})) = P_{AB}^{k-1} \prod_{j=1}^{k-1} 1_{(i_j \leq q)}$$

Note here that the upper index in P_{AB}^k and P_{AB}^{k-1} denote a power, not the time.

Hence

$$\begin{aligned} P(\bigcup_{i=1}^t (A_t^i | (e_t \text{ is not a loop}))) &= \sum_{k=1}^t (-1)^{k-1} P_{AB}^k \sum_{\{i_1,\ldots,i_k\} \subseteq \{1,\ldots t-1\}} \prod_{j=1}^k 1_{(i_j \leq q)} \\ &+ \sum_{k=1}^t (-1)^{k-1}.1_{(t \leq q)} \sum_{\{i_1,\ldots,i_{k-1}\} \subseteq \{1,\ldots t-1\}} \prod_{j=1}^{k-1} 1_{(i_j \leq q)} \\ &= \sum_{k=1}^t (-1)^{k-1} P_{AB}^k \binom{\min(t-1,q)}{k} + 1_{(t \leq q)} \sum_{k=1}^t (-1)^{k-1} P_{AB}^{k-1} \binom{\min(t-1,q)}{k-1} \end{aligned}$$

Let us now compute $P((E_{e_t}^t \geq w_{e_t}^t) \cap \overline{\mathcal{Q}_{e_t}^{t-1}} \cap (e_t \text{ is not a loop}))$.

$$\begin{aligned} P((E_{e_t}^t &\geq w_{e_t}^t) \cap \overline{\mathcal{Q}_{e_t}^{t-1}} \cap (e_t \text{ is not a loop})) = P((E_{e_t}^t \geq w_{e_t}^t) \cap \overline{\mathcal{Q}_{e_t}^{t-1}} | (e_t \text{ is not a loop}))(1-\alpha) \\ &= 1_{(t \leq q)} (1 - P(\mathcal{Q}_{e_t}^{t-1} | (e_t \text{ is not a loop})))(1-\alpha) \end{aligned}$$

with

$$
\begin{aligned}
P(\mathcal{Q}_{e_t}^{t-1}|(e_t \text{ is not a loop})) &= P(\bigcup_{i=1}^{t-1}(A_t^i|(e_t \text{ is not a loop})) \\
&= \sum_{k=1}^{t-1}(-1)^{k-1}\sum_{\{i_1,\dots,i_k\}\subseteq\{1,\dots t-1\}} P_{AB}^k \prod_{j=1}^{k} P(E_{e_t}^{i_j} > w_{e_t}) \\
&= \sum_{k=1}^{t-1}(-1)^{k-1}P_{AB}^k \sum_{\{i_1,\dots,i_k\}\subseteq\{1,\dots t-1\}} \prod_{j=1}^{k} 1_{(i_j\leq q)} \\
&= \sum_{k=1}^{t-1}(-1)^{k-1}P_{AB}^k \binom{\min(t-1,q)}{k}
\end{aligned}
$$

We thus get the general formula (recall $P_{AB} = \frac{1}{n(n-1)}$):

$$
\begin{aligned}
E(profits) \;=\; & \alpha N(g_{dom} - c_{dom}) + (1-\alpha)\sum_{t=1}^{N}(g_{cross} - c_{cross})\{\sum_{k=1}^{t}\frac{(-1)^{k-1}}{n^k(n-1)^k}\binom{\min(t-1,q)}{k} \\
& +1_{(t\leq q)}\sum_{k=1}^{t}\left(\frac{-1}{n(n-1)}\right)^{k-1}\binom{\min(t-1,q)}{k-1}\} \\
& -(1-\alpha)w\sum_{t=1}^{N}1_{(t\leq q)}(1-\sum_{k=1}^{t-1}\frac{(-1)^{k-1}}{n^k(n-1)^k}\binom{\min(t-1,q)}{k}) - C_{fixed}
\end{aligned}
$$

Applying the result for the complete network amounts to set $w = 0$ (assuming all the links are already there amounts to setting a uniform zero cost for building any link, with the consequence that any link needed to be built to realise an idea of trade will indeed be built). Hence $q = N$ and so $t \leq N$ for any period of the simulation. Using the binomial formula, we see that the previous expression for the profits collapses into:

$$
E(profits) = \alpha N(g_{dom} - c_{dom}^{comp}) + (1-\alpha)N(g_{cross} - c_{cross}^{comp}) - C_{fixed}^{comp}
$$

Subtracting the profits of the previous network shows that the benefits of joining the network can be expressed as:

$$
\begin{aligned}
& \alpha N(c_{dom} - c_{dom}^{comp}) + (1-\alpha)\{N(g_{cross} - c_{cross}^{comp}) - \sum_{t=1}^{N}(g_{cross} - c_{cross})\{\sum_{k=1}^{t}\frac{(-1)^{k-1}}{n^k(n-1)^k}\binom{\min(t-1,q)}{k} \\
& +1_{(t\leq q)}\sum_{k=1}^{t}\left(\frac{-1}{n(n-1)}\right)^{k-1}\binom{\min(t-1,q)}{k-1}\}\} \\
& -(1-\alpha)w\sum_{t=1}^{N}1_{(t\leq q)}(1-\sum_{k=1}^{t-1}\frac{(-1)^{k-1}}{n^k(n-1)^k}\binom{\min(t-1,q)}{k}) + C_{fixed} - C_{fixed}^{comp}
\end{aligned}
$$

6.3 Annex 3: the case of cost-recovery from network usage solely

Assumes the complete network cost-recovery is done entirely through transactions fees, and entails no fixed costs. That is, in Section 3.3 we have $c = 0$ and $c_{AB}^{comp} = c_{AB}^{comp}(n_0)$ is dependent of the trade volume (which is directly linked to the number of agents joining the network in our model). Hence $z = z(n_0)$ and the relation

$$
n_0 z - c/n_0 - c^A > 0 \quad (C)
$$

becomes

$$
n_0 z(n_0) - c^A > 0
$$

The previous aggregate costs c to recover is thus included in each transaction realised in the new network and depends thus of the size n_0 of the new complete network. More precisely:

$$c^{comp}_{AB\,new} = c^{comp}_{AB}(n_0) = \frac{c}{N\left(\frac{n_0^2}{n^2}\right)} + c^{comp}_{AB\,old}$$

Indeed, the ratio $\frac{n_0^2}{n^2}$ is nothing else than the expected portion of trade being settled entirely in the network (this is the probability, in our model, of an idea of the trade of type AB with both A and B agents having joined the network to be selected at a given period; since we run the model on N periods $N\left(\frac{n_0^2}{n^2}\right)$ is thus the expected number of transactions settled in the network). Now it can easily be checked that replacing $c^{comp}_{AB\,new}$ by this expression in $n_0 z(n_0) - c^A > 0$ gives exactly the condition (C), and thus the same values for n_l^A. This does not need to be the case in general, when some agents settle more trades than others on average.

6.4 Annex 4: An informal discussion concerning the dynamics implied by the one-stage game

In what follows we give a glimpse of the adjustment process as implied by the dynamic, myopic best-response version of our previous game. Such a game is defined by playing the previous game repeatedly a certain number of times, where each agent plays its best-response, given its belief n_0 about the number of agents joining. Consider the case described in Figure 4 and assume the agents, because of imperfect information, had formed the common belief that $n_0 = 1$. For example, agent A thought agent B would join, while agents B and C thought agents A would join. Acting on this belief leads agent A to join while agents B and C stay out. Hence an equilibrium is of the static one-period game (hence of the repeated game) is immediately reached with $n_0 = 1$. The dynamic game is then very similar to the static one, player A belief was wrong but with no detrimental consequences for any player. The Nash equilibrium reached is sub-optimal as it is Pareto-dominated by the equilibrium where all the three players join. But a distinct possibility for the second round of this game is that agent B and C both adjust their believes and expect the other to enter the network, which prompt the two of them to do so (self-fulfilling prohphecy). Without these strong believes there would be a coordination problem among agents there, because agent B needs to be sure agent A will enter to make the complete network solution competitive, and vice-versa. Furthermore, what would happen if n_l^A was not between 0 and 1 but between 1 and 2? After its first period move agent A would have been better not joining the new network, while in the second step agent B and C would need the insurance that agent A will not leave the network, on top of coordinating their moves, to make their decision to enter profitable. Reaching the Pareto efficient Nash equilibrium on the second period of the game now requires both the commitment of agent A to stay in the network and the joint coordination of agent B and C to enter.

Remark: One needs to modify the rules of the iterated game to model better the decision to leave. A realistic modification could be that a given agent A adaptation costs c^A when the agent first enter in the network, since these adaptation costs should be seen as an upfront payment paid only once and not as a on-going fee. Hence, a agent having made the wrong decision and paid the adaptation costs once would be more reluctant to leave the network after that and the network naturally becomes more attractive to it in the second period. For example in the case described by Figure 8 , where at the first step agent A joined the new network while agent B did not, it will become more probable that in the next period agent A will stay and agent B will join (hence making n_0 to go to $2 > n_l^B$), rather than agent A would leave.

To illustrate further the possible need for coordination among agents as well as greater transparency concerning agents intentions (and adaptation costs), we consider three different cases: all three cases starts with a pessimistic belief about the number of agents joining in the first wave: $n_0 = 1$. The first case evolves naturally, without the need for synchronisation between agents, towards the socially optimal

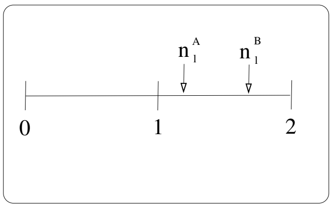


Figure 8: Agent B would be more willing to join than agent A to leave.

Nash equilibrium. The second case illustrates a coordination problem and the last the situation where a complete network cannot and should not be reached, since it is not a optimal solution.

Consider the case illustrated by Figure 9, with common initial belief $n_0 = 1$, and where A and B have entered. One might ask how in real life a common belief could be formed while some agents ignored others would enter: consider for example the simple case where A was not aware of B entrance and vice-versa, and where other market participants thought only A would join and not B (for example B had given wrong signals concerning its adaptation costs c^B, indicating a higher c^B than the real one). At the next step $n_0 \geq 2$ and agent C would join. Then at the next period $n_0 = 3$ and D will join. The process goes on until all agents belong to the complete network. No coordination was needed to reach this efficient equilibrium where all agents have joined.

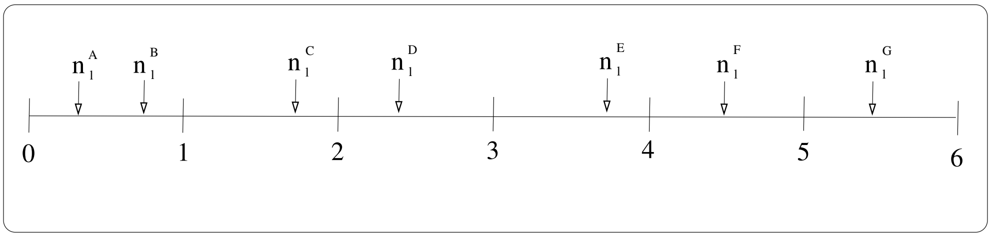


Figure 9: Case without any synchronization problem in the dynamic game.

Now alter the repartiton of the quantities n_l^A, n_l^B... such that we are in the situation depicted in Figure 10 below, with agent A and B having joined:

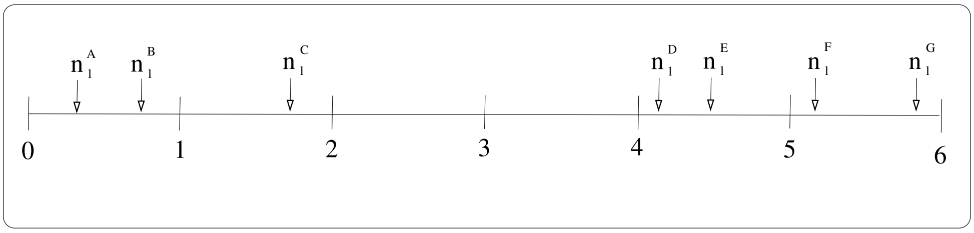


Figure 10: Situation requiring coordination from a subset of agents.

The beginning of the dynamics are the same: agent C joins and $n_0 = 3$. Now we have a synchronisation problem: indeed coordination would be needed between agent D and E to make it advantageous for them to join: because of their relatively high adaptation costs they need to bring both their volumes to make it more profitable for them to join than opt out. Moreover, so long that these agents have not solved their coordination problem and joined, there is little hope for agents F or G, which have higher adaptation costs, to join. Hence, a coordination problem among a strict subset of agents could put in jeopardy the evolution of the whole process towards the socially efficient Nash equilibrium, which is the equilibrium where all agents have joined.

Last, but not least, some subset of agents might have adaptation costs so high that there will not be a NE where all agents join. In Figure 11 below, the equilibrium resulting from the same set of initial assumptions is the one where only agents A, B, and C joined, while agent D and E stay outside the network:

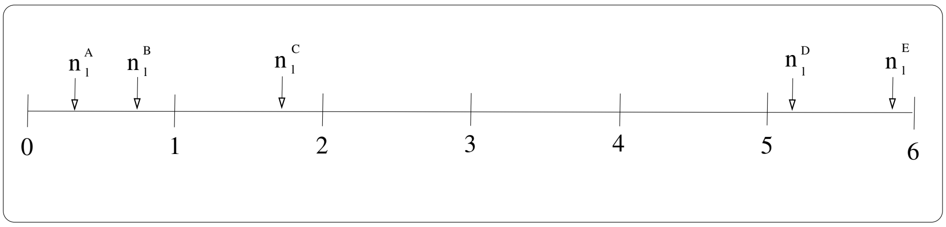


Figure 11: Situation where all agents joining would not result in a Nash equilibrium.

Without other agents in the market than the five agents represented here, and with no growth of their volumes, it is not profitable for agents D and E to join at any period. Note this model does not take into account the dynamic effects of the network on volume growth, nor the potential of new entrants. Now suppose there is another agent F, with $n_l^F = n_l^D$ for example. Because its adaptation costs are too high for him to join (that is, $4 < n_l^F$), it opts to stay out. Assume furthermore than the other agents believe (falsely) that it will join. Then they may want to synchronise their own entrance in the network to join. In such a case in the next period both agent D and E will have joined, contrary to agent F- whose only appearance made them join in the first place. Because of the additional volume now brought by the two recently joining agents, it is now more advantageous for agent F to join This illustrate how the simple participation of more agents in the network-related discussions could actually prompt other agents to join (on false believes) and allow to reach a better Nash equilibrium. Also note, in the three cases, the tendency of the process to drift to the right, that is, towards a bigger volume being dealt through the network, because adaptation-costs are seen as an upfront payment that can hardly be taken back once paid (see previous Remark).

6.5 Annex 5: proofs of various graph theoretic theorems

Proposition 15 If $w < \frac{g-c}{2}$ then any stable network is strongly connected and its diameter satisfies $\Delta \leq \min(k_0 - 1, \frac{\frac{g+c}{2}+w}{\frac{g+c}{2}-w})$.

Proof. We first prove that any stable network is strongly connected, and then derives the upper-bound for its diameter. Let G be a stable network and suppose it is not strongly connected. Let A and B be two nodes such that there exists no path going from A to B. Building the edge (A, B) would incur a cost of w to node A but an additional profit of at least $\frac{g-c}{2}$. Indeed, the trade with B itself already provides A with a profit of $\frac{g-c}{2}$, to which may be added other intermediation fees from the possible admissible paths using the newly created edge (A, B). Since $w < \frac{g-c}{2}$ the node A would be willing to build (A, B). This contradicts the stability of the network G.

Consider a minimum path P of length Δ, $P = A_0, A_1, A_2, ..., A_\Delta$ and assume $\Delta \geq 2$ (the case where $\Delta \leq 1$ can be readily checked since for network with at least two nodes, $w < \frac{g-c}{2}$ implies $\Delta = 1$ and $\frac{\frac{g+c}{2}+w}{\frac{g+c}{2}-w} > 1$ since $w > 0$, while by assumption $k_0 \geq 2$.) Hence by minimality of P, $A_0 A_\Delta$ is not an edge. Since $w < \frac{g-c}{2}$, A_0 would certainly be willing to pay for creating an edge towards A_Δ if it could not use the minimum path $A_0 = A$, A_1, A_2,..., A_Δ for carrying out trades with A_Δ, in which case we would have a contradiction with the minimality of P. Stability thus implies that this path can be used, hence that its length Δ is at most $k_0 - 1$. Also, stability requires that the benefits derived by A_0 from trading with A_Δ by using this minimum path is no less than the benefits it would get if it chose to establish a direct link with A_Δ, *idem est* that $\frac{g-\Delta c}{\Delta+1} \geq \frac{g-c}{2} - w$, which provides the condition $\Delta \leq \frac{\frac{g+c}{2}+w}{\frac{g+c}{2}-w}$. Hence $\Delta \leq \min(k_0 - 1, \frac{\frac{g+c}{2}+w}{\frac{g+c}{2}-w})$. ■

Notice in passing that another (weaker) necessary condition for the minimum path to be able to carry out trade is that $g - \Delta c > 0$ ie $\Delta \leq g/c$. Hence the above results also contains the somehow simpler condition that $\Delta \leq g/c$.

Proposition 21 If $w > \frac{g-c}{2} + 2(n-2)\frac{g-2c}{3}1_{(k_0>2)} + (n-2)(n-3)\frac{g-3c}{4}1_{(k_0>3)} =: l_0$, then the empty network is the only strict stable network (and the only strict Nash equilibrium).

Proof. Suppose $g \leq 2c$. Without loss of generality $k_0 \leq 2$ and the assumption becomes that $w \geq \frac{g-c}{2}$. This implies that the direct gains for A from establishing a link towards B are $\frac{g-c}{2} - w < 0$. Hence no node A will ever choose to maintain a link. This results in a empty network.

Suppose $g > 2c$ and that $w > l_0$. It is enough to show that l_0 is the maximal additional profit a node A can hope for by establishing a link. Since the costs w are superior to l_0 it will then follow that no node A will ever want to maintain a link, and that the empty network is the only NE for such parameter constellations. To show that l_0 is an upper bound for the additional profits derived by establishing a single edge AB, we start from any graph G and prove by successive reductions than the additional profit a node A can derive by establishing an edge AB cannot be more than the additional profit node A would derive from edge AB in a particular graph K_0, whose structure is known and in which A derives from AB an additional profit of precisely l_0.

Additional profits for A derived from establishing a new edge AB stems from:

- the direct trade with B, which amounts to at most $\frac{g-c}{2}$ – this upper-bound corresponding to the case where there were no admissible path going from A to B prior to the establishment of the link (A, B).

- the profits from trades involving more than two intermediaries and using a chain P of intermediaries which finishes with the edge AB. Assume such chain contains more than three intermediaries, and let $P =: C...DAB$. Then A derives less profits in such a graph from trades CB than it would if C was a in-neighbours of A, the reason being it has to share the profits from trade CB with more intermediaries (for example, with D). Hence A obtains no less profits in the graph G' obtained from G by adding all edges pointing towards A than in G itself. These profits amounts to $(n-2)\frac{g-2c}{3}$ in G'.

- the profits from trades involving more than two intermediaries and using a chain P of intermediaries which starts with the edge AB. Similarly, A obtains no less profits in the graph G'' obtained from G' by adding all edges starting from the node B than in G'. These profits amounts to $(n-2)\frac{g-2c}{3}$ in G''.

- the profits from trades involving more than three intermediaries and using a chain P of intermediaries in which both A and B are strict intermediaries. Note these profits can only be positive if $k_0 > 3$. Also, these profits are maximal when no node in $V(G)\backslash\{A, B\}$ is linked to another node of $V(G)\backslash\{A, B\}$. Indeed, in such a configuration any node in $V(G)\backslash\{A, B\}$ has to use the node A to carry out a trade with another node from $V(G)\backslash\{A, B\}$. Hence, A obtains no less profits in the graph G_0 obtained from G'' by deleting all edges between two nodes in $V(G)\backslash\{A, B\}$ than in G''. These profits amounts to $(n-2)(n-3)\frac{g-3c}{4}1_{(k_0>3)}$ in G_0.

It results from the above that $\frac{g-c}{2} + 2(n-2)\frac{g-2c}{3} + (n-2)(n-3)\frac{g-3c}{4}1_{(k_0>3)}$ is the additional profit derived by node A by establishing the link AB in the graph K_0, and that it is an upper-bound for the additional profit derived by node A by establishing the link AB in the initial graph G. Since this holds for all G, we have a general upper-bound. ■

Proposition 24: Assume $g > 2c$. Then stable complete stars are also Nash equilibria.

Proof. Let G be a complete stable star.

By symmetry of the network, the center's payoff depends uniquely on the number of links it creates towards the other nodes. Assume it creates k links, and let f be its payoff. Then

$$f(k) = \{-kw\} + \{k\left(\frac{g-c}{2}\right) + (n-1)\left(\frac{g-c}{2}\right)\} + \{k(k-1)\left(\frac{g-2c}{3}\right) + (n-1-k)k\left(\frac{g-2c}{3}\right)\}$$

Indeed, the first term between brackets is the costs of creating the links, the second term between brackets the profits from direct transactions and the third term between brackets the profits from intermediation. $f(k)$ can be rewritten as:

$$f(k) = k(-w + \left(\frac{g-c}{2}\right) + (n-2)\left(\frac{g-2c}{3}\right)) + (n-1)\left(\frac{g-c}{2}\right)$$

Since $-w + \frac{g-c}{2} + (n-2)\left(\frac{g-2c}{3}\right) \geq 0$ the maximum is obtained for the maximum value of k, that is, $n-1$, which means creating all possible links. Note that, in the above reasoning, instead of computing the total profits, we could have only computed the additional profits obtained by establishing k links, that is. $k(-w + \frac{g-c}{2} + (n-2)\left(\frac{g-2c}{3}\right))$. Because $-w + \frac{g-c}{2} + (n-2)\left(\frac{g-2c}{3}\right) \geq 0$ the maximum is obtained for the maximum value of k, which is $n-1$, and we conclude similarly. This approach being somehow simpler, we proceed in this fashion for the leaves, computing the *additional* benefits of having a given number of links in each case.

For computing one of the leaf A's best-response we will distinguish two cases:

In the first case the leaf A plays a strategy where it builds a link towards the center and k other links, for $k \in \{0, 1, ..., n-2\}$. Its *additional* profits from building k edges towards other leaves is: $k(-w + \frac{g-c}{2} - \frac{g-2c}{3})$. Because $\frac{g+c}{6} - w \leq 0$, the maximum is obtained for $k = 0$, which means building no link except towards the center.

In the second case it does not link to the center but to k other links, for $k \in \{0, 1, ..., n-2\}$. If $k \geq 1$ then the additional profit obtained by A if it severs all existing links and connects to the center instead is: $(k-1)w - (k-1)\frac{g-c}{2} + (k-1)\frac{g-2c}{3} + (n-2-k)(\frac{g-2c}{3} - \frac{g-3c}{4}1_{(k_0>3)})$. Indeed, the first term represents the costs saved from having to pay only for a single link instead of k links, the second the loss of profits from direct trades towards other nodes, the third the compensating gains from indirect trade steping up to replace the previously direct trades, and the last term the gains in what was before, and still is, indirect trade ($1_{(k_0>3)}$ is 1 if $k_0 > 3$ and 0 otherwise). This amounts to $(k-1)(w - \frac{g+c}{6}) + (n-2-k)(\frac{g-2c}{3} - \frac{g-3c}{4}1_{(k_0>3)})$, a non-negative quantity since $w - \frac{g+c}{6} \geq 0$ and $\frac{g-2c}{3} - \frac{g-3c}{4}1_{(k_0>3)} > 0$. Hence strategies sending edges towards leaves and not towards the center are dominated by the strategy of just linking to the center.

The only check left is the suboptimality of not building any edge. Building a single edge towards the center would bring an additional profit of $-w + \frac{g-c}{2} + (n-2)\left(\frac{g-2c}{3}\right) \geq 0$, where $\frac{g-c}{2}$ stems from direct trade of the leaf A with the center and $(n-2)\left(\frac{g-2c}{3}\right)$ from indirect trades starting at A. This concludes the proof. ■

Proposition 28:Assume $g > 2c$ Then a d-regular network is stable if, and only if,

$$2d\left(\frac{g-2c}{3}\right) + \left(\frac{g+c}{6}\right) \leq w \leq \frac{g-c}{2} + 2d\left(\frac{g-2c}{3}\right)$$

The proof of Proposition 28 relies both on Theorem 27 and on the following Lemma:

Lemma 35 *Suppose $g > 2c$. Suppose $N^+(A) \cap N^-(B) \neq \varnothing$. Then node A earns, by creating the link (A, B), an additional profit of*

$$(|N^-(A)| - |N^-(A) \cap N^-(B)| + |N^+(B)| - |N^+(A) \cap N^+(B)|)\left(\frac{g-2c}{3}\right) + \frac{g+c}{6} - w$$

Proof. The additional profits of creating such a link are (see Figure 12):

1) those derived from intermediating transactions using chains of type DAB where D is any node in $N^-(A)$ but not in $N^-(B)$: $(|N^-(A)| - |N^-(A) \cap N^-(B)|)\left(\frac{g-2c}{3}\right)$,

2) those derived from intermediating transactions using chains of type ABC where C is any node in $N^+(B)$ but not in $N^+(A)$: $(|N^+(B)| - |N^+(A) \cap N^-(B)|)\left(\frac{g-2c}{3}\right)$,

3) the additional profit $\frac{g-c}{2} - \frac{g-2c}{3} = \frac{g+c}{6}$ of carrying out directly trades (A, B) instead of using one strict intermediary.

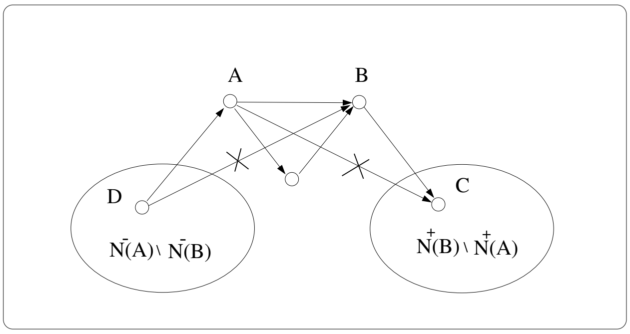


Figure 12: Illustration of the different chains of intermediaries alluded to in the proof

The costs are w. The result follows. ■

We can now proceed to prove that the condition of Proposition 28 is necessary and sufficient:

Suficiency: Suppose (A, B) is not an edge. Since G is a uninetwork we thus have a node C such that the path ACB allows to settle trades of type (A, B), and $C \in N^+(A) \cap N^-(B) \neq \varnothing$. We have, since $|N^-(A)| = d = |N^+(B)|$:

$$(d - |N^-(A) \cap N^-(B)| + d - |N^+(A) \cap N^+(B)|)\left(\frac{g-2c}{3}\right) + \frac{g+c}{6} \leq \left(\frac{g-2c}{3}\right)(2d) + \left(\frac{g+c}{6}\right) \leq w$$

Hence by Lemma 35 node A does not want to create this edge.

Suppose (A, B) is an edge. Deleting it would incur a gain of precisely $w - \frac{g-c}{2} - d\left(\frac{g-2c}{3}\right) \leq 0$, since there is no other path than (A, B) that can be used by A to trade with B nor other path than ABC which can be used to trade with C for any $C \in N^+(B)$. Hence A would be worse off deleting (A, B). Hence the condition is sufficient.

Necessity: The most important part of the proof is indeed to establish the necessity of the condition, as it will allow us to conclude the network we are looking for does not exist. Indeed, for such a network to be stable it is necessary that no node A wants to create a link towards B for any $AB \notin E(G)$. Let $A \in V(G)$. Let A' be the unique node of G which is both an in-neighboor and an out-neighboor of A. The existence and unicity of A' is assured by Theorem 27 and the definition of a line network of a complete network (Definition 26). Since $d \geq 2$ there exists v in $N^-(A)\backslash\{A'\}$. Hence $vA \in E(G)$.

Suppose $N^-(A) \cap N^-(v) \neq \varnothing$. Then there exists z in $N^-(A) \cap N^-(v)$ and zvA and zA are two distinct paths going from z to A, a contradiction with the definition of uninetwork. Hence $N^-(A) \cap N^-(v) = \varnothing$. Similarly, we can show that $N^+(A) \cap N^+(v) = \varnothing$. Since we suppose that the network G is stable, and that $Av \notin E(G)$, node A should not want to create an edge towards v. Applying then Lemma 35 to the pair (A, v) yields the condition $2d\left(\frac{g-2c}{3}\right) + \left(\frac{g+c}{6}\right) \leq w$.

For the other part of the inequality, notice again that if (A, B) is an edge in the network, deleting it would incur a gain of precisely $w - \frac{g-c}{2} - 2d\left(\frac{g-2c}{3}\right)$, since there is no other path than (A, B) that can be used by A to trade with B nor other path than ABC which can be used to trade with C for any $C \in N^+(B)$, nor other path than DAB which can be used by $D \in N^-(A)$ to trade with B. Assuming this quantity to be nonpositive is thus necessary. This concludes the proof of the necessity of the condition.

Theorem 31: Assume $g > 2c$ and $w > \frac{g+c}{6}$. Let G be a strict Nash equilibrium. Then there is a threshold $d_0 \leq \frac{3}{g-2c}(w - \frac{g+c}{6}) + 1$ such that, if G has a partition $\{S_i,\, i \in I\}$ of complete stars, where s_i

denotes the center of star S_i, and such that, for all $i \in I$, we have:

$$|(N^+(s_i) \cap S_i) \backslash \bigcup_{w \notin S_i} N^+(w)| > d_0 \quad (C)$$

then G is a multistar with partition $\{S_i,\, i \in I\}$ of complete stars.

Proof. Let $\mathcal{S}_i := N^+(s_i) \backslash \bigcup_{w \notin S_i} N^+(w)$ be the set of vertices of S_i that can only be reached through s_i.

Let $j \in I$. Let $v \in V(G) \backslash \{s_j\}$.

a) Assume first that $v \notin \mathcal{S}_j$. If v did not had a link towards s_j, then creating it would provide v with an additional profit of at least

$$-w + \frac{g-c}{2} - \frac{g-2c}{3} + |\mathcal{S}_j| \left(\frac{g-2c}{3} \right)$$

Indeed, w is the cost of establishing a new link towards s_j, $\frac{g-c}{2} - \frac{g-2c}{3} = \frac{g+c}{6}$ is a lower bound for the additional gains from direct trade with s_j and the last term represents the additional gains from intermediating trade with the $|\mathcal{S}_j|$ nodes that can only be accessed only through s_j. It is thus a lower bound for the additional (strict) intermediation profits.

b) If $v \in \mathcal{S}_j$ then the additional profit of creating a link to s_j will be at least

$$-w + \frac{g-c}{2} - \frac{g-2c}{3} + (|\mathcal{S}_j| - 1) \left(\frac{g-2c}{3} \right)$$

This lower bound is lower than in a) because v does not trade with itself in the model. This last lower bound, and hence both lower bounds, are positive as soon as $|\mathcal{S}_i| > \frac{3}{g-2c}(w - \frac{g+c}{6}) + 1 =: d_0$ which is the case by assumption. This contradicts the stability of a Nash equilibrium. Hence every $v \in V(G) \backslash \{s_j\}$ has a link to the center s_j of S_j, and this result holds for all $j \in I$.

c) Let $v \in V(G) \backslash \{s_i,\, i \in I\}$. Let k_i be the number of links that v built towards $V(S_i)$. By the above, we already know that v is linked to $s_i \in S_i$, hence $k_i \geq 1$. Assume by contradiction that $k_i > 1$ for at least one $i \in I$. Then $\sum_{i \in I} k_i - |I| > 0$. We can write the additional profit of node v derived by these extra-edges as a function of the k_i:

$$-w(\sum_{i \in I} k_i - |I|) + (\sum_{i \in I} k_i - |I|)(\frac{g-c}{2} - \frac{g-2c}{3}) = (\sum_{i \in I} k_i - |I|)(-w + \frac{g+c}{6})$$

which is strictly negative since $\sum_{i \in I} k_i - |I| > 0$ and $-w + \frac{g+c}{6} < 0$, a contradiction with G NE. Hence $k_i = 1$ for all $i \in I$ and v sends only a single link towards each star S_i, which is directed towards the center s_i of S_i.

d) Let now $i_0 \in I$ and consider s_{i_0}. The same argument as in (c), where for $i \in I \backslash \{i_0\}$ we define k_i to be the number of links that s_{i_0} built towards $V(S_i)$, can be used by considering $\sum_{i \in I \backslash \{i_0\}} k_i$ instead of $\sum_{i \in I} k_i$, and $I \backslash \{i_0\}$ instead of I. This concludes the proof. ■

Conversely, it can be checked that this unique possible structure is indeed a NE when we had the now well-known necessary condition that $w \leq \frac{g-c}{2} + (n-2)\frac{g-2c}{3}$, together with a sufficiently high degree for each center of the stars.

Proposition 32: Let G be a multistar with p*artition* $\{S_i,\, i \in I\}$ *of complete stars* S_i. Assume moreover than $d^+(s_i) - |I| + 1 > d_0$ for all $i \in I$, and that $\frac{g+c}{6} < w < \frac{g-c}{2} + (n-2)\frac{g-2c}{3}$. Then G is a strict Nash equilibrium.

Proof. The condition $d^+(s_i) - (|I| - 1) > d_0$ is precisely the condition (C) of the previous theorem once the overall structure has been deduced. The same arguments thus allow to prove that, given the other nodes strategies, each node strategy is a best-response among the strategies where the centers of the stars do maintain the links towards their leaves. More precisely, the *additional* benefit of the center

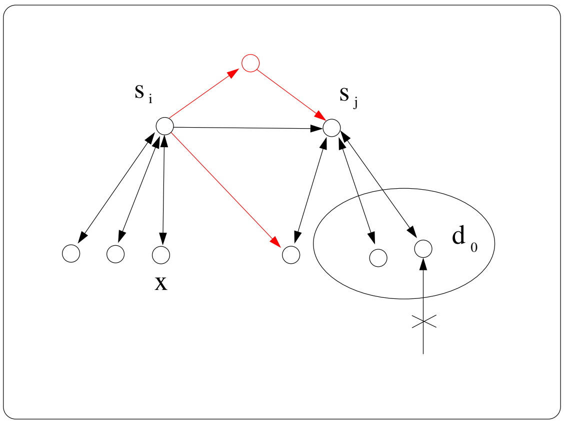


Figure 13: Illustration of the proof

s_i of forming k links towards its $d^+(s_i) - (|I| - 1)$ leaves is, assuming other nodes do not deviate from the strategy profile:

$$(-w + \frac{g-c}{2})k + (k(k-1) + k(n-1-k))\frac{g-2c}{3}$$

Indeed, $-wk$ is the cost of maintaining these links, $\frac{g-c}{2}k$ the additional profit from direct trade, $k(k-1)\frac{g-2c}{3}$ the intermediation profits from being the strict intermediary of trade between vertices both in $N^+(s_i) \cap S_i$ and $k(n-1-k)\frac{g-2c}{3}$ the intermediation profits from nodes outside S_i towards nodes in $N^+(s_i) \cap S_i$. This amounts to

$$k(-w + \frac{g-c}{2} + (n-2)\frac{g-2c}{3})$$

This additional benefit is increasing with the number of links created towards the leaves if, and only if, $w < \frac{g-c}{2} + (n-2)\frac{g-2c}{3}$. By assumptions this condition is satisfied. Hence each center maintains a link towards each of its leaves at equilibrium.

Creating links towards leaves of other stars would result in a additional profit of $-w + \frac{g+c}{6} < 0$ for the center s_i. Hence no center will maintain such links at equilibrium.

Any vertex (center or leaf) not in $\mathcal{S}_j$ linking towards the center s_j of another star S_j will benefit from an additional profit of at least

$$-w + \frac{g-c}{2} - \frac{g-2c}{3} + |\mathcal{S}_j|\left(\frac{g-2c}{3}\right)$$

while any vertex in $\mathcal{S}_j := N^+(s_j) \setminus \bigcup_{w \notin S_j} N^+(w)$ linking towards the center s_j of star S_j will benefit from an additional profit of at least

$$-w + \frac{g-c}{2} - \frac{g-2c}{3} + (|\mathcal{S}_j| - 1)\left(\frac{g-2c}{3}\right)$$

The definition of l_0 insures doing so brings additional benefits.

We are only left with checking that linking directly to another leaf is not beneficial for a leaf, which is true since $-w + \frac{g+c}{6} < 0$. We can thus conclude that the network G is a NE. ■

Theorem 36 *Assume $g > 2c$ and $w > \frac{g+c}{6}$. Let G be a strict Nash equilibrium. Then there is a threshold $d_0 \leq \frac{12}{g+c}(w - \frac{g+c}{6}) + 1$ such that, if G has a partition $\{S_i, i \in I\}$ of complete stars, where s_i denotes the center of star S_i, and such that, for all $i \in I$, we have:*

$$|(N^+(s_i) \cap S_i) \setminus \bigcup_{w \notin S_i} N^+(w)| > d_0 \quad (C)$$

then G is the unique network such that 1) it contains the $\{S_i,\, i \in I\}$ as subnetworks, and 2) each of the node $v \notin \{s_i,\, i \in I\}$ is linked to precisely all the $\{s_i,\, i \in I\}$. 3) Each center s_i is linked to precisely all the nodes in $S_i \cup \{s_j,\, j \in I \backslash \{i\}\}$.

Proof. The proof is similar to the one of Theorem 31, and thus omitted there. ■

Remark 37 *We suspect the bound d_o of Theorem 31 and Theorem 36 not to be optimal, since it is increasing with w, which is contrary to the intuition that higher costs of building edges would make intermediation through an oligopolistic more profitable. Hence, there may be a better bound than the one given by the above theorem.*

Remark 38 *In case $3c < g$ which is the least to assume in the 4 intermediaries case, the upper bound for d_0 is lower (thus better) in Theorem 31 than in Theorem 36*

Theorem 33: Let $h(c,w) = 5n/4 - 1 - (n-2)w/c$. Let i be the closest positive number to $h(c,w)$. Then among all the multistars, the ones with presicely i centers are the most efficient networks.

Proof. First notice that since all the latent profits are realised, aggregate efficiency ranking can be obtained by simply inverting the ranking of the costs of the links: the parameter g will thus drop from all the conditions derived here.

Simple combinatorics arguments allow to show that the number of edges in a multistar whose set of centers is I is simply $n + |I|\,(n-2)$. Each of this edge cost w to build.

In a multistar there are $(n - |I|).\,|I| + (n - |I|) + |I|\,(|I| - 1) = |I|\,(n-2) + n$ direct trades. This is because there are $(n - |I|).\,|I|$ direct trades from a leave to a center, $(n - |I|)$ direct trades from a center to a leaf (as there is a unique edge from I to any if the $n - |I|$ leave, and $|I|\,(|I| - 1)$ edges linking the different centers of the multistar. Each of the direct trade is realised at a cost c.

In a multistar there are $(n - |I|).(n - |I| - 1)$ indirect trades, which follow routes of length 2. This is because each leaf indireclty trade with each other leaf of the multistar. Each of the indirect trade costs $2c$.

Hence the total costs associated to a multistar with set of centers I is:

$$(n + |I|\,(n-2))w + (|I|\,(n-2) + n)c + (n - |I|).(n - |I| - 1)2c$$

Define $f(x) = (n + x(n-2))w + (x(n-2) + n)c + (n - x).(n - x - 1)2c$. f is a polynomial of degree 2 whose minimum is attained in $h(c,w)$, hence the result. ■

Theorem 39 *Assume $c < w$. Let G be a multistar with partition $\{S_i,\, i \in I\}$ of complete stars S_i of center s_i and let j be such that $s_j = \max(s_i,\, i \in I)$. Then a star with the same number of vertices as the multistar G Pareto-dominates G as soon as*

$$6(g-c-w)(|I| + s_i - 2) + 3(g-c)(n - |I| - s_i - 1) + 2(g-2c)\{-2(n-2) + (s_i - 1)(2n - s_i - i - 1\} \text{ for all } i \neq j$$

$$(s_i - 1)(n - s_i)(g - 2c) > 3($$

Proof. Leaves are better off in a star than in a multistar since in a multistar each leaf does $|I| - 1$ more direct trades than in a star which save the transaction costs of $c(|I| - 1)$ but building those additional links are paid by the leaves concerned hence a cost of $c(|I| - 1)$. There is this a net additional cost of $(w - c)(|I| - 1) > 0$ in a multistar compared to the star for the same profits being realised.

The center of a star is better than one of the center of the multistar s_i if, and only if, $(s_i - 1)(n - s_i)(g - 2c)/3 > (n - s_i - i + 2)w$. Indeed using simple combinatoric argumentswe see that the center of the multistar, compared to the center of a star, (a) misses $(s_i - 1)(n - s_i)$ paths of length 2 which realise profits from one leaf of their star S_i to leaves not in S_i; (b) it builds $(n - (s_i - 1) - (|I| - 1))$ less edges than the center of a star hence a saving of $(n - s_i - i + 2)w$. Writing that profit losses are higher than cost cutting benefits is equivalent to the second condition.

The profit of a center of a star s_i is the sum of the cost of its $|I|+|S_i|-2$ edges: $-w(|I|+|S_i|-2)$, the profits from direct trades to and from the $|I|-1$ other centers and its $|S_i|-1$ leaves: $2.(|I|+|S_i|-2)\frac{g-c}{2}$, the direct profit from trades from leaves outside S_i: $(n-|S_i|-|I|+1)\frac{g-c}{2}$, and the indirect profits. The indirect profits can be further decomposed into: the indirect profits from intra-S_i trades: $(|S_i|-1)(|S_i|-2)\frac{g-2c}{3}$, from trades from outside S_i to a leaf in S_i: $(n-|S_i|)(|S_i|-1)$, from trades from a leaf in S_i to a node outside: $(|S_i|-1)(n-|S_i|-|I|+1)$. This results in an overall profit of

$$(g-c-w)(|I|+|S_i|-2)+\frac{g-c}{2}(n-|I|-|S_i|+1)+\frac{g-2c}{3}((|S_i|-1)(|S_i|-2)+(|S_i|-1)(2n-2|S_i|-|I|+1)$$

Writing that those profits are lower than those of the leaf of star is equivalent to the first condition.

Under the first condition for all $i \neq j$ we have that all the centers of a multistar with the possible exception of the center eaning the most profits earns less profits in a multistar than the leaves of a simple star. Under the second condition we have that the center of the star earns more than the center eaning the most profits in the multistar. Since the leaves are better off in a star than being leafs in a multistar anyway, any relabelling of the centers of one of the center of the multistar which earns the most into the center of the star, and of the other centers into the name of simple leaves of the star, will do. ■

Remark 40 *The profit of a leaf x of the multistar G in S_i is $-w|I|+\frac{g-c}{2}|I|+2\frac{g-2c}{3}(n-|I|-1)+\frac{g-c}{2}+\frac{g-2c}{3}(|I|-1)$. Indeed, $-w|I|$ is the cost of the leaf x to maintain a link towards each of the centers, $\frac{g-c}{2}|I|$ the benefits from direct trade towards those centers, $2\frac{g-2c}{3}(n-|I|-1)$ the benefits of trading to and from the $n-|I|-1$ other leaves of the multistar, $\frac{g-c}{2}$ the benefits from direct trade from s_i and $\frac{g-2c}{3}(|I|-1)$ for indirect trades from the s_j with $j \neq i$ to x. This can be rewritten as $|I|(-w+\frac{g-c}{2}-\frac{g-2c}{3})+\frac{g-c}{2}+(2n-3)\frac{g-2c}{3}$. Since $-w+\frac{g-c}{2}-\frac{g-2c}{3}=-w+\frac{g+c}{6}<0$, the lower $|I|$, the higher the profits for the leaf x. Economically, this expresses its preferences to only have to pay for one single link to get access to the whole network, in contrast to having to establish a link towards each center of a multistar. Note that, of course, by plugging $|I|=1$ we find again the profit of a star of size n.*

Article 5: Models of default rates: current models review and an alternative approach using Hidden Markov Chains

1 Motivation of the present paper

1.1 Credit claims in the European context

The Eurosystem is currently accepting both individual credit claims and pools of credit claims as collateral for its monetary operations. Individual credit claims have always been accepted in the general collateral framework. In December 2011, the Governing Council decided to allow National Central Banks (NCBs) to temporarily extend eligibility to other performing credit claims based on NCB-specific eligibility and risk control frameworks, albeit on a non-loss-sharing basis. These credit claims eligible on the basis of such specific frameworks are referred to as the Additional Credit Claims (ACC). They can be individual credit claims or whole pools of credit claims. The Governing Council approved ACC frameworks, including their associated haircuts. Hence, the three main features of ACC frameworks are that (1) they operate under a *non-risk sharing regime*, meaning that the potential losses in a given ACC framework will solely be borne by its NCB, (2) they were conceived as a *temporary measure* to expand collateral availability in the Eurosystem, and (3) entire pools of credit claims can be pledged. These pools can contain credit claims with a higher probability of default than the individual credit claims which are eligible as collateral through the general collateral framework. This is because diversification effects are taken into account at the pool level through the use of a model for default rates in pools of assets[1].

These NCB-specific frameworks are still subject to minimum risk-control requirements, in the sense that the Governing Council has set minimum eligibility criteria for credit claims to be accepted.

As of 29 May 2013, the resulting aggregate value of credit claims after haircuts pledged to the Eurosystem (both through the general framework and through the ACC frameworks) was 454 billion. This compares to 330 billion for ABSs, for example, and represents about 15% of the total amount of liquidity provided to the Eurosystem. Hence, the volumes are large, which increases the associated risks (in amount terms).

1.2 Scope of the paper

The present paper is targeting both the broad audience, which will gain a valuable, self-contained knowledge of the models described, and the more specialist audience, which will find many relevant details about specific models. The paper focuses on the modelling of defaults. This can serve a variety of purposes such as assessing quantitatively the credit risk associated to the collateral pledged to the Eurosystem by its counterparties. As such, it does not try to take into account other model features that would be necessary for the *pricing* of loans. As a consequence, it does not need to investigate the behaviour of recovery rates, or the additional subordination structures for creditor claims[2], or the waterfall structure for ABSs. All these layers can easily be added later, depending of the purpose of the model being built. The paper mentions some of them, when deemed relevant, but only to the extent where these concepts

[1] Nevertheless the maximum probability of default for any individual credit claim to be eligible for ACC was set to 1.5% by the Governing Council.

[2] Only un-subordinated (i.e. senior) credit claims, as accepted in the general collateral framework or in the ACC frameworks, will be considered. Also, it is assumed all obligators are distinct. Hence, in the case different loans of a pool have been granted to the same obligator, the models developed here should be applied to a synthetic loan portfolio obtained from the real portfolio by regrouping all loans granted to the same obligator in a single loan.

can be *directly* linked to the default rate modelling (for example, when time of prepayment can be made an inversely correlated quantity to the default time as in [29]).

The default rates models detailed in this paper are not specific to credit claims, and could as well be applied for corporate bonds. Nevertheless at the current juncture they are particularly wanted for credit claims, for the reasons detailed in the previous section. Models for both pools of credit claims and individual credit claims – still distinguishing by sector and country – are described and compared in the paper.

The contributions of this graduate project paper are twofold. First, it reviews the main default rate models used in the Eurosystem, as well as the links between them, using homogeneous notations for their mathematical descriptions. This allows a unified presentation of most of the credit models widely used. By emphasising each model assumptions and limitations, the paper aims at providing the reader with a better understanding of each model relevance. It also explain the links between these models. Second, it introduces the use of hidden Markov chain models (HMM) to estimate the probability of defaults and the credit cycle of credit claims in the European context. A Matlab library has been developed and documented for that purpose, and is available from the author on demand. HMM allows to estimate numerical probability of defaults for a given type of credit claims (for example, Spanish loans to medium enterprises).

1.3 Detailed outline of the structure of the paper

The first part of the paper describes, *inter alia*, the structural Vasicek one-factor model used in the Capital Requirement Directive of Basel II. Other widely used models, such as the statistical (Gaussian) co-dependency model, will be described in this first section. It will also describe a generalisation of the Gaussian co-dependency approach for modelling default rates [1], which allow to induce inverse correlation between default rates and prepayment rates through Monte-Carlo simulations, a very appreciated feature in, for example, the pricing of asset backed-securities (ABSs).

Each model is described in a way to provide the general reader with the main intuition behind it, such as the model underlying assumptions and shortcomings, while allowing the more specialist audience to understand the technical details involved. More generally, the present paper provides the general reader with a unified description, both in terms of terminology and in terms of mathematical notation, of the main core models of credit risk, and thoroughly explains the links between the models. To improve the clarity of the paper, all mathematical derivations are placed in Annex. These derivations follow the order of the paper, and should be considered as part of it.

It is shown how all the models described in this first section (the Vasicek, the Gaussian-copula and its generalisation [1]), ultimately result in a common risk factor impacting simultaneously all the loans of the pool involved. Hence, the question of knowing what really constitutes this risk factor arises. Specifying explicitly this risk factor by using a proxy variable, or by assuming a theoretical distribution for it, or even by inferring it from asset prices, implies model risks. Indeed the proxy arbitrarily chosen may be inappropriate.

In the second part of the paper we propose a different type of model which does not make any assumption about the risk factor but derives it only from the default data. This model is based on the theory of Hidden Markov chains, and has been successfully used in the context of US corporate bonds default data [14]. It is not specific to this asset class, but corporate bonds were probably used because of the easiest availability of default data. The hidden Markov model approach provides a current estimate of the credit risk cycle (low or high) together with the probability of defaults in each of the period of the cycle. Interestingly, the European Data Warehouse has been operational since January 2013, and will be collecting loan-level data on all ABSs pledged to the Eurosystem, as loan-level data has been made an eligibility criteria for ABSs. This data will allow, once it covers enough observations in a steady state of the economy and in a high default-risk state, to estimate directly the probability of default of a given sector/jurisdiction as well as the past credit cycle of the sector, which can be quite distinct from the business cycle. Because the implementation part is far from trivial, having available in a whole Matlab library for this purpose, properly documented, is a valuable asset. In the short-term, these algorithms will already be used for assets for which data exists such as corporate bonds. It is also used on a synthetic index built on the delinquency data obtained from Bloomberg.

2 Notation and preliminaries

This short section details preliminaries in terms of notations, basic probability facts, commonly used risk-measures, and discuss the concept of portfolio granularity. It is mainly needed for the first section of the paper, which gives a broad overview of the current use of default models. It is also useful as introductory reading for the literature of default model in general. The reader interested only in the Hidden Markov Chains model could skip this preliminary section and go directly to its relevant section.

2.1 Random variables for modelling default

2.1.1 General preliminaries

For any random vector $X = (X_1, ..., X_k)$ we denote by F_X its (multivariate) cumulative distribution function (cdf), that is,

$$F_X(a) = P(X_1 \leq a_1, ..., X_k \leq a_k)$$

for every real vector $a = (a_1, ..., a_k)$.

If F is a multivariate cdf we often denote by F_i its ith marginal. We denote by Φ the cdf of the standard normal law, that is, the Gaussian law with mean 0 and variance 1, and by ϕ its density.

We define the inverse of a univariate cdf F by:

$$F^{-1}(a) = \inf\{t,\ F(t) \geq a\}$$

In case F is invertible then F^{-1} coincides with the usual definition of the inverse of a bijective function. For $a \in [0, 1]$, the quantity $F^{-1}(a)$ is called the a-quantile of the distribution. If F is the cdf of a random variable X, $F^{-1}(a)$ is also known as the a-quantile of X.

For a given asset or loan i and time t denote by $\lambda_{i,t}$ the instantaneous hazard rate of loan i at time t. It is the average probability of default of a loan which has not defaulted until time t over a very short interval of time $[t, t + dt]$, hence the name "instantaneous". The $\lambda_{i,t}$ can be a deterministic quantity, or a random variable. Let τ_i be the time of default of loan i. The τ_i are modelled as (non-deterministic) random variables. Let $PD_i(t)$ be the cumulative distribution function of τ_i, that is,

$$PD_i(t) := F_{\tau_i}(t) = P(\tau_i \leq t)$$

2.1.2 Intensity models

A so-called intensity model links the cumulative probability of default $PD_i(t)$ to the instantaneous hazard rate $\lambda_{i,t}$ by:

$$PD_i(t) := P(\tau_i \leq t) = 1 - \exp(-\int_0^t \lambda_{i,s} ds)$$

In case PD_i is differentiable, the above formula can be easily proved from the definition of the hazard rate (Annex 6.1).

Most of the time, the $\lambda_{i,s}$ are in fact assumed non-stochastic. They are often even assumed constant: $\lambda_{i,s} = \lambda_{i,0} =: \lambda_i$ for all borrower i. In that case the intensity model reduces to:

$$PD_i(t) = P(\tau_i \leq t) = 1 - \exp(-t\lambda_i)$$

In practice, intensity model are very useful to transform a single point data $p_{\widetilde{t}}$ of the (cumulative) probability of default of debtor i over the time interval $[0, \widetilde{t}]$ into a whole term-structure of the cumulative probability of default $t \to PD_i(t)$, which is a function defined on the interval $[0, T]$ and of values in $[0, 1]$, for the chosen horizon T needed in the model, and can be seen as a distribution of the timing of default. Most of the time $p_{\widetilde{t}}$ is estimated from market data, using the excess spread on bonds or a Credit Default Swap (CDS) spread, or it can be taken from the rating agencies tables of probability of default, or inferred from the commercial bank borrower rating score. Deriving λ_i from $p_{\widetilde{t}}$ is done by setting the empirical quantity $p_{\widetilde{t}}$ equal to the theoretical quantity $1-\exp(-\widetilde{t}\lambda_i)$ where $\widetilde{t}$ is the time horizon of $p_{\widetilde{t}}$. For example the credit rating agencies tables indicate the cumulative probability of default for each rating class with an horizon of one year, hence $\widetilde{t} = 1$. For example one can look at the most liquid CDS when such instruments are available for the borrower, and very often these are the three or five year CDS. Hence $\widetilde{t} = 3$ or 5, depending of the most liquid CDS maturity. Solving for λ_i in:

$$p_{\widetilde{t}} = 1 - \exp(-\widetilde{t}\lambda_i)$$

yields

$$\lambda_i = -\frac{\ln(1 - p_{\widetilde{t}})}{\widetilde{t}}$$

Admittedly, other methods exist to derive the curve $t \to PD_i(t)$, for example using bonds of different maturities. In the context of credit claims from small borrowers, this is often not possible, hence a one point estimate $p_{\widetilde{t}}$ is used to derive λ_i, then the whole curve $t \to PD_i(t)$ is deduced through the intensity model.

Two simple mathematical facts are constantly used in the context of default modeling and simulations:

Lemma 1 *Let U be a uniformly distributed variable on $[0, 1]$. Let F_X be the cumulative distribution function of some random variable X. Then $F_X^{-1}(U)$ follows the same law than X, that is, $F_{F_X^{-1}(U)} = F_X$.*

Proof. $P(F_X^{-1}(U) \leq a) = P(U \leq F_X(a)) = F_X(a)$ for any real a since $F_X(a) \in [0, 1]$. ■

Lemma 2 *Let X be a random variable and F_X its cumulative distribution function. Then $F_X(X)$ is uniformly distributed on $[0, 1]$.*

Proof. $P(F_X(X) \leq a) = P(X \leq F_X^{-1}(a)) = F_X(F_X^{-1}(a)) = a$ for all real a, by definition of F_X. ■

2.2 Commonly used risk-measures

The Value At Risk (VaR) at the α confidence level of a given risky asset (which can be an individual asset or a portfolio), denoted by VaR(α), is the αth percentile[3] of the risky asset loss distribution L:

$$\text{VaR}(\alpha) = \inf\{l \,|\, P(L \leq l) \geq \alpha\} = F_L^{-1}(\alpha)$$

Intuitively, one could expect that the losses on the asset value would not exceed the VaR(α) of the asset more than $1-\alpha$ percent of the time[4], since VaR(α) is the smallest l such that $P(L > l) = 1 - P(L \leq l) \leq$

[3] When the loss distribution is not continuous, its upper α quantile $\inf\{l \,|\, P(L \leq l) \geq \alpha\}$ could be higher than its lower α quantile $\inf\{l \,|\, P(L \leq l) > \alpha\}$. We will (arbitrarily) always opt for the upper quantile, which is more conservative as losses are counted positively ($L \geq 0$).

[4] Note that such frequency interpretation, to be valid, implicitely needs the rather unrealistic assumption that no auto-correlation exists in the sequence of losses. The correct mathematical interpretation is the probabilistic one.

$1 - \alpha$. The VaR has remained a widely used measure of risks among portfolio managers; Bassel II also uses the VaR(0.999) for defining its capital requirements. Nevertheless, it has two main drawbacks: it says nothing about the size or average size of the extreme losses of the αth percentile, and it is in general not subadditive, meaning the value at risk of a portfolio can be more than the sum of the values at risk of its individual components.

The Expected Shortfall (ES) at the α confidence level is simply the average loss conditional on the losses being in the worse α quantile[5]:

$$ES(\alpha) = E(L \mid L \geq \mathrm{VaR}(\alpha))$$

Alternatively, it can be proved[6] that:

$$ES(\alpha) = \frac{1}{1-\alpha} \int_{\alpha}^{1} \mathrm{VaR}(u) du$$

As an example of its use, the Expected Shortfall at the 99% confidence level is the chosen Eurosystem tail risk measure. In particular, this is the measure used to calibrate the haircuts on the collateral the Eurosystem receives from the counterparties participating in refinancing operations. Setting the haircut equal to the expected shortfall at the 99% confidence level insures that, on average, the Eurosystem will make a zero loss even if all the losses realisations are drawn from the worse percentile of the tail distribution of losses.

Contrary to the VaR, it can be proved that the Expected Shortfall is subadditive. By definition, it also gives a good indication of the size of the average loss in the worse αth percentile.

Notice that the VaR and the ES only depend on the loss distribution L of the considered asset, hence they are sometimes defined in relation to the loss distribution L only, with no mention of the underlying portfolio.

2.3 Granularity of a portfolio and Herfindahl index

Intuitively, the concept of granularity refers to the extent to which a system can be broken into smaller parts, or that a larger entity can be subdivided. Granularity is certainly relevant in a risk management context. If the system is "granular", the aggregate risk of the portfolio depends less on individual, idiosyncratic shocks, as no single loan is given an extreme weight in the overall portfolio risk. Hence, a granular portfolio risk would be mainly driven by the common factors which affect all the loans simultaneously than by individual loan characteristics. The Herfindahl index, which we define below, is a measure of granularity commonly used by market participants.

Let consider a portfolio of n loans and let w_i be the share of the ith loan in the total loan amount at the current time t. The Herfindahl index (at time t) is, by definition:

$$H\,\mathrm{erf} = \sum_{i=1}^{n} w_i^2$$

Notice low values of the index indicate higher granularity[7].

[5] This is in fact the definition of the Tail Conditional Expectation, to which the Expected Shortfall is equal in case of a continuous distribution. The interested reader can refer to [20] for the general (and cumbersome) definition of Expected Shortfall and its comparison to Tail Conditional Expectation. When no distinction needs to be made for the purpose of the present paper we always present the simplest definition.

[6] This general fact is admitted here. For a proof the reader can refer to Acerbi and Tasche (2002), "On the cooherence of expected shortfall" J Bank Fin 26 (7), page 1492.

[7] The CEPH [?] defines the *inverse* of the above quantity to be the Herfindahl index. Hence higher Herfindahl index values, in the context of the CEPH, would mean higher granularity.

2.4 Definitions of defaults

The notion of default is given very different definitions depending on the overall context, type of asset and jurisdiction. Even within a given framework, the definition can vary. For example, in the Basel II framework defaults are usually defined as the event of the borrower being more than 90 days late in a payment to the creditor, or being unlikely to pay. But in the case of retail (including credit card) and public sector entities exposures, the 90 days figure can be replaced by figures up to 180 days. National discretion lists provided by national supervisors indicate the cases where a bank should apply a longer than 90 days definition for defining the event of default on certain products[8]. Rating agencies have a different approach. Both Standard and Poor's and Moody's characterise default as encompassing any missed or delayed disbursement of interest and/or principal, bankruptcy, and distressed exchange[9]. Hence rating agencies do not allow for a systematic grace period after a missed or delayed payment, contrary to Basel II and most bank internal default accounting. Instead, they would only report the credit event as a default if a monetary loss becomes sufficiently probable.

Overall, for Corporates the definitions in use are usually inspired by the rating agencies and thus more or less consistent, whereas for retail there is a variety of different definitions. A practical consequence is that it is probably better to use delinquency data than default data for quantitative studies. For example, the delinquency three months rate is the number of loans which are more than three months late in repayment of principal or interest, and this gives rise to a more homogeneous statistic across jurisdictions than defaults[10].

3 Description of default models widely used: Vasicek, Gaussian copulas and generalisations...

3.1 A structural, single-firm model: the Merton model

3.1.1 Model description [25]

Simply put, Merton model (1974) assumes that a firm will default if, at the time of servicing its debt, the value of its assets stands below the value of its debt. The Merton model of default is considered to be the first modern model of default. It originally applies to a *single* borrower i. Still, for consistency of notations with the next sections, as well as for the extension to multi-borrowers in this section, we will indicate the index i on each variable involved. Merton model is called a *structural* model, or *asset value* model, because the probability of default PD_i of the borrower i is directly linked to the evolution of the value of its assets $A_{i,t}$ and of its debt $B_{i,t}$. The model was originally formulated for corporate debt. The intuition underpinning Merton's reasonning is the following: at maturity T of the debt, which is modelled for simplicity as a zero-coupon bond of value $B_{i,t}$, the shareholders of the company can decide whether or not to honour their obligations by paying back their debt. If the value of the firm is below the value of the debt $B_{i,T}$ of the company at time T, the shareholders will not repay and let the company default. Bondholders will be entitled to liquidate the company assets and will share the profits from the sale, whereas shareholders will make a zero profit. If the value of the firm is above the debt level $B_{i,T}$ of the company at time T, then the firm will repay. Hence, Merton simple model sees the shareholders as

[8] See the Basel Committee on Banking Supervision FAQ, http://www.bis.org/bcbs/qis/qis3qa_f.htm.

[9] As an example, Moody's definition of default for Corporate includes three types of credit events:

1. A missed or delayed disbursement of interest and/or principal;
2. Bankruptcy, administration, legal receivership, or other legal blocks (possibly by regulators) to the timely payment of interest and/or principal; or
3. A distressed exchange occurs where: (i) the issuer offers debt holders a new security or package of securities that amount to a diminished financial obligation (such as preferred or common stock, or debt with a lower coupon or par amount, lower seniority, or longer maturity); and (ii) the exchange has the effect of allowing the issuer to avoid a bankruptcy or payment default. See for example "Corporate Default and Recovery Rates", 1920-2008, page 51.

[10] Homogeneity across jurisdictions is one of the two reasons the CEPH actually uses delinquency data, and not default data, the other reason being the data quality of the time series itself.

having a European call option on the firm's assets with maturity T (the maturity of the debt) and strike price $B_{i,T}$ (the value of the debt at time T). The shareholder's payoff at time T is thus:

$$\max(0, A_{i,T} - B_{i,T})$$

where $A_{i,t}$ is the value of the firm at time t and $B_{i,t}$ the debt level of the firm at time t. This allows Merton to use the framework developped by Black and Scholes for option pricing [5].

It is important to understand Merton's original idea since it leads to the broader concept of a threshold $B_{i,T}$ under which default occurs. In the more advanced models which are detailed below, this threshold will be time varying, the quantity $A_{i,T}$ which is compared to it will not necessarily be a function of the value of the debtor assets, and defaults will be allowed to occur at any time t instead of just at maturity of the debt T. Nevertheless, the idea of a threshold to which some creditworthiness index of the firm compares still forms the basis of most default models.

The value $A_{i,t}$ of the borrower's assets being not observable[11], a model for its evolution has to be assumed. Merton assumes that $A_{i,t}$ follows a Ito stochastic process of the form:

$$dA_{i,t} = r_i A_{i,t} dt + \sigma_i A_{i,t} dW_{i,t}$$

This allows Merton to directly use Black and Scholes framework for option pricing, in which the probability of default is:

$$PD_i = P(A_{i,T} < B_{i,T}) = P(Y_i < c_i) = \Phi(c_i)$$

where

$$c_i = \frac{\ln(B_{i,T}) - \ln(A_{i,0}) - r_i T + \frac{\sigma_i^2}{2} T}{\sigma_i \sqrt{T}}$$

and Y_i is a standard normal random variable. The derivation is presented in Annex 6.2. Because default is triggered in scenarios where the realisation of the variable Y_i falls below the value c_i, the variable Y_i can be seen as indicating the borrower credit-worthiness: the higher it is, the more credit-worthy the borrower. For this reason we will from now on call Y_i the *credit-worthiness index* of loan i. Notice Merton model implies a standard normal distribution for Y_i, as indicated above.

3.1.2 Limitations, extensions and current use

In this section we explain the limitations of the Merton model [25], and how they were dealt with by the financial industry. The proposed extensions to the Merton model pave the way for the joint simulation of default covered in the next section, such as the Vasicek one-factor model [32].

The main limitations of the simple Merton model are the following:

1) It is a *one-period default model*, or a *static* model: although the dynamics of Y_i follows a continuous time-process on $[0, T]$, default can only occur at time T in the model.

2) The process driving the value $A_{i,t}$ of the firm i is not observable.

3) It is a *single-borrower* default model.

4) The debt structure assumed for the firm is very simple (a zero coupon bond), and refinancing conditions are not taken into account in the model.

5) It does not take into account liquidity effects on the default of firms.

Despite these limitations, two widely used model of credit risk, CreditMetrics™ and KMV™, both rely on Merton model.

CreditMetrics™ was developped in 1997 by the risk management research division of JP Morgan, which eventually became the RiskMetrics™ group. The quality of its publicly available methodology

[11] Even for listed companies, all that can be observed are the stock exchange values. Those are often taken as a proxy for the value of the firms itself, but this is an approximation.

description [11], stemming from the philosophy of its authors[12] to make the product transparent, contributed to its success and influence in many bank-internal developments of credit risk models. The CreditMetrics^{TM} approach is used by many central banks [4], either directly using the CreditManager software, or through in-house systems developed in Matlab or Excel using a similar methodology.

KMV was a small company who specialised in credit and portfolio tools which was acquired by Moody's. Most of large banks and insurance companies use at least one of the major KMV products.

The first limitation was addressed as early as 1976 by Black and Cox [6] extension of the model. It simply consists in assuming that default occurs as soon as the value of the assets falls below the debt threshold. Hence the time of default τ_i of the ith borrower is just $\tau_i := \inf\{t, A_{i,t} < B_{i,t}\}$. For that reason, it is called a *first-passage* model. Defaults can take place at any time, and using properties of the brownian motion such as the reflection principle allows to derive a closed analytical formula for the default probability during any time interval. Nevertheless, this is at the cost of a higher mathematical complexity. Moreover, it introduces the problem of the predictability of default: because the underlying asset value process is a continuous process (with no jumps), and that default occurs only when this process hits the level of the default threshold, default does not come as a surprise, in contrast to what usually happens in real life. Attempts in the financial literature to tackle the predictability of default in first-passage models have consisted either of incorporating jumps in the value process $A_{i,t}$, which can thus unexpectedly fall below the default threshold [17], [34], or by allowing only for imperfect information from the lender side, who will not know the exact position of process $A_{i,t}$ anymore, and thus can be surprised by the borrower default [9], [15], [22].

Concerning the second limitation[13], notice that what is really required in the model is an estimation of the credit-worthiness process Y_i. Although this estimation stems from the value of the firm in Merton's original model, with underlying dynamics given by the stochastic process $dA_{i,t} = r_i A_{i,t} dt + \sigma_i A_{i,t} dW_{i,t}$, other specifications for Y_i can be used. For example, CreditMetrics^{TM} make the variable Y_i linearly dependent on a vector of factors X, as decribed in Annex 6.3. The parameter c_i can then be calibrated based on the empirical probability of default PD_i by noticing that $PD_i = \Phi(c_i)$ implies $c_i = \Phi^{-1}(PD_i)$. Notice that estimating the c_i in that way also allows to address the fourth limitation, as the debt level B_t completely disappears from the model[14]. Nevertheless, this is at the cost of the loss of the dependence of default on the current debt level that was in the original Merton model. Hence, the influence of the borrower debt level on its probability to default will only be taken into account insofar his debt level is already captured through the empirical PD_i (or by the borrower rating used to derive the PD_i). Similarly, the refinancing conditions are only taken into account in this model insofar they are captured by PD_i[15].

Similarly, the third limitation can be addressed within the framework of Merton model by specifying *dependent*, *inter-related* processes Y_i for each borrower i. Different default thresholds c_i can still be set for each borrower, depending on their credit quality rating, but the dependence introduced between the dynamics of the credit-worthiness indexes Y_i will induce correlation in the joint modelling of the borrowers defaults. Hence, Merton model already yields a pool default model when the driving processes

[12] The authors were Gupton, Finger and Bhatia, although the technical note has been improved, as well as the underlying methodology, many times since its first issue in 1997.

[13] Admittedly, when the borrower is a company publicly traded, the Merton model can be calibrated using stock prices by solving a system of two equations: one is the price of the equity as implied by this model, which considers equity as a call option, and results from the Black and Scholes formula, and the other stems from an application of Ito's lemma. In the context of credit claims such information is not available, and hence not developed further in the present paper. The interested reader can refer to Hull [21], page 490.

[14] Admittedly, there had been other attempts to introduce more complex debt structures than a simple zero-coupon bond, such as Geske multi-coupon debt [13] where the decision to default or not is taken at each coupon payment date. The main common problem of these models is the calibration. Even if calibration were correct they would be very firm-specific and unusable in practice from a comprehensive, global risk-management perspective.

[15] Because most of the time these probability of default estimates are backward-looking (eg realised default frequencies from rating agencies tables), a change in the access to liquidity of the borrower is thus *completely neglected* by all the default models based on Merton's approach. This model drawback has thus potentially contributed to the mis-pricing of risk preceding the financial crisis.

of the borrowers' asset values are made co-dependent. It is precisely the specification of different, yet inter-related processes Y_i for each borrower i which opens the door to Vasicek default models for *pools* of loans, which is explained in the next section (Section 3.2).

An illustration of the use of the Merton model by CreditMetrics^{TM} is provided in Annex 6.3. The KMV^{TM} methodology is roughly similar and hence omitted in the present paper. Asset correlations between the different counterparties are solely captured by the stock price correlations of the respective composite indices the borrowers are mapped to. Indeed, each borrower, be it listed on a stock exchange or not, is mapped, depending on both its industry and country, to a different index[16]. Notice also correlation is not estimated directly in the model, but will result from the correlations between the synthetic stock market indices created for each borrower.

3.2 The Vasicek one-factor model for pools

3.2.1 Model description

The Vasicek one-factor model was one of the first models proposed for modelling defaults in *pools* of loans. It was presented by Vasicek in its seminal paper [32] as linked to the Merton [25] structural model. Nevertheless, as we will see, the asset dynamics can be taken out of the model easily, and realised default taken as an input. The only assumptions still shared with the Merton model will be the concept of a threshold under which the loan defaults, on the one hand, and the probability distribution of the variable involved, on the other hand. We choose to present the shorter version of the Vasicek model, the interested reader can refer to Annex 6.4 for the link between the Vasicek and the Merton's structural model.

Vasicek model consists in assuming that each firm i is associated to a variable Y_i indicating its credit-worthiness and that this variable Y_i can be written as:

$$Y_i = \sqrt{\rho} Z + (\sqrt{1-\rho})\epsilon_i$$

where Z and the ϵ_i are mutually independent, identically distributed random variables following a standard normal law. Note that, by independence, the Y_i are themselves Gaussian, of variance 1 and expected value 0.

Default of firm i occurs in the time period $[0, T]$ if the realisation of the variable Y_i is below the Merton's threshold c_i. The formula giving c_i in terms of the parameters of the Merton model is given in Section 3.1.1 and is proved in Annex 6.2. But most of the applications of the Vasicek model do not model explicitly c_i. Instead, they use the following shortcut: Merton's relation states that $\Phi(c_i) = PD_i$. Hence, if PD_i is an (unconditional) probability of default of loan i, then c_i can be recovered by: $c_i = \Phi^{-1}(PD_i)$. This way of only *partially* using some underlying structural model results is typical of the field of default modelling. We will find similar transformations using the inverse quantile function in, for example, the copula approach (Section 3.3).

Hence, the unconditional probability of default of loan i becomes:

$$PD_i = P(Y_i < c_i) = P(Y_i < \Phi^{-1}(PD_i))$$

Borrowers do not default independently from one another, as all the Y_i are impacted by the same common risk factor Z. From independence of the Z and the ϵ_i, it is trivial to see that $Cor(Y_i, Y_j) = \rho$ for all $i \neq j$.

[16] The index can be a synthetic one made of different existing broad market indexes, depending on the borrower industry participation and the jurisdiction of the market it operates in. For example, a borrower might be mapped as 80% French and 20% Germany, and as 70% chemicals and 30% finance, resulting in its reference index being 56% French chemicals, 24% French finance, 14% German chemicals and 6% German finance. See Table 8.8 and 8.9 of [11], p 94-95.

Assuming a portfolio from which all loans are of equal size and whose number of loans tends to infinity, and that all loans have the same probability of default $PD_i = PD$, Vasicek proves that the distribution of the pool default ratio is:

$$P(L \leq x) = \Phi(\frac{\sqrt{1-\rho}\Phi^{-1}(x) - \Phi^{-1}(PD)}{\sqrt{\rho}})$$

The proof is given in Annex 6.5. It relies on the fact that conditional to the realisation of the common factor, the variables Y_i are independent, and thus the Law of Large Numbers can be applied to find the limiting distribution.

Vasicek model belongs to the class of so-called *one-factor conditionally independent default models*, since, conditional to the realisation of the single common systematic risk factor Z, all the Y_i, and hence all the defaults, are independent. Not only is this crucial fact used for the derivation of the above formula giving the distribution of losses of the portfolio, but it allows to easily compute the value at risk and expected shortfall of the portfolio, through a similar limit argument, as the idiosyncratic components factor out in the limit.

3.2.2 Calibration and current use

The Vasicek one-factor model is often used in conjunction with Gordy's framework [16]. This is because under a certain set of technical conditions, but with very broad assumptions in terms of individual probability of default, of loss given default and of loan size distribution, Gordy provides a formula for the portfolio loss distribution, its value at risk as well as its expected shortfall (Annex 6.6). Notice Gordy's framework still requires that the Herfindahl index[17] of the porfolio tend to 0 when the number of loans tends to infinity, a condition already mentioned in Vasicek original work [32]. Hence, Gordy's framework is very similar to Vasicek's framework in the sense that the loss distribution is derived asymptotically, meaning that in practice it only holds for a "very large" number of loans.

The Vasicek model is used to derive the conditional probability of default which is then plugged into Gordy's formula. The derivation[18] is detailed in Annex 6.6. We describe below one of the most important application of this composite approach.

Basel II capital requirements computations The very core of the Internal Rating Based (IRB) approach of Bassel II to obtain the levels of regulatory capital relies on the mixed Vasicek/Gordy framework described above. Basel II requires banks to hold a minimum level of capital, also referred to as regulatory capital. The more risky a bank's assets, the higher the value of the regulatory capital it should hold. Basel II allows banks, subject to a certain number of conditions, to use the IRB approach to derive their level of capital requirements[19]. Basel II[20] divides a bank's assets into five asset classes: corporate, sovereign, bank, retail and equity[21]. It considers each of this asset class as forming a separate portfolio, and uses the Vasicek/Gordy model to derive the amount of regulatory capital to be held to cover for the risk of each of these portfolios. Hence, the total amount of regulatory capital is simply the sum of the amounts to cover each asset class portfolio. The capital required by each of these five asset classes is the VaR at the 99.9% confidence level as expressed in Annex 6.6 and indicated below; LGD_i is the random

[17] As explained in Section 2.3, the Herfindahl index is a measure of the granularity of a portfolio.

[18] Contrary to the other derivations found in Annex, this proof is not self-contained as it starts from one of the main result from Gordy [16] and proceeds to describe what we believe are *the most meaningful steps* of the mixed Gordy/Vasicek approach. The reader interested in the proof of the (here admitted) technical preliminary result can refer to [16] or [20].

[19] An incentive for banks to implement this approach as opposed to the standard Basel approach is that the resulting regulatory capital is generally lower.

[20] See the Basel Committee on Banking Supervision document: "International convergence of capital measurement and capital standards: a revised framework" [3], paragraph 215 to 243.

[21] Within the corporate asset class, five sub-classes of specialised lending are separately identified. Within the retail asset class, three sub-classes are separately identified. Within the corporate and retail asset classes, a distinct treatment for purchased receivables may also apply provided certain conditions are met. For the purpose of ease of explanation, only five classes with no subclasses and no exceptions are assumed here.

variable representing the loss given default of loan i, and w_i the assumed fixed portion of loan i as a percentage of the total nominal amount of the portfolio:

$$VaR(\alpha) = \sum_i w_i E(LGD_i \,|\, Z = F_Z^{-1}(1-\alpha))\Phi(\frac{\Phi^{-1}(PD_i) + \sqrt{\rho_i}\Phi^{-1}(\alpha)}{\sqrt{1-\rho_i}})$$

The final capital requirement slightly differs from this exact formula as

- the unexpected loss, defined as the difference between the VaR and the expected loss, is considered in place of the single VaR, so as to account for the provisioning of the non-performing loans expected to default.

- The expected loss given default factor $E(LGD_i \,|\, Z = F_Z^{-1}(1-\alpha))$ is approximated by an expected loss given default in "severe economic downturn".

- The one-year probability of default PD_i is scaled for loans of longer maturity, a process called *maturity adjustment*.

- The result from the formula is multiplied by a *scaling factor* equal to 1.06 based on a quantitative impact study[22].

These adjustments are described in more detail in [20].

The individual probability of default PD_i, the expected loss given default as well as the exposure at default are all inputs provided by the bank itself or given by the regulatory rules[23]. The internal model is assessed by national regulators. Another important parameter input of the Vasicek model are the correlations ρ_i. The Basel II framework assumes the following, empirically-based estimation of the correlation coefficient:

$$\rho_i = 0.24 - 0.12\frac{1-\exp(-50PD_i)}{1-\exp(-50)}$$

3.2.3 Limitations of the Gordy/Vasicek model

In both applications, asset correlations are introduced to obtain stressed values of the probability of default through the use of a Vasicek model applied to the pool of all assets from the same sector. Hence the correlations ρ_i which appear in the formulas of the model are not correlations between the different assets of the pool considered (the pool of all the bank lendings in the case of Basel II), but a correlation on all existing similar assets. Its use is to obtain stressed probability of defaults, probably to capture the impact of (unknown) common factors on the default probabilities of similar loans.

Also, both approaches assume an infinitely granular portfolio, and hence can only in practice be applied to very granular portfolio.

3.3 Statistical models: the copula approach

All the previous models are labelled as *structural*, as they are assuming some type of underlying process related to the borrower's fundamentals which explains (triggers) the borrower's default. Another approach avoids modelling the reasons of the defaults, but instead focuses solely in reproducing a given default pattern. The models resulting from this approach are often called *statistical* default models, and the most commonly used for pools of loans are *copula-based models*. Copulas are useful as they are mathematical tools which allow to capture virtually any type of dependence between the loans' defaults. That being said, in most applications a specific type of dependence is assumed through the *Gaussian copula*. Another advantage of copula models is that they are *dynamic* models of default in the sense that default can occur at any time. This is particularly relevant for pricing instruments whose values are dependent on a pool of loans such as ABSs, as the timing of defaults becomes very important in that context.

[22] 365 banks participated in the study, which focuses on the impact of Bassel II in terms of capital requirement as compared to Bassel I.

[23] More precisely, in the *advanced IRB approach*, all those paramteters have to be estimated by the bank, whereas in the *foundation IRB approach*, the loss given default is given by the regulatory rules.

3.3.1 Model description

Copulas as multivariate cumulative distribution functions

A copula can be simply defined as the joint distribution of uniform random variables on the interval $[0,1]$:

Definition 1 *A function $C : [0,1]^k \to [0,1]$ is a (k-dimensional) copula if there exists k random variables U_1,...,U_k uniformely distributed on $[0,1]$ such that C is their joint distribution function, idem est such that:*

$$C(u_1, ..., u_k) = P(U_1 \leq u_1, ..., U_k \leq u_k) \quad \textit{for all } (u_1, ..., u_k) \in [0,1]^k$$

The use of copula functions to model joint probability distributions is theoretically justified by Sklar's theorem:

Theorem 1 *Let $F = F_{(Y_1,...,Y_k)}$ be a k-dimensional cdf, and F_i its marginals. Then there exists a k-dimensional copula C such that:*

$$F(y_1, ..., y_k) = C(F_1(y_1), ..., F_k(y_k))$$

Moreover, there is unicity of the copula on the cartesian product of the ranges of the F_i.

In particular, Theorem 1 implies the unicity of the copula associated with F when each of the marginals F_i are continuous. Note the *existence* part of the theorem is easy to prove as setting $u_i := F_i(y_i)$ allows to write:

$$\begin{aligned} F(y_1, ..., y_k) &= P(Y_1 \leq y_1, ..., Y_k \leq y_k) \\ &= P(Y_1 \leq F_1^{-1}(u_1), ..., Y_k \leq F_k^{-1}(u_k)) \\ &= P(F_1(Y_1) \leq u_1, ..., F_k(Y_k) \leq u_k) \end{aligned}$$

Because each $F_i(Y_i)$ is uniformly distributed by Lemma 2, let $U_i := F_1(Y_1)$, and let

$$C(u_1, ..., u_k) := P(U_1 \leq u_1, ..., U_k \leq u_k)$$

This gives the relevant copula.

Hence a copula can model virtually any kind of dependence between random variables.

Strictly speaking, a k-dimensional cdf F is entirely defined by its copula C and its marginal densities F_i. This is because, as seen in the proof above:

$$F(y_1, ..., y_k) = C(F_1(y_1), ..., F_k(y_k)) \quad (1)$$

And by Sklar theorem, the converse is true on the cartesian product of the ranges of the marginals: the cdf F entirely defines the copula as

$$C(u_1, ..., u_k) := P(F_1(Y_1) \leq u_1, ..., F_k(Y_k) \leq u_k) \quad (2)$$

Relations (1) and (2) are very important for understanding the use and simulations of random variables through the copula approach.

From relation (2) define a *Gaussian* copula as a copula satisfying:

$$C(u_1, ..., u_k) := P(F_1(Y_1) \leq u_1, ..., F_k(Y_k) \leq u_k)$$

where the F_i are the marginals associated to a k-variate Gaussian cumulative distribution F. It happens that without loss of generality, we can assume the F_i to be all standard normal Gaussian variables: this actually yields the same Gaussian copula[24].

Similarly, one can define other types of copulas, depending on the family of the multidimensional probability distribution functions being considered.

Quantile-to-quantile correspondence

A copula allows to model the joint probability of uniform random variables U_i. Hence it is natural to look for a model of joint default where the U_i would *somehow* represent the time of default τ_i. Because the U_i are all related to each others through the copula which generates them, this would allow to obtain correlated variables for the time of default of the diverse loans of the pool. Nevertheless the times of default variables τ_i probably do not follow a uniform distribution. Moreover, the cumulative distribution functions of time of default $PD_i(t)$ are inferred from borrower credit score or market data as explained earlier, and a good modelling requires using those curves $PD_i(t)$ – possibly different for each borrower – as an input to the copula model. Ideally, a joint modelling of the probabilities of default of the pool loans should be consistent with the empirically determined marginal distribution functions $PD_i(t)$. That is, the model should not impose additional structure on the marginal times of default than the $PD_i(t)$ curves previously determined through the intensity model, and should yield precisely those marginal probabilities of default. **The main interest of the (general) copula approach is precisely that it allows to specify *any* dependence structure between the time to default while being compatible with *any* given unconditional (ie marginal) individual loan distribution of default.**

Recall that PD_i is the cumulative distribution function of the time of default τ_i of loan i and define $U_i := PD_i(\tau_i)$. By Lemma 2, U_i is uniformly distributed. Hence, the copula approach can be followed, but with $U_i = PD_i(\tau_i) = F_i(Y_i)$ (see proof of Theorem 1). This correspondence is called quantile-to-quantile correspondence, as $y_i = F_i^{-1}(PD_i(e_i))$ is a correspondence which allows to go from the empirical distribution e_i to the theoretical one (given by the copula) y_i. Notice this correspondence can be set between virtually *any* empirical or inferred distribution PD_i to any distribution F_i. Hence the choice of the copula is absolutely unconstrained by the PD_i.

Example of the Gaussian *one-factor* copula model

Li's [24] proposes Gaussian copulas as a mean to systematically model credit risk. The Gaussian one-factor copula model is still widely used by the financial industry.

It can easily be shown that a Gaussian copula $C(u_1, ..., u_k) := P(F_1(Y_1) \leq u_1, ..., F_1(Y_1) \leq u_k)$ with

$$Corr(Y_i, Y_j) =: \rho, \text{ for all } i \neq j$$

can be simulated from the following model:

$$Y_i = \sqrt{\rho} Z + (\sqrt{1-\rho})\epsilon_i$$

where Z and ϵ_i are mutually independent, identically distributed standard normal Gaussian variables, and default has occurred at time t for borrower i whenever t satisfies:

$$Y_i < F_i^{-1}(PD_i(t))$$

The variable Y_i can be interpreted as the credit-quality of loan i, which is influenced both by an idiosyncratic factor ϵ_i, and by a systematic factor Z which affects all the loans of the portfolio. Notice that by independence of Z and ϵ_i, the two-dimensional random vectors (Z, ϵ_i) are Gaussian, and hence the variable Y_i is also Gaussian. Also, the variable Y_i is compared to the *default threshold* $a_i(t) := F_i^{-1}(PD_i(t))$,

[24] This is the consequence of the fact that copulas are invariant by transformation of the marginals by increasing functions. Because the functions $y \rightarrow \frac{y-\mu}{\sigma}$ are increasing and transform a normal law of mean μ and variance σ^2 to a standard normal Gaussian law, the result follows.

in a way similar to the Vasicek one-factor model. This is the reason why the variable Y_i is sometimes called a *credit-worthiness indicator*, as in [29]: when it falls below the threshold, the associated loan defaults.

Since by Lemma 2 the variable $F_i(Y_i)$ is uniformly distributed, it follows that

$$P(Y_i \leq F_i^{-1}(PD_i(t))) = P(F_i(Y_i) \leq PD_i(t)) = PD_i(t)$$

as wanted. Hence only the *joint* probability of default depends on the copula, and the unconditional probability of default matches exactly the empirical input $PD_i(t)$.

The time of default random variable τ_i is thus simply the function which solves:

$$Y_i = F_i^{-1}(PD_i(\tau_i))$$

That is,

$$\tau_i = PD_i^{-1}(F_i(Y_i))$$

Notice τ_i is both a function of the empirical curve $t \to PD_i(t)$ and of the copula since $U_i := F_i(Y_i)$ is the uniform random variable which is the ith marginal of the copula[25].

3.4 Lévy process approach

3.4.1 Model description

Before introducing the Lévy process approach proposed in 2008 by Dobransky and Schoutens [1], notice that the previous Gaussian one-factor framework consists in assuming the following model for the credit-worthiness of the loan i:

$$Y_i = X_\rho + X_{i,1-\rho}$$

where X_ρ and $X_{i,1-\rho}$ are independent random variables, X_ρ follows a Gaussian distribution of mean 0 and variance ρ and the $X_{i,1-\rho}$ are identically distributed and follow a Gaussian distribution of mean 0 and variance $1-\rho$. It follows that

$$Corr(Y_i, Y_j) =: \rho, \text{ for all } i \neq j$$

The one-factor Lévy process framework from Dobransky and Schoutens [1] is simply a generalisation of the above Gaussian framework: instead of using a Gaussian distribution for the processes X_ρ and $X_{i,1-\rho}$, another Lévy process is chosen. For example, if a drifted Gamma process is chosen instead of Gaussian processes, then the model can be written as:

$$Y_i = X_\rho + X_{i,1-\rho}$$

with X_ρ and $X_{i,1-\rho}$ are defined by

$$\begin{aligned} X_\rho &= \sqrt{a}\rho - G_\rho \\ X_{i,1-\rho} &= \sqrt{a}(1-\rho) - G_{i,1-\rho} \end{aligned}$$

where G_ρ and $G_{i,1-\rho}$ are independent random variables following different Gamma distributions: the variable G_ρ follows a Gamma distribution of parameter $\rho a > 0$ and $b = \frac{1}{\sqrt{a}}$, while the variables $G_{i,1-\rho}$ follow a Gamma distribution of parameter $(1-\rho)a > 0$ and $b = \frac{1}{\sqrt{a}}$. Other processes, in particular Poisson processes with unexpected jumps, could be used.

[25] Hence, to obtain a realisation t_i of the time of default of the ith borrower, that is, a realisation of the random variable τ_i, all that is needed is to obtain a realisation u_i from the copula marginal distribution U_i. Then the time of default is derived as $t_i = PD_i^{-1}(u_i)$. The way copulas are simulated depends on their nature. The Gaussian copula is particularly easy to simulate: just simulate Z and ϵ_i as defined above, and get $U_i := F_i(\sqrt{\rho}Z + (\sqrt{1-\rho})\epsilon_i)$.

Recall that the density function of a Gamma distribution $\mathcal{G}(a,b)$ of parameter $a > 0$ and $b > 0$ is defined by:

$$f(x;a,b) = \frac{b^a}{\Gamma(a)} x^{a-1} \exp(-xb)$$

with $\Gamma(a)$ such that

$$\int_{-\infty}^{+\infty} f(x;a,b)dx = 1$$

As in the Gaussian case, the model assumes default at time t if, and only if, the credit-worthiness index falls below some *default threshold* $a_i(t)$. To calibrate this default threshold, the default probabilities under this model are matched against those empirically obtained, that is, the theoretical probability of default $P(Y_i \leq a_i(t))$ of the model is set equal to $PD_i(t)$[26], in the same way that would be done for a Gaussian copula.

$$P(Y_i \leq a_i(t)) = PD_i(t)$$

is successively equivalent to:

$$\begin{aligned} F_{X_1}(a_i(t)) &= PD_i(t) \\ a_i(t) &= F_{X_1}^{-1}(PD_i(t)) \end{aligned}$$

The time of default for loan i is thus obtained, as in the Gaussian approach, by solving

$$Y_i = F_{X_1}^{-1}(PD_i(t_i))$$

That is,

$$t_i = PD_i^{-1}(F_{X_1}(Y_i))$$

4 An alternative: using Hidden markov chain models (HMM) for modelling defaults

4.1 Motivation for the use of HMM in default modelling

In all the models described in the previous section, correlation between defaults is introduced, explicitly or implicitly, by the mean of a "common risk factor", which is assumed to weigh on the creditworthiness of all the borrowers in the pool at the same time. This common factor can be estimated directly from the macroeconomic variables deemed relevant, as in a one-factor model CreditMetrics framework, or be assumed to follow a given theoretical probability distribution, as in the Vasicek one-factor model. The correlation parameter, in any case, is simply the dependence of the creditworthiness of the borrower, and hence of his probability of default, to the common factor[27]. When an attempt is made to specify the common factor, an issue arises. If the common factor is assumed to follow a (theoretical) probability distribution, which probability distribution should be chosen? The Vasicek one-factor model and the Gaussian copula model choose a normal distribution. If on the contrary the common factor is assumed to depend on observable macroeconomic data, the issue of correctly selecting the relevant explanatory

[26] As explained in section 2.1.2, $PD_i(t)$ was obtained from the intensity model calibrated with the empirical, point in time data p_T.

[27] Admittedly, using industry and country specific factor loadings, CreditMetrics allows to go farther than the framework of a single *one*-factor model. Nevertheless, even then, it does not calibrate the model on actual defaults, but on stock market indices which are assumed to represent the (hidden) firm value process as in a Merton's model.

variables arises. When the focus is shifted away from the identification of the common factor to the correlation factor, the problem simply translates into the issue of estimating this correlation factor.

A Hidden Markov chain Model (HMM) consists of an unobservable Markov chain which determines the risk-state of the economy. The risk-state influences the individual and/or collective loan probability of default. Hence, in a HMM, correlation between the loans' defaults is induced only by the common factor, which is the hidden risk state. In this sense, HMM models are similar to many of the models detailed in the first section, be it the Vasicek model (Section 3.2), the Gaussian copula model (Section 3.3), or the generalisation of the copula approach with Lévy processes (Section 3.4). As explained earlier, all these models indeed introduced correlation in the default events through a single unobserved common factor. The rationale for using HMM models in estimating default rates is that it does not need to make any assumptions on the possible set of variables driving the common factor, nor on the common factor distribution. The common factor is determined *endogenously from the default data* by maximum likelihood techniques. In particular, it could be that the credit cycle follows different dynamics than the usual macroeconomic variables that we associate with expansions and recessions like GDP growth[28]. In that case, HMM will allow to get to very different results in terms of probability of default estimations compared to any other model based drectly on macroeconomic variables, as it will allow a default-cycle to be fit to the data, which can then be used for predictions. Admittedly, a k-state HMM only allows the common factor to take k distinct values. But these values do not appear as numerical values in a default triggering inequality as in the case of the previous models reviewed. Instead, each value for a "HMM common factor" determines another "state" of the economy, in which *any* probability distribution of default can be assumed.

In this section we provide an illustration of the use of the HMM approach in a way similar to [14], but on European default data.

4.2 Description of the two-states binomial HMM and implementation

In a two-state HMM, the risk-state of the economy is modelled via two different (unobserved) states: a *low risk state*, labelled 1, and an *high risk state*, labelled 2. The risk state determines the probability distribution of defaults. Loans are assumed to default independently conditional to the risk state. They are also assumed to be so similar than they have the same probability of default. Hence if p_i is the individual loan probability of default in a given state i in $\{1, 2\}$, the law governing the number of defaults $\widetilde{O}_t$ at time t is simply the binomial distribution:

$$P(\widetilde{O}_t = k) = \binom{n_t}{k} p_i^k (1 - p_i)^{n_t - k}$$

where n_t is the number of bonds in the sample at time t, and k any integer in $\{1, ..., n_t\}$. We will call this model a *binomial two-state HMM*. In the high risk state the individual probability of default p_2 is expected to be higher than in the low risk state.

Very often time series of defaults are only available through the time series of the corresponding default *rates*. A way to use the two-states binomial HMM described above in the case of default *rate* series is to assume that the default rate multiplied by some (high enough) constant n and rounded to the closest integer corresponds to the number of actual individual defaults of some (proxy) portfolio of constant size $n_t = n$. Hence it is assumed that there exists some portfolio for which each time k loans default in one period, k new loans of similar risk-profile are introduced before the start of the new period. Admittedly, this approach may fail to provide some useful information about the loans' default behaviour compared

[28] For an example of how the credit cycle of a specific industry can differ substantially from the overall economy business cycle (as defined by the IMF), the reader can refer to the US Energy sector study performed in [14]. Assuming a higher probability of default for the energy companies in the index during recessions would lead to results opposed to reality, as the two cycles are almost completely disjoint.

to a "real default data series" approach. For example, periods of high defaults could have made the least credit-worthy loans default, and in the next period one could expect less defaults, a feature which would be captured by a unwarranted fall in the riskiness of the state if the model is applied to the rate series. Nevertheless, empirically we did not find any significant difference between the two models. The reason might be that the number of defaults being usually low compared to the overall size of the sample, working with a decreasing sample size or on the default rate series is roughly equivalent.

The transition from state i to state j is governed by the *state transition matrix* $A = (a_{ij})$, where a_{ij} is the probability of moving to state j, conditional to being in state i. Given an observation sequence $O_1...O_T$, where for each t in $\{1, ..., T\}$ the number O_t is the realisation of the random variable $\widetilde{O}_t$ defined as the number of defaults at time t, the model is re-estimated repeatedly until the log-likelihood of observing $O_1...O_T$ is maximised.

After estimation, retrieving the implied "most-likely" state sequence can be done using the Viterbi algorithm which finds, through dynamic programming, the single state sequence with the highest probability of occurring among all state-sequences of length T. Alternatively, it can be done by maximising the expected number of correct states, an algorithm which makes use of some previously computed quantities of the estimation, and that we choose to call the *HMM algorithm* for state-retrieval[29]. Both state-retrieval algorithms outputs depend not only of the observation sequence but also on the estimated parameters of the model. The converse does not hold: retrieving the state sequence is only an additional analysis which does not affect in any way the fitted parameters of the model.

There are three main issues in the practical implementation of HMMs: the efficiency of the algorithms used, underflow[30], and the non-global character of the maximum likelihood estimates found which makes them depend on the assumed initial conditions.

Efficient algorithms we borrow from the Speech Recognition literature [30] and we implement in Matlab. These algorithms involve computing the log-likelihood, given the model, using the so-called *forward* and *backward* paths, which contrary to direct computation results in an efficient (polynomial time in input size) algorithm. The methodology used in the paper is explained in detail in Annex 7.1.1.

For the computation of forward and backward paths underflow is dealt with by rescaling the corresponding quantities by an appropriate factor, at each step, after verifying that the re-scaled parameters satisfy the recursive relation of the efficient algorithm. For the Viterbi algorithm, underflow is simply dealt with by taking the logarithms.

The possible existence of different (local) maxima is dealt with by perturbing the initial conditions in an *ad-hoc* manner. This is by no mean a complete answer to the problem, but can certainly allow detecting instable solutions, that is, local maxima which depend very much on the starting point used for the parameter estimations.

4.3 An example with binomially-simulated data

To provide an illustration of the methodology as well as a performance check of the (non-trivial) algorithms involved, we apply it to data simulated by a binomial two-state HMM. The data-generating process has a state-transition probability matrix:

$$A = \begin{bmatrix} 0.9 & 0.1 \\ 0.1 & 0.9 \end{bmatrix}$$

Hence, being in any state at a given period, there is a 90% probability to stay in the same state in the following period and a 10% probability to change state. We also set the individual loan probability of default at $p_1 = 0.004$ in the low risk state and at twice this number in the enhanced state: $p_2 = 0.008$.

[29] The two approaches do not theoretically necessarily yield the same result. In practice we noticed the state sequences retrieved closely resemble each other.

[30] Underflow occurs when the quantities computed become so small that they are (wrongly) assimilated with 0 in the machine memory.

We simulate $T = 80$ observations with a sample of constant size $n = 1000$. The initial state is the low risk state (state 1), that is, we set the initial probability distribution to $\pi = (1, 0)$.

To start our estimation, we make the following (very rough) guess: $A_{guessed} = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}$, $\pi_{guessed} = (0.5, 0.5)$ and $p_{1,guessed} = 0.001$, $p_{2,guessed} = 0.020$. We allow for 100 iterations for maximising the log-likelihood as described in Annex 7.1.1, which results in the estimated parameter indicated, together with our starting guess and the true model parameters, in the table below[31]:

	initial guess	parameter estimates	true data-generating process
π_1	0.5	1	1
π_2	0.5	0	0
p_1	0.0001	0.0034	0.004
p_2	0.0020	0.0079	0.008
a_{11}	0.5	0.8823	0.9
a_{22}	0.5	0.9123	0.9

A state-sequence retrieval algorithm can then be applied to the fitted model to try to recover this true state sequence. The graph on the top left-hand corner of Figure 1 below indicates the default series generated and the underlying true state sequence. The other graphs indicate the retrieved state sequence, through either the Matlab built-in Viterbi algorithm, the author's own implementation of the Viterbi algorithm, or the author's HMM algorithm which maximises the *expected number of correct states*, as an alternative to maximising the likelihood of a given chain sequence among all possible state sequence as in the Viterbi algorithm.

Because we simulate the numerical example on which the calibration is applied, we know the true underlying state sequence. We can compute the Hamming distance between the true sequence and the retrieved one, and translate it into the score of the total number of correctly retrieved states divided by the total number of states (which is 80 in our example). The results for this particular simulation are reported below:

	author's Viterbi	Matlab Viterbi	author's HMM
score	0.8750	0.8750	0.8875

The goodness of fit of the model can be assessed through the computation of the so-called mid-pseudo residuals (see Annex 7.1.5). If the fitted model is correct, these residuals should be standard normal. Figure 2 below indicates, from top to bottom and left to right: the data series of the residuals of our simulation, their fit vis-a-vis of the standard Gaussian law via a histogram and a quantile-to-quantile plot, and their partial autocorrelation function (PACF). Notice in HMM models it is perfectly normal for residuals to be correlated: this does not have negative implications for the validity of the model.

Furthermore, normality tests can be carried out on these pseudo-residuals. For example, Jarque-Bera normality tests do not reject the null of a standard Gaussian for the mid-pseudo residuals at the 10% confidence level.

We will be able to compare this fit to those obtained on real world default data in the next sections.

One can also use parametric bootstrap to obtain the standard deviation as well as the variance-covariance matrix of the estimated parameters. Parametric bootstrap is described in Annex 7.1.3. It generates 100 time-series of default from the fitted model, fits a HMM model to each of the scenario using the same starting guess than for fitting the model for the first time, and 100 iterations for the maximisation of the log-likelihood in each of the fit, and allows to obtain statistics on the fitted parameters. We obtain

[31] To ensure reproducibility of results we used the Matlab command rng(5) to set the seed of the random number generator.

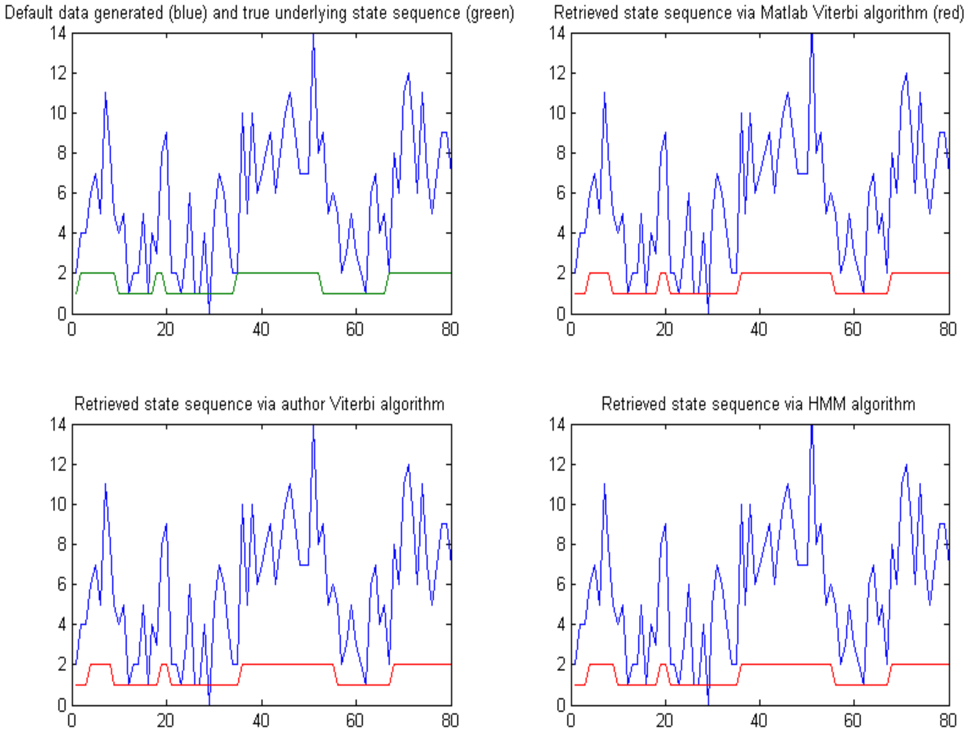


Figure 1: Simulated default data and state sequences (true one and retrieved)

the following correlation matrix:

	p_1	p_2	a_{11}	a_{22}
p_1	1	0.5431	0.3562	-0.1164
p_2		1	0.1571	-0.2162
a_{11}			1	0.4693
a_{22}				1

as well as the mean and standard deviation of each parameter, which are reported in the table below:

	true generating model	mean parameter estimate	standard deviation
p_1	0.0038	0.0039	0.0005
p_2	0.0076	0.0079	0.0009
a_{11}	0.7780	0.8668	0.0876
a_{22}	0.836	0.8535	0.1221

Figure 3 below indicate the histograms of the empirical distribution of each fitted parameter p_1, p_2, a_{11} and a_{22}. As can be seen from the top two histograms, p_1 and p_2 can reasonably be assumed normally distributed. From this we deduce the 95% confidence interval as 1.96 times the standard deviation: $p_1 = 0.0039 \pm 0.00098$ and $p_2 = 0.0079 \pm 0.001764$. As the upper bound for p_1 (0.00488) is lower than the lower bound for p_2 (0.006136), we can safely say that the two distinct states are well identified (at a 95% confidence level).

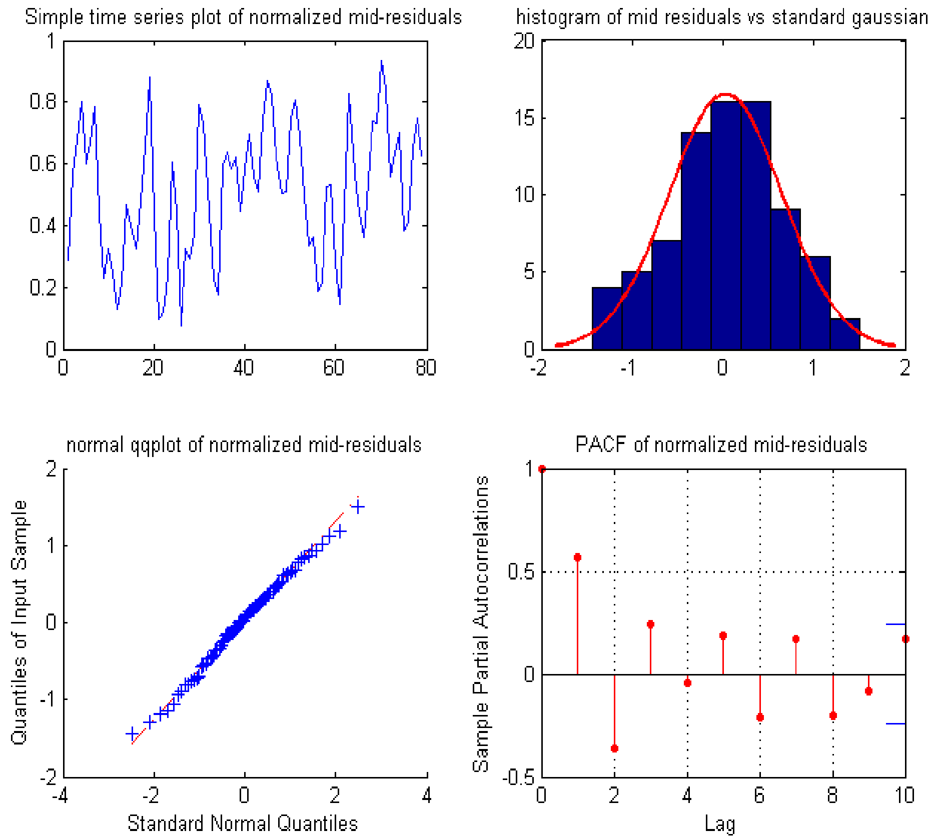


Figure 2: Analysis of the goodness-of-fit of the model using mid-pseudo-residuals

4.4 European speculative grade corporate defaults from Moody's

The data sample consists of 12 month trailing default rates, at a monthly frequency, of European speculative grade corporates bonds, as provided by Moody's on its website[32]. The default rate series was converted into a synthetic bond sample as explained in Section 4.2, using a factor $n = 1000$, meaning there are at each time period 1000 bonds in the sample and that the first period number of defaults is 48 (corresponding to the rate of default of 4.8%). The sample starts on January 1999 and ends on January 2013. The speculative grade corporate bonds follow a credit cycle more closely related to individual corporate loans than investment grade corporate bonds. Hence, they are taken here as a proxy for corporate loans. Nevertheless, drawback of this time series with respect to HMM are twofold:

- it is a trailing rate, meaning the series is bound to exhibit persistence, and the fitted HMM will reflect this persistence. Hence, if the goal is to predict future defaults, using point in time (in this case, monthly) default rate would have been more appropriate.

- It does not distinguish between the different sectors (consumer, etc.) and thus is not granular enough to be useful for estimating precisely credit claims probability of defaults, for example. The aggregation potentially explains that the high risk states are distinguishable by simple graphical inspection of the data time series, without requiring the more sophisticated use of HMM provided here.

Initial values for the estimation are fixed in the most objective way as possible as follows: the initial probability distribution is $\pi = (0.5, 0.5)$. This means, we assume there is no more chance to be in state 1

[32] www.moodys.com

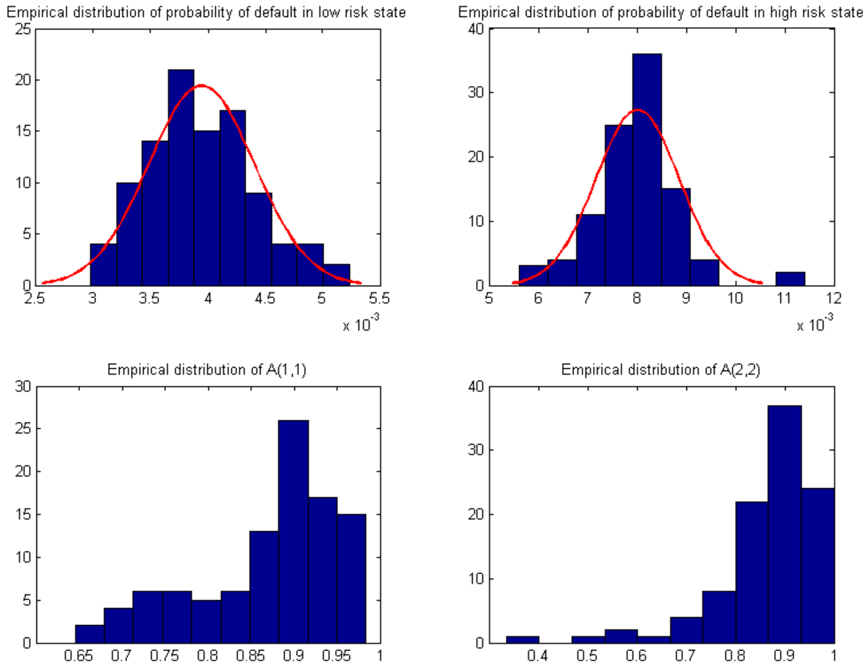


Figure 3: Empirical estimates dispersion

than in state 2. Also, we assume that at each state there is an equal chance to move to the other state, that is, the initial state transition matrix is:

$$A = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}$$

Concerning the individual loan probability of default in the low risk state and the high risk state, they are fixed the following way: according to the initial distribution and transition probability indicated above, there is a 50% chance of being in state 1 and a 50% chance of being in state 2. More precisely by the property of Markov chain such an initial chain would be 50% of the time in state 1 and 50% of the time in state 2. The average probability of default is thus $\frac{1}{2}p_1 + \frac{1}{2}p_2$, which we make equal to the realised frequency of default of the sample. We then just need to specify the ratio $\frac{p_2}{p_1}$ and the values of both p_1 and p_2 will be determined. We decide to choose $\frac{p_2}{p_1} = 2$, which is the single arbitrary choice in our initial guess. The table below indicates the initial guess, and the parameter estimates obtained after 100 iterations to maximise the log-likelihood.

	initial guess	parameter estimates
π_1	0.5	1
π_2	0.5	0
p_1	0.0315	0.0206
p_2	0.0629	0.0922
a_{11}	0.5	0.9810
a_{22}	0.5	0.9523

The results are interesting. The maximum likelihood estimation has separated further two risk-state and has fully characterised them. The initial ratio of 2 between the low and high risk state was not a

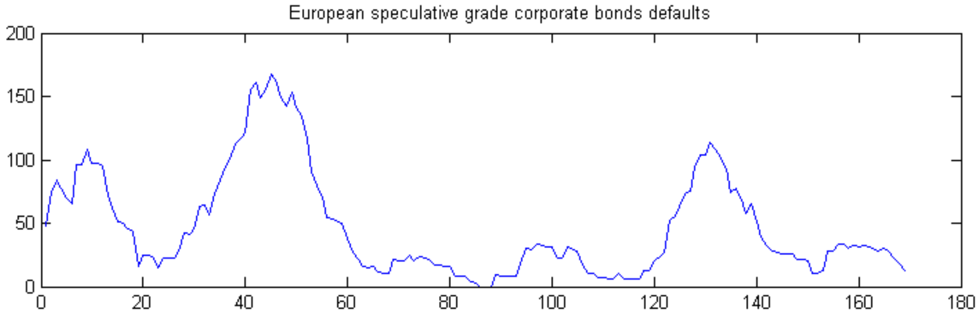


Figure 4: 12 month trailing default series on all European speculative grade bonds rated by Moody's

local maxima and has been increased as high as $\frac{0.0922}{0.0206} = 4.4757$. Hence probability of default appears in this sample very much dependent on the state, which gives credit to muti-state models for defaults. Similarly, the permanence of being in a given risk-state and hence the "inertia" of the model has been increased, from the null inertia we started with (as in any given state we assumed a 50% chance to go to the other state) to as high as 98%, for the low risk state, to 95% for the high risk state. Hence in a low risk state there is only a 2% chance to go to a high risk state, and from a high-risk state there is a 5% chance to go to a low risk state.

Figure 5 below indicates the default series and the underlying retrieved state sequence. Notice the Viterbi algorithm and the HMM algorithm retrieve the same state sequence of low and high default states

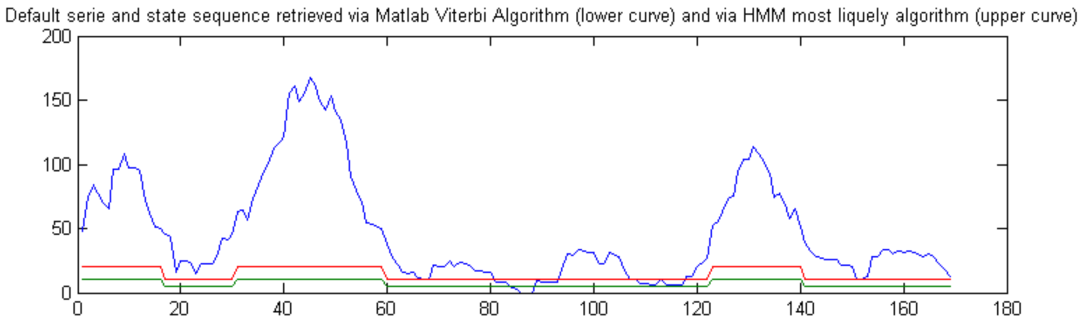


Figure 5: Default series and retrieved state sequence

The goodness of fit of the model can be assessed through the computation of the so-called mid-pseudo residuals (see Annex 7.1.5). If the fitted model is correct, these residuals should be standard normal. Figure 6 below indicates, from top to bottom and left to right: the data series of the residuals of our simulation, their fit vis-a-vis of the standard Gaussian law via a histogram and a quantile-to-quantile plot, as well as their partial autocorrelation function (PACF). As indicated earleir in HMM models it is perfectly normal that residuals be correlated: this does not have negative implications for the model.

Furthermore, normality tests can be carried out on these pseudo-residuals. For example, Jarque-Bera normality tests reject the null of a standard Gaussian for the mid-pseudo residuals at the 10% confidence level, but not at the 5% level.

One can also use parametric bootstrap to obtain the standard deviation as well as the variance-covariance matrix of the estimated parameters. Parametric bootstrap is described in Annex 7.1.3. The

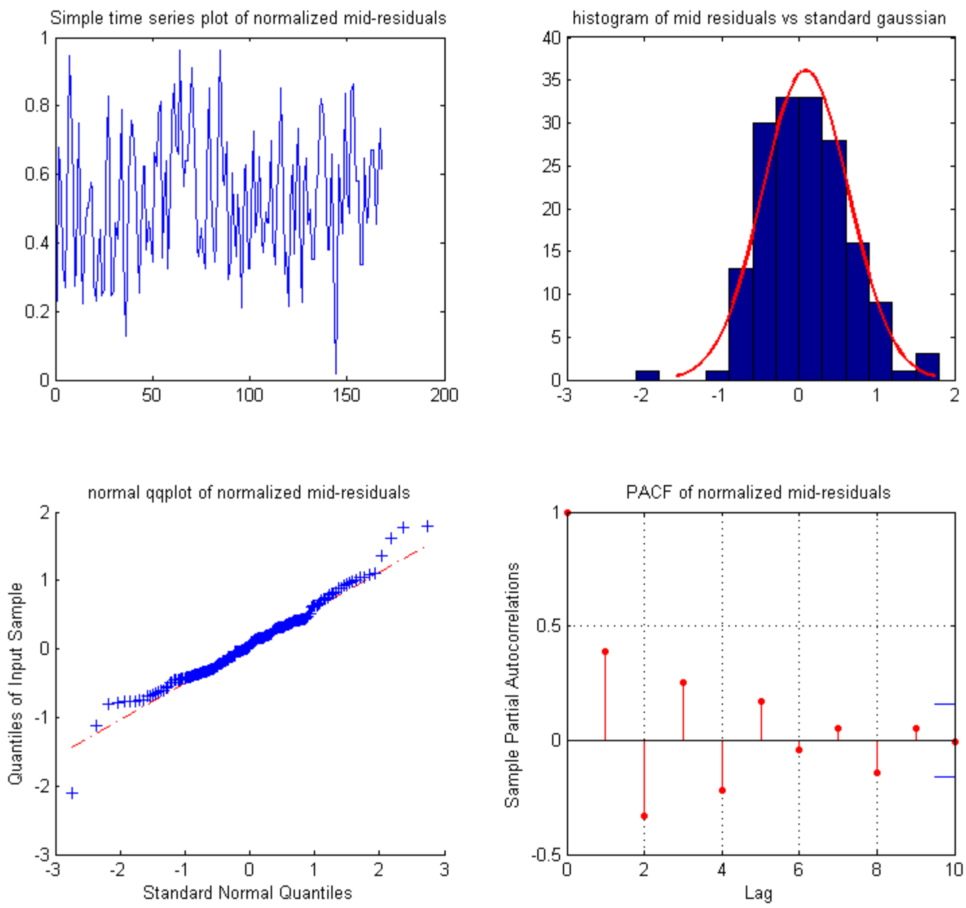


Figure 6: Analysis of the goodness-of-fit of the model for corporates using mid-pseudo-residuals

set up is the same as explained in Section 4.3, and we obtain the following correlation matrix:

	p_1	p_2	a_{11}	a_{22}
p_1	1	0.0091	0.1181	-0.1253
p_2		1	-0.0564	0.1351
a_{11}			1	-0.1531
a_{22}				1

as well as the mean and standard deviation of each parameter. The table below reports the previously indicated fitted model (which generates the scenario for our bootstrap), the means of the different parameters of the fitted models and their standard deviation.

	parameter estimates	mean parameter estimate	standard deviation
p_1	0.0206	0.0206	0.0005
p_2	0.0922	0.0920	0.0015
a_{11}	0.9810	0.9788	0.0147
a_{22}	0.9523	0.9340	0.0445

Figure 7 below indicates the empirical parameter distribution obtained for each of the fitted parameter

p_1, p_2, a_{11} and a_{22}. As can be seen from the top two histograms, p_1 and p_2 cannot reasonably be assumed normally distributed. Nevertheless they seem more concentrated around the mean than a normal law. Hence we can deduce a 95% confidence interval as (less than) 1.96 times the standard deviation: $p_1 = 0.0206 \pm 0.0005$ and $p_2 = 0.0922 \pm 0.0015$. As the upper bound for p_1 (0.0211) is lower than the lower bound for p_2 (0.0907), we would say that the two distinct states are well identified (at a 95% confidence level).

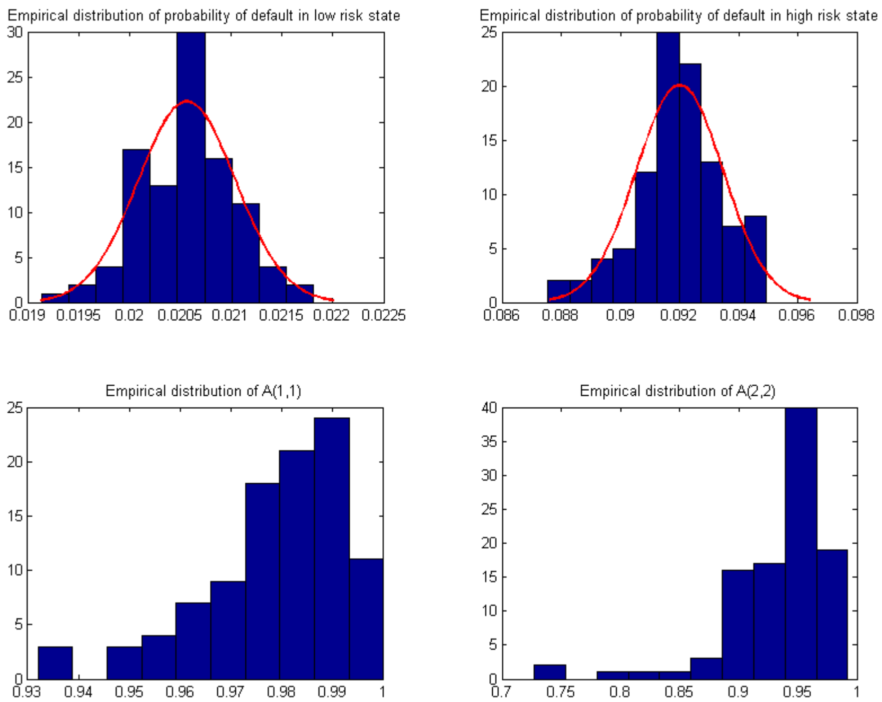


Figure 7: Empirical estimates dispersions

4.5 European pool data from Bloomberg

To be useful for the practical purpose of assessing the probability of defaults of credit claims, there is the need for credit claims default series data. A minimum requirement on the data is that it should be split by type of loans, as for example mortgage loans have a very different risk profile than, say, loans to Small and Medium Enterprises (SMEs). Also, divergence in overall European country risk profiles as well as difference in national legislation concerning loans would indicate that country-specific indexes are needed, at least for a risk analysis purpose.

Default data is in general very difficult to obtain. Defaults are rare events, hence very long default series are required to perform the HMM analysis to uncover risk states and their conditional default probabilities. Nevertheless, if the pool sample is large enough, one can expect to be able to observe enough defaults in shorter time periods. The new European Data Warehouse will be providing loan by loan data, including default, for all Eurosystem eligible ABSs. As pools of credit claims constitute the underlying backing ABSs, this data will provide a valuable source of defaults for the assessment of the risk of credit claims in general.

Currently, this data is not available with a sufficiently long history. We obtained instead the one month delinquency rate for pools of credit claims backing eligible ABSs from Bloomberg. Observations can date back as far as to 2006. This is unfortunately no loan-by-loan level data, but it provides the percentage of default of the pool in percent of the nominal outstanding the pool, at each point in time[33]. We aggregated this data over all the pools of a given country and asset class at each point in time, weightened by the proportion of the pool outstanding amount to the overall outstanding amount of all ABSs, to obtain our generic default rate indexes (using delinquencies as a proxy for defaults).

We chose here to illustrate the HMM methodology on Spanish SMEs.

The data is presented in Figure 8 below:

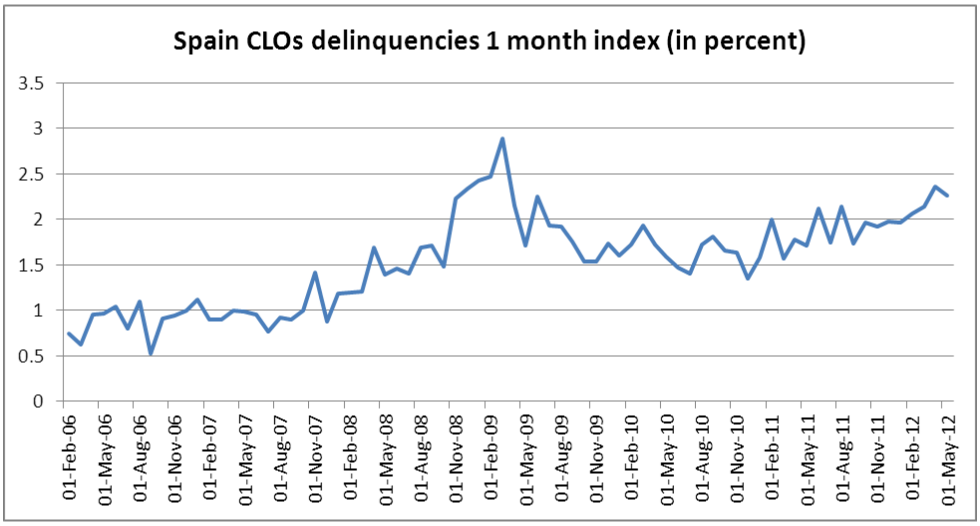


Figure 8: Spain CLOs one month delinquencies

Initial values for the estimation are fixed in the same way as in the previous section (Section 4.4). The parameter estimates are reported in the table below, together with the initial values assumed:

	initial guess	parameter estimates
π_1	0.5	1
π_2	0.5	0
p_1	0.0100	0.0091
p_2	0.0200	0.0179
a_{11}	0.5	0.9613
a_{22}	0.5	1.0000

The results are interesting. The maximum likelihood estimation has separated further two risk-state and fully characterised them. The initial ratio of 2 between the low and high risk state was almost a local maxima and has only been decreased to $\frac{0.0179}{0.0091} = 1.967033$. Probabilities of default are, for this sample, very much dependent on the state. Similarly, the permanence of being in a given risk-state and hence the "inertia" of the model has been increased, from the null inertia we started with (as in any given state we assumed a 50% chance to go to the other state) to as high as 96%, for the low risk state, to 100% for the high risk state. This 100% probability was estimated because the sample only exhibits a single change from state 1 to state 2. Hence it does not provide the HMM with any information about the possibility of leaving the high risk state, and in consequence the HMM indicates a structural change, as it predicts the new high probability of default will persist indefinitely.

[33] Admittedly, this is not quite the same as the percentage of defaults in terms of number of loans, with the exception of the particular case where all loans would be of the same size.

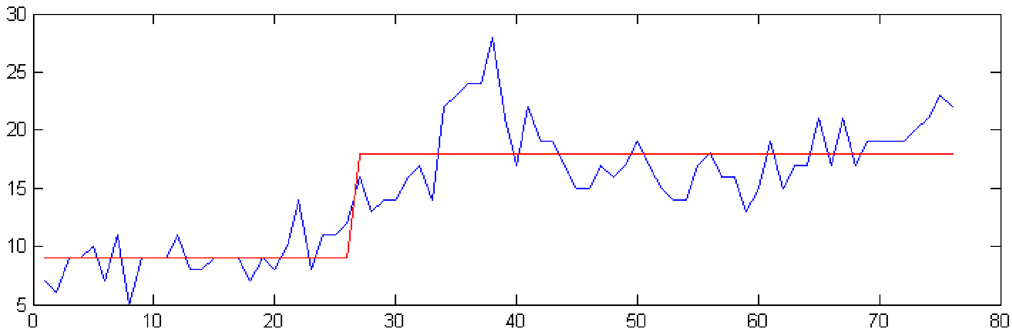


Figure 9: Default series with retrieved state sequence.

Figure 9 below indicates the default series and the underlying retrieved state sequence. Notice the Viterbi algorithm and the HMM algorithm retrieve the same state sequence of low and high default states.

The goodness of fit of the model can be assessed through the computation of the so-called mid-pseudo residuals (see Annex 7.1.5). If the fitted model is correct, these residuals should be standard normal. Figure 10 below indicates, from top to bottom and left to right: the data series of the residuals of our simulation, their fit vis-a-vis of the standard Gaussian law via a histogram and a quantile-to-quantile plot, as well as their partial autocorrelation function (PACF). Notice in HMM models it is perfectly normal that residuals be correlated: this does not have negative implications for the model.

Furthermore, normality tests can be carried out on these pseudo-residuals. For example, Jarque-Bera normality tests reject the null of a standard Gaussian for the mid-pseudo residuals at the 10% confidence level, but not at the 5% level.

One can also use parametric bootstrap to obtain the standard deviation as well as the variance-covariance matrix of the estimated parameters. Parametric bootstrap is described in Annex 7.1.3. The set up is the same as explaned in Section 4.3, and we obtain the following correlation matrix:

	p_1	p_2	a_{11}	a_{22}
p_1	1	0.1073	-0.3755	-0.0587
p_2		1	-0.0915	0.7561
a_{11}			1	-0.0149
a_{22}				1

A very strong positive coefficient (75%) is found between p_2 and a_{22}. This indicates that the results at this level should be taken with caution: the parameters do not seem uniquely determined.

We also obtain the mean and standard deviation of each parameter. The table below reports the previously indicated fitted model (which generates scenario for our bootstrap), the means of the different parameters of the fitted models and their standard deviation.

	parameter estimates	mean parameter estimate	standard deviation
p_1	0.0091	0.0094	0.0017
p_2	0.0179	0.0176	0.0015
a_{11}	0.9613	0.8770	0.1831
a_{22}	1.0000	0.9523	0.0148

Figure 11 below indicates the empirical parameter distribution obtained for each of the fitted parameter p_1, p_2, a_{11} and a_{22}. As can be seen from the top two histograms, p_1 and p_2 cannot reasonably

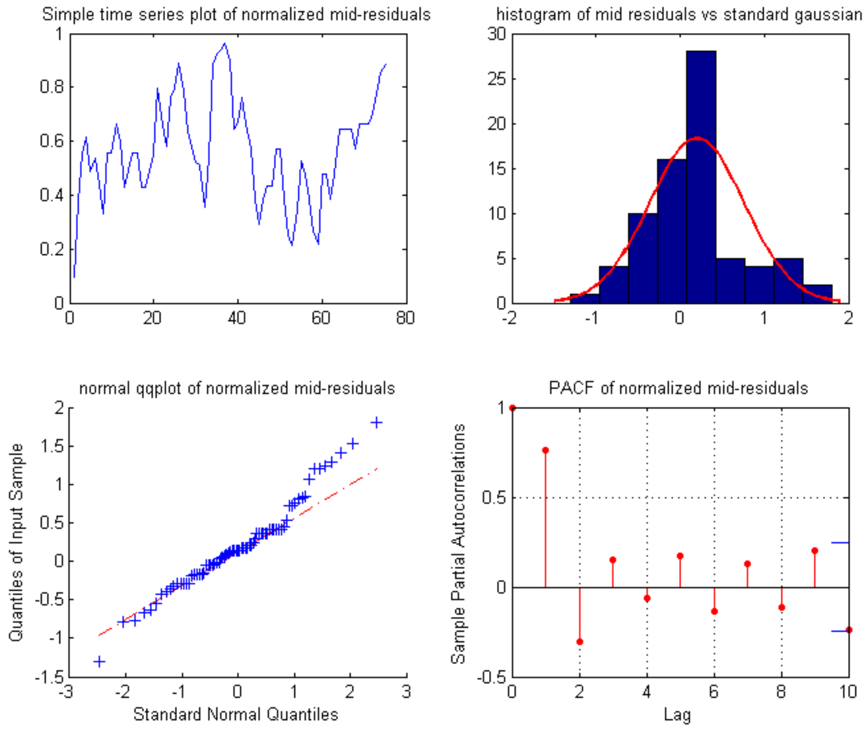


Figure 10: Analysis of the goodness-of-fit of the model for Spanish CLOs using mid-pseudo-residuals

be assumed normally distributed. They seem much more concentrated around the mean than a normal law. Hence we can deduce a (conservative) 95% confidence interval as (less than) 1.96 times the standard deviation: $p_1 = 0.0091 \pm 0.0017$ and $p_2 = 0.0179 \pm 0.0015$. As the upper bound for p_1 (0.0108) is lower than the lower bound for p_2 (0.0164), we conclude that two risk states are indeed disentangled.

5 Conclusion

In this paper we reviewed the most commonly used models of defaults, and explained them in a unified framework with consistent concepts and notations. These models often rely on some creditworthiness indicators which have to exceed some threshold for the loan to default, and on correlation coefficients which reflect common shocks to all the creditworthiness indicators of the loans within the pool. We then proposed an alternative approach for modelling loans default through Hidden Markov Chains. Correlation in that approach is induced by the common state of the credit cycle, which impacts all loans, thus sharing similar features with the previous approaches. A main difference of the Hidden Markov Chain model is that the economy state sequence is not assumed a priori to follow a given theoretical distribution, or to follow a macroeconomic proxy, but estimated directly from the time-series of default pattern. Empirical evidence in the European corporate bond market and the Spanish CLOs market suggested that default series clearly exhibit multi-state properties, which support the use of multi-state Markov model to get more precise estimate of the probability of defaults in each state.

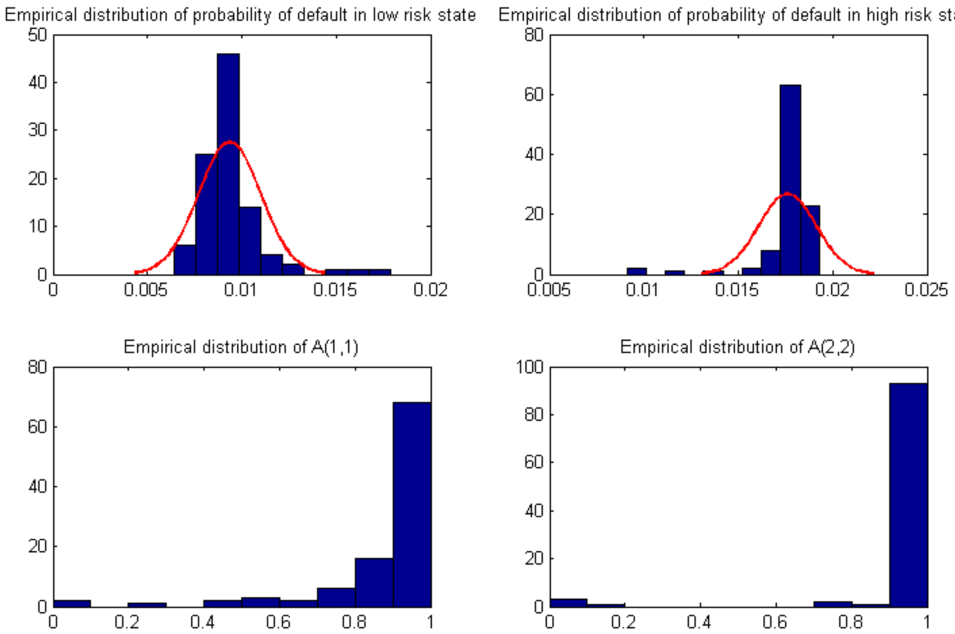


Figure 11: Empirical estimates dispersions

References

[1] Albrecher, H., Ladoucette, S. A., Schoutens W., 2007, "A generic one-factor Lévy model for synthetic CDOs", Advances in Mathematical Finance, R.J. Elliott et als edt., Birkhaeuser

[2] Banachewicz, K., van der Vaart, A., Lucas, A., "Modeling portfolio defaults using hidden Markov models with covariates", Tinbergen Institute Discussion Paper, TI 2006-094/2

[3] Basel Committee on Banking Supervision, 2006, "International convergence of capital measurement and capital standards: a revised framework"

[4] Bindseil et al (task force of the Market Operations Committee), 2007, "The use of portfolio credit risk models in central banks", Occasional Paper Series 64

[5] Black, F., Scholes, M., 1973. "The pricing of options and corporate liabilities", Journal of Political Economy vol 81, issue 3 , 637-654

[6] Black, F., Cox, J., C., 1976. "Valuing corporate securities: some effects of bond indenture provisions ", Journal of Finance 31, 351-367

[7] Bluhm C., Overbeck L., Wagner C., 2002, "An introduction to credit risk modeling", Chapman & Hall/CRC Financial mathematic series

[8] Ching, W-K, Siu, T-K, Li, L., Li, T., Li W-K, 2007, "Modeling default data via an interactive hidden Markov model"

[9] Duffie, D. ; Lando, D., 2001, “Term structure of credit spreads with incomplete accounting information,” Econometrica 69, 633-664

[10] Cariboni, J., Schoutens, W., 2009, "Lévy processes in credit risk", Wiley Finance

[11] CreditMetrics^{TM} –Technical Document

[12] Dobransky P., Schoutens W., 2008, "Generic Lévy one-factor models for the joint modelling of prepayment and default: modelling LCDX"

[13] Geske, R., 1977, "The valuation of corporate liabilities as compound options", Journal of Financial and Quantitative Analysis 7, 63-81

[14] Giampieri, G., Davis, M., Crowder, M., 2005, "Analysis of default data using hidden Markov Models", Quantitative Finance, vol 5, issue 1

[15] Giesecke, K., 2005, "Default and information", Working Paper, Cornell University.

[16] Gordy, M. B., 2002, "A risk-factor model fundation for the ratings-based bank capital rules"

[17] Hilberink, B.; Rogers, L. C. G., 2002, "Optimal capital structure and endogenous default," Finance and Stochastics 6, 237-263.

[18] Ericsson, J., Reneby, J., 2005, "Estimating structural bond pricing models", Journal of Business 78,707-736

[19] Heidorn, T., Kahlert, D., 2010, "Implied correlations of iTraxx tranches during the financial crisis," Frankfurt School Working Paper 145

[20] Hibbeln M., 2010, "Risk management in credit portfolios", Physica-Verlag

[21] Hull, J. C., 2006, "Options, futures and other derivatives", Sixth edition, Pearson

[22] Jarrow, R. A.; Protter, P., 2004, "Structural versus reduced form models: a new information based perspective," Journal of Investment Management 2, 1-10

[23] Kyprianou, E. A., 2006, "Introductory lectures on fluctuations of Lévy processes with applications", Springer

[24] Li, D. X.,1999, "On default correlation: a copula function approach," Working Paper 99-07, The Risk Metrics Group

[25] Merton, R., 1974, "On the Pricing of Corporate Debt: the Risk Structure of Interest Rates," Journal of Finance 29, 449-470

[26] Moody's, "Methodology for forecasting and stress-testing ABS and RMBS deals", Moody 'sAnalytics, 5 August 2010

[27] Moody's, "Moody's approach to real estate analysis for CMBS transactions in EMEA: portfolio analysis (MoRE Portfolio)"

[28] Moody's, "Update on Moody's approach to real estate analysis for CMBS transactions in EMEA"

[29] Nedeljkovic, J., Rosen, D., D Saunders, D. "Pricing and hedging collateralized loan obligations with implied factor models", The Journal of Credit Risk (53–97). vol 6, n. 3

[30] Rabiner L R, "A tutorial on hidden Markov models and selected applications in speech recognition", Proc. IEEE 77 257-286

[31] Schönbucher, P. J., and Schubert, D., 2001, "Copula-Dependent Default Risk in Intensity Models," Working Paper, Department of Statistics, Bonn University

[32] Vasicek, O. A., 2002, "The distribution of loan portfolio values", Risk

[33] Vasicek, O. A., 1997, "The loan loss distribution", Technical Report, KMV Corporation

[34] Zhou, C., 1997, "A jump-diffusion approach to modelling credit risk and valuing defaultable securities", Federal Reserve Board, Washington

6 Annex 1: formal proofs, derivations

6.1 Proof of the intensity model formula when $PD_i(t)$ is derivable

By definition, the instantaneous hazard rate $\lambda_{i,t}$ is the average probability of default of the loan i in over a very short time interval $[t, t+dt]$, knowing the loan has not defaulted before time t, hence we have:

$$\begin{aligned}
\lambda_{i,t} &= \lim_{dt\to 0}\left(\frac{P(t<\tau_i\le t+dt\,|\,t>\tau_i)}{dt}\right) = \lim_{dt\to 0}\left(\frac{P(t<\tau_i\le t+dt)}{P(t>\tau_i).dt}\right)\\
&= \lim_{dt\to 0}\left(\frac{PD_i(t+dt)-PD_i(t)}{(1-PD_i(t)).dt}\right) = \frac{1}{1-PD_i(t)}\lim_{dt\to 0}\left(\frac{PD_i(t+dt)-PD_i(t)}{dt}\right)\\
&= \frac{PD_i'(t)}{1-PD_i(t)} = -\left(\frac{-PD_i'(t)}{1-PD_i(t)}\right)
\end{aligned}$$

Hence $\lambda_{i,t}$ is integrable and integrating both sides yields:

$$\int_0^t \lambda_{i,s}ds = -\ln(1-PD_i(t)) + \ln(1-PD_i(0))$$

From which we deduce the expression of $PD_i(t)$ as a function of the trajectory $(\lambda_{i,t})_t$:

$$PD_i(t) := P(\tau_i \le t) = 1 - \exp\left(-\int_0^t \lambda_{i,s}ds\right)$$

6.2 Merton threshold derivation

Merton assumes the value process $A_{i,t}$ of firm i is driven by:

$$dA_{i,t} = r_i A_{i,t}dt + \sigma_i A_{i,t}dW_{i,t}$$

Using Ito's lemma to find the dynamics of $d\ln(A_{i,T})$ and integrating one can derive the process driving $\ln(A_{i,t})$:

$$\ln(A_{i,T}) = \ln(A_{i,0}) + rT - \frac{\sigma_i^2}{2}T + \sigma_i\sqrt{T}Y_i$$

where Y_i is a standard normal gaussian variable.

Now, Merton assumes default if, and only if, the value of the firm falls below its debt level, idem est $A_{i,T} < B_{i,T}$. This is successively equivalent to

$$\begin{aligned}
\ln(A_{i,T}) &< \ln(B_{i,T})\\
\ln(A_{i,0}) + r_iT - \frac{\sigma_i^2}{2}T + \sigma_i\sqrt{T}Y_i &< \ln(B_{i,T})\\
Y_i &< \frac{\ln(B_{i,T}) - \ln(A_{i,0}) - r_iT + \frac{\sigma_i^2}{2}T}{\sigma_i\sqrt{T}}
\end{aligned}$$

Hence we can define:

$$c_i = \frac{\ln(B_{i,T}) - \ln(A_{i,0}) - r_iT + \frac{\sigma_i^2}{2}T}{\sigma_i\sqrt{T}}$$

6.3 CreditMetrics as a Merton model of default

CreditMetrics is based on the Merton model, although it does not model c_i explicitly. Indeed, the parameter c_i is not derived from Merton model, but calibrated using the corresponding empirical probability of default over the time horizon T, that is, $PD_i(T)$. Indeed from $\Phi(c_i) = PD_i(T)$ it follows that $c_i = \Phi^{-1}(PD_i(T))$.

Because by definition (see Annex 6.2)

$$\ln(A_{i,0}) + r_i T - \frac{\sigma_i^2}{2} T + \sigma_i \sqrt{T} Y_i = \ln(A_{i,T})$$

we have:

$$Y_i = \frac{\ln(A_{i,T}/A_{i,0}) - r_i T + \frac{\sigma_i^2}{2} T}{\sigma_i \sqrt{T}}$$

Because under the Merton model Y_i is assumed to follow a standard normal distribution, the asset-value $\ln(A_{i,T}/A_{i,0})$, under the Merton model, has to follow a normal distribution of mean $r_i T - \frac{\sigma_i^2}{2} T$ and standard deviation $\sigma_i \sqrt{T}$.

CreditMetrics maps each borrower i to a country and a sector stock index. Then it uses this stock index as a proxy for the (unobservable) asset value process $A_{i,t}$. The obtained variable $\ln(A_{i,t}/A_{i,0})$ is then standardised, which as explained above results, under the assumptions of the Merton model, in a standard normal variable Y_i. This is the reason why in the CreditMetrics context Y_i is sometimes referred to as the *standardised log-return level*.

Then CreditMetrics decomposes Y_i as $Y_i = R_i \Psi_i + \epsilon_i$ where the ϵ_i are independent, identically distributed standard Gaussians and R_i is a Gaussian vector representing a systematic risk-factor independent from the ϵ_i, and Ψ_i is the corresponding vector of factor loadings, that is, the weight given to each risk factor in the model.

In Merton model default occurs if, and only if, $Y_i < c_i$, that is, if and only if $\epsilon_i < c_i - R_i \Psi_i$. Hence CreditMetrics obtains the probability of default conditional to the realisation of X as $P(\epsilon_i < c_i - R_i \Psi_i) = \Phi(c_i - R_i \Psi_i)$.

6.4 From the Merton one firm model to the Vasicek model multi-firm model

Because Merton model is a model for a single firm, it makes no assumption on the cross-correlation of the Y_i derived above (Annex 6.2). Vasicek model consists simply in assuming that the Y_i can be written as

$$Y_i = \sqrt{\rho} Z + (\sqrt{1-\rho}) \epsilon_i$$

where Z and the ϵ_i are mutually independent, identically distributed random variables following a standard normal law. Note that, by independence, the Y_i are themselves Gaussian, of variance 1 and expected value 0.

6.5 Vasicek derivation of pool default rate distribution for large homogenous portfolios

Let $p_i(Z)$ be the probability of default of loan i conditional to a realisation of the common factor Z. Then

$$\begin{aligned} p_i(Z) &= P(Y_i < c_i \,|\, Z) = P(\sqrt{\rho} Z + (\sqrt{1-\rho}) \epsilon_i < c_i \,|\, Z) \\ &= P(\epsilon_i < \frac{c_i - \sqrt{\rho} Z}{\sqrt{1-\rho}}) \\ &= \Phi(\frac{c_i - \sqrt{\rho} Z}{\sqrt{1-\rho}}) \end{aligned}$$

Remark: Very often, this result is used to generalize the Vasicek one-factor model by letting Z follow any distribution (possibly not the standard normal Gaussian distribution).

The aim of the Vasicek model is to derive the default distribution of a given pool or porfolio of loans. To reach this aim it proceeds to make a certain number of assumptions. First, all the loans have equal size, and same probability of default $PD_i =: PD$. This implies $c_i = \Phi^{-1}(PD_i) = \Phi^{-1}(PD) =: c$. Second, the limiting distribution of losses is computed when n tends to infinity, meaning in practice that a high number of loans is required to apply the Vasicek model.

Let L_i be the default indicator of the ith loan, that is, L_i is the random variable equal to 1 if loan i defaults and 0 otherwise. Let $L = \frac{1}{n}\sum_i L_i$ be the pool default ratio. The core of Vasicek proof consists in noticing that since $P(L_i = 1 \,|\, Z) = P(Y_i < \Phi^{-1}(PD) \,|\, Z) = \Phi(\frac{\Phi^{-1}(PD)-\sqrt{\rho}Z}{\sqrt{1-\rho}})$, in the model the variable L_i are independently distributed when conditioned to the value of the common risk-factor Z. Hence, by the Law of Large Numbers, the arithmetic average $L = \frac{1}{n}\sum_i L_i$ converges to the distribution of L_i when n tends to infinity. Hence

$$P(L \le x) = P(p_i(Z) \le x) = P(Z \ge p_i^{-1}(x)) = \Phi(-p_i^{-1}(x))$$

since $p_i(x) = \Phi(\frac{c_i-\sqrt{\rho}x}{\sqrt{1-\rho}})$ is a decreasing function of x and Z is assumed to be a standard Gaussian law. Since

$$\begin{aligned} x &= p_i(y) \Leftrightarrow x = \Phi(\frac{c_i - \sqrt{\rho}y}{\sqrt{1-\rho}}) \\ &\Leftrightarrow \frac{c - \sqrt{1-\rho}\Phi^{-1}(x)}{\sqrt{\rho}} = y = p_i^{-1}(x) \end{aligned}$$

we have

$$\begin{aligned} P(L &\le x) = \Phi(\frac{\sqrt{1-\rho}\Phi^{-1}(x) - c}{\sqrt{\rho}}) \\ &= \Phi(\frac{\sqrt{1-\rho}\Phi^{-1}(x) - \Phi^{-1}(PD)}{\sqrt{\rho}}) \end{aligned}$$

6.6 Derivation of Value at Risk and Expected shortfall in the Gordy/Vasicek framework

Two main results from Gordy's framework [16] are the derivation of the VaR(α) and of the $ES(\alpha)$ for the limiting loss distribution of a portfolio whose Herfindahl index tends to 0, assuming a single factor is able to capture all the co-dependence between loans in the portfolio. In the examples of the use of that framework chosen in this paper (Basel II capital requirements), losses *after recovery* are taken into account. Hence this section makes an exception compared to the rest of the paper and explicitly considers recoveries values. Let LGD_i be the loss given default of loan i. Hence a total monetary loss following default corresponds to the special case where $LGD_i = 100\%$. Let D_i be the event of default of loan i. Hence, if the time-frame of our one-period model is $[0, T]$, the event D_i is linked to the previously defined random variables τ_i and L_i by:

$$D_i = (\tau_i \le T) = (L_i = 1)$$

The VaR of the limiting loss distribution is then:

$$\text{VaR}(\alpha) = E(\sum_i w_i LGD_i 1_{D_i} \,|\, Z = \Phi^{-1}(1-\alpha))$$

This result is admitted here[34]. Crucial to Gordy's derivation is that the default events once conditioned on the realisation of the common factor Z are independent. Assuming the exposures of each individual

[34] The above formula also holds for Z following a different distribution than a standard gaussian one: just replace Φ^{-1} by F_Z^{-1}. A proof, in the general setting, can be found in [20], Annex 2.8.9, page 53, or in Gordy original article [16].

loan w_i are fixed and that each LGD_i is, conditionally to Z, independent from the default event D_i, we can write:

$$\mathrm{VaR}(\alpha) = \sum_i w_i E(LGD_i \,|\, Z = \Phi^{-1}(1-\alpha)) E(1_{D_i} \,|\, Z = \Phi^{-1}(1-\alpha))$$

Now the expression $E(1_{D_i} \,|\, Z = \Phi^{-1}(1-\alpha)) = E(1_{D_i} \,|\, Z = -\Phi^{-1}(\alpha)) = P(L_i = 1 \,|\, Z = -\Phi^{-1}(\alpha))$ is the probability of default of loan i subject to a given value $(-\Phi^{-1}(\alpha))$ of the risk factor Z. It is thus a known expression in the Vasicek framework (see Annex 6.5):

$$E(1_{D_i} \,|\, Z = -\Phi^{-1}(\alpha)) = \Phi(\frac{c_i - \sqrt{\rho_i}x}{\sqrt{1-\rho_i}})$$

with $c_i = \Phi^{-1}(PD_i)$ and $x = -\Phi^{-1}(\alpha)$ and where ρ_i *corresponds to the correlation within the assets of a (imaginary) pool made of loans similar to loan i*. This probability of default conditional to a systemic factor event is then plugged into Gordy's formula:

$$\mathrm{VaR}(\alpha) = \sum_i w_i E(LGD_i \,|\, Z = \Phi^{-1}(1-\alpha)) \Phi(\frac{\Phi^{-1}(PD_i) + \sqrt{\rho_i}\Phi^{-1}(\alpha)}{\sqrt{1-\rho_i}})$$

The ES of the limiting loss distribution can be expressed as[35]

$$ES(\alpha) = \frac{1}{1-\alpha}\int_\alpha^1 \mathrm{VaR}(u)du$$

Using the previous formula for expressing $\mathrm{VaR}(u)$, and assuming a fixed value for the stressed (conditional) expected loss given defaut $E(LGD_i \,|\, Z = \Phi^{-1}(1-\alpha)) =: ELGD_i^{stressed}$ yields:

$$\begin{aligned} ES(\alpha) &= \frac{1}{1-\alpha}\int_\alpha^1 \sum_i w_i ELGD_i^{stressed} \Phi(\frac{\Phi^{-1}(PD_i) + \sqrt{\rho_i}\Phi^{-1}(u)}{\sqrt{1-\rho_i}})du \\ &= \frac{1}{1-\alpha}\sum_i w_i ELGD_i^{stressed} \int_\alpha^1 \Phi(\frac{\Phi^{-1}(PD_i) + \sqrt{\rho_i}\Phi^{-1}(u)}{\sqrt{1-\rho_i}})du \end{aligned}$$

But the change of variable $u = \Phi(-x)$ gives, since $\phi(-x) = \phi(x)$:

$$\begin{aligned} \int_\alpha^1 \Phi(\frac{\Phi^{-1}(PD_i) + \sqrt{\rho}\Phi^{-1}(u)}{\sqrt{1-\rho}})du &= -\int_{-\Phi^{-1}(\alpha)}^{-\Phi^{-1}(1)} \Phi(\frac{\Phi^{-1}(PD_i) - \sqrt{\rho_i}x}{\sqrt{1-\rho_i}})\phi(-x)dx \\ &= -\int_{-\Phi^{-1}(\alpha)}^{-\infty} \Phi(\frac{\Phi^{-1}(PD_i) - \sqrt{\rho_i}x}{\sqrt{1-\rho_i}})\phi(x)dx \\ &= \int_{-\infty}^{-\Phi^{-1}(\alpha)} \Phi(\frac{\Phi^{-1}(PD_i) - \sqrt{\rho_i}x}{\sqrt{1-\rho_i}})\phi(x)dx \\ &= \Phi_2(-\Phi^{-1}(\alpha), \Phi^{-1}(PD_i), \sqrt{\rho_i}) \end{aligned}$$

where we used that

$$\Phi_2(x, z, a) = \int_{-\infty}^z \Phi(\frac{x - ay}{\sqrt{1-a^2}})\phi(y)dy$$

Hence

$$ES(\alpha) = \frac{1}{1-\alpha}\sum_i w_i ELGD_i^{stressed} \Phi_2(-\Phi^{-1}(\alpha), \Phi^{-1}(PD_i), \sqrt{\rho_i})$$

[35] This general fact is admitted here. For a proof the reader can refer to Acerbi and Tasche (2002), "On the cooherence of expected shortfall" J Bank Fin 26 (7), page 1492.

6.7 Simulation of the reduced Gaussian copula model and link to the Vasicek model

Here we study a particular case of Gaussian copula, show how simulations of this reduced Gaussian copula model can easily be obtained, and how it is consistent with the structural one-factor Vasicek model. We also explain in detail how to simulate time to default within this framework.

Let $C(u_1, ..., u_k) := P(F_1(Y_1) \leq u_1, ..., F_1(Y_1) \leq u_k)$ be a Gaussian copula; it is associated to a Gaussian cdf $F_{(Y_1,...Y_k)} = F$. Without loss of generality, we can assume the Y_i are standard normal variables. We seek to impose the following simple linear correlation structure: the images of marginal $F_i^{-1}(U_i) = Y_i$ of the copula C have to satisfy:

$$Corr(Y_i, Y_j) =: \rho \text{ for all } i \neq j$$

The coefficient ρ is often called "asset correlation", because asset price correlations are often used as a proxy for ρ. It is thus a parameter of the copula, which has been reduced in its generality by assuming only a *single* cross-asset correlation, instead of the general $n(n-1)/2$ possibly distinct cross-correlations. This simplifying assumption is often made in default modelling.

A way to obtain the condition $Corr(Y_i, Y_j) =: \rho$ for all $i \neq j$, is to simulate $k+1$ independent and identically distributed random variables $U'_1, ..., U'_{k+1}$ and then compute $Y_i = \sqrt{1-\rho}\Phi^{-1}(U'_i) + \sqrt{\rho}\Phi^{-1}(U'_{k+1})$, where Φ is the cdf of a standard normal Gaussian.

Indeed, by bilinearity of the covariance and by independence we have:

$$Cov(Y_i, Y_j) = \rho Var(\Phi^{-1}(U'_{k+1})) = \rho\sqrt{Var(\Phi^{-1}(U'_i))}\sqrt{Var(\Phi^{-1}(U'_j))}$$

and, since by independence of the $\Phi^{-1}(U'_i)$ with the $\Phi^{-1}(U'_{k+1})$:

$$\begin{aligned} Var(Y_i) &= Var(\sqrt{1-\rho}\Phi^{-1}(U'_i) + \sqrt{\rho}\Phi^{-1}(U'_{k+1})) \\ &= (1-\rho)Var(\Phi^{-1}(U'_i)) + \rho Var(\Phi^{-1}(U'_{k+1})) \\ &= Var(\Phi^{-1}(U'_i)) \end{aligned}$$

we do get

$$Corr(Y_i, Y_j) = \frac{Cov(Y_i, Y_j)}{Var(Y_i)} = \frac{Cov(Y_i, Y_j)}{\sqrt{Var(Y_i)}\sqrt{Var(Y_j)}} = \rho$$

This derivation explains the choice of the form for the coefficients $\sqrt{1-\rho}$ and $\sqrt{\rho}$. It is important to note that the correlation coefficient ρ, which is indeed a Pearson-correlation coefficient at the level of the random variables $Y_i = \Phi^{-1}(U_i) = \Phi^{-1}(PD_i(\tau_i))$, is not necessarily the Pearson correlation coefficient for the final variables explained (the τ_i), nor the "uniformalized" ones (the $U_i = PD_i(\tau_i)$). Note that in the case of a Gaussian copula, because the cross-correlation between a standard multivariate Gaussian vector determine entirely the distribution, the assumption that $Corr(Y_i, Y_j) = \rho$ for all i, j entirely characterizes the distribution.

Now, notice that we wrote:

$$Y_i = \sqrt{1-\rho}\Phi^{-1}(U'_i) + \sqrt{\rho}\Phi^{-1}(U'_{k+1})$$

Letting $\epsilon_i := \Phi^{-1}(U'_i)$ and $Z := \Phi^{-1}(U'_{k+1})$ this is

$$Y_i = \sqrt{\rho}Z + (\sqrt{1-\rho})\epsilon_i$$

ϵ_i is called the idiosyncratic factor associated to loan i while Z is called the systematic risk factor, as it is common to all Y_i. Hence simulating random variables for the purpose of using the (statistical) copula model yields a structural model, in which Y_i, which can be interpreted as the credit-quality of loan

i, is influenced both by an idiosyncratic factor ϵ_i and a systematic factor Z which affects all the loans of the portfolio. Because ϵ_i and Z are normally distributed by Lemma 1, this is precisely the *Vasicek one-factor model.* Notice that by independence of Z and ϵ_i, the two-dimensional random vectors (Z, ϵ_i) are Gaussians, and hence as a linear combination of them the variable Y_i is also Gaussian. This is the reason why the resulting model is sometimes called the *standard single-factor Gaussian copula model.*

The simulated default time t_i (realisation of the variable τ_i) is the time which solves:

$$PD_i(t_i) = u_i$$

where u_i is a realisation of $U_i = \Phi(Y_i)$. This solves into

$$\begin{aligned} 1 - \exp(-t_i\lambda_i) &= u_i \\ t_i &= \frac{\ln(1-u_i)}{-\lambda_i} \end{aligned}$$

7 Annex 2: Hidden Markov Chains for default rates

7.1 Efficient algorithms for the maximum likelihood (ML) estimation

7.1.1 Expressing the likelihood

The practical strength of hidden Markov chains (HMMs) is that there exists an efficient algorithm to compute the likelihood of the model being in a given state knowing the observation sequence. We describe below in a succinct but self-contained way the main results that served as a basis to implement the paper HMMs for default series.

Let N be the number of hidden states of the underlying Markov model. Without loss of generality we can label these states from 1 to N. In the paper we only considered two states: a low risk state, and a high risk state, hence $N = 2$, but the results and algorithms presented in this section can apply to an arbitrary number of states N. Let T be the length of the sample of observations. Without loss of generality, time is labelled from 1 to T. Let $A = (a_{ij})_{i,j}$ be the state-transition probability matrix of the model time-invariant hidden Markov chain; it is defined by:

$$a_{ij} = P(S_{t+1} = j \mid S_t = i)$$

For each state i in $\{1, 2, ..., N\}$, the number of defaults at time t, assuming the underlying state is i, is the realisation of a random variable which follows a discrete probability distribution function $B_t(i)$ (or, in the continuous-observation case, a probability distribution of *density* $B_t(i)$) supported by $\{1, 2, ..., n_t\}$ (in the discrete case) where n_t is the size of the sample at time t. For k in $\{1, 2, ..., n_t\}$ we denote by $B_t(i)(k)$ the probability of observing the value k under $B_t(i)$ (or, in the continuous-observation case, the density of k under $B_t(i)$).

Let $\pi = (\pi_1, \pi_2)$ be the initial distribution, and let $\lambda = (A, B, \pi)$ be the model. Let $S = S_1...S_T$ be the underlying sequence of states of the HMM. Let $O = O_1...O_T$ be the sequence of observations. Hence, O_t is the realisation of a random variable of law $B_t(S_t)$ and belongs to $\{1, 2, ..., n_t\}$. Using basic properties of Markov chains and the model definition it is clear that the probability $P(O, S)$ of observing the sequence O while having the underlying state sequence S is simply:

$$P(O, S) = \pi_{S_1} a_{S_1S_2} B_1(S_1)(O_1) a_{S_2S_3} B_2(S_2)(O_2)...a_{S_{T-1}S_T} B_T(S_T)(O_T)$$

Re-grouping similar terms yields:

$$P(O,S) = \pi_{S_1} \prod_{t=1}^{T-1} a_{S_t S_{t+1}} \prod_{t=1}^{T} B_t(S_t)(O_t)$$

Hence the log-likelihood can be written as:

$$\log(P(O,S)) = \log(\pi_{S_1}) + \sum_{t=1}^{T-1} \log(a_{S_t S_{t+1}}) + \sum_{t=1}^{T} \log(B_t(S_t)(O_t))$$

At this point it is convenient to introduce the following indicator functions $\gamma_t(i)$, for t in $\{1, 2, ..., T\}$, and $\gamma_t(i,j)$, for t in $\{1, 2, ..., T-1\}$, which are dependent on the random variables S_t, and defined by:

$$\gamma_t(i) = 1_{(S_t=i)} = \begin{cases} 1 \text{ if } S_t = i \\ 0 \text{ otherwise} \end{cases}$$

and

$$\gamma_t(i,j) = 1_{(S_t=i, S_{t+1}=j)} = \begin{cases} 1 \text{ if } S_t = i \textit{ and } S_{t+1} = j \\ \quad 0 \text{ otherwise} \end{cases}$$

The log-likelihood can be re-written as (notice the re-writting only consists in adding terms of zero values):

$$\log(P(O,S)) = \sum_{i=1}^{N} \gamma_1(i) \log(\pi_i) + \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{t=1}^{T-1} \gamma_t(i,j) \log(a_{ij}) + \sum_{i=1}^{N} \sum_{t=1}^{T} \gamma_t(i) \log(B_t(i)(O_t))$$

The algorithm consists in repeating a certain number of times the following two steps.

First, compute estimates $\widehat{\gamma_t(i)}$ and $\widehat{\gamma_t(i,j)}$ of the parameters $\gamma_t(i)$ and $\gamma_t(i,j)$, knowing $O = O_1...O_T$ and assuming the current values of the model (A, B, π). To this aim use efficient algorithms based on the recursive relations described below for computation, but rescal at each step to avoid underflow. These estimates $\widehat{\gamma_t(i)}$and $\widehat{\gamma_t(i,j)}$ replace, in the expression above, the true values $\gamma_t(i)$ and $\gamma_t(i,j)$.

Second, each term of the resulting maximum likelihood estimate is maximised with respect to the model parameters (A, B, π) separately. This is possible because π only appears in the first term, A in the second and B_t in the third. Moreover, notice that in the second term the $\sum_{j=1}^{N} \sum_{t=1}^{T-1} \gamma_t(i,j) \log(a_{ij})$ can be maximised independently (subject to the linear constraint $\sum_{j=1}^{N} a_{ij} = 1$), and that similarly the third term is made of the N separate sums $\sum_{t=1}^{T-1} \gamma_t(i) \log(B_t(i)(O_t))$ which can be maximised separately.

7.1.2 The first step

In what follows we detail efficient algorithms for computing the estimates of $\gamma_t(i)$ and $\gamma_t(i,j)$. For simplicity of notation, we will drop the hat, hence $\gamma_t(i)$ and $\gamma_t(i,j)$ stand for $\widehat{\gamma_t(i)}$ and $\widehat{\gamma_t(i,j)}$.

By definition,

$$\begin{aligned} \gamma_t(i) &= P(S_t = i \,|\, O, \lambda) = \frac{P(S_t = i\,, O \,|\, \lambda)}{P(O \,|\, \lambda)} \\ \gamma_t(i,j) &= P(S_t = i,\, S_{t+1} = j \,|\, O, \lambda) = \frac{P(S_t = i\,, S_{t+1} = j,\, O \,|\, \lambda)}{P(O \,|\, \lambda)} \end{aligned}$$

To compute $P(S_t = i\,, O \,|\, \lambda)$ efficiently, as well as $P(O \,|\, \lambda)$, we introduce the so-called "forward paths". Let i be in $\{1, ..., N\}$ and t be in $\{1, ..., T\}$. We define the forward path $\alpha_t(i)$ as:

$$\alpha_t(i) = P(S_t = i\,, O_1...O_t \,|\, \lambda)$$

Clearly

$$\alpha_1(i) = P(S_1 = i\,, O_1 \,|\, \lambda) = \pi_i.B_1(i)(O_1)$$

Then the others $\alpha_t(i)$, for t in $\{2, ..., T\}$ can be computed recursively by noticing that:

$$\alpha_t(i) = \sum_{j=1}^{N} \alpha_{t-1}(j).a_{ji}.B_t(i)(O_t)$$

The recursive formula is obtained by simply noticing that the only way to end up in state i at time t, for t in $\{2, ..., T\}$, while producing the observation sequence $O_1...O_t$ is to have been in some prior state j at time $t-1$ and having produced $O_1...O_{t-1}$ (probability $\alpha_{t-1}(j)$) and moving from state j to state i in the next step (probability a_{ji}) while producing the observation O_t in state i at time t (probability $B_t(i)(O_t)$). Independence of these events implies the product $\alpha_{t-1}(j).a_{ji}.B_t(i)(O_t)$ of these probabilities is the probability of such an event.

There are many ways to express $P(O \,|\, \lambda)$ as a function of the other introduced quantities:

$$\begin{aligned} P(O \,|\, \lambda) &= \sum_{i=1}^{N} P(S_T = i\,, O_1...O_T \,|\, \lambda) = \sum_{i=1}^{N} \alpha_T(i) \\ &= \sum_{i=1}^{N}\sum_{j=1}^{N} P(S_t = i\,, S_{t+1} = j,\, O \,|\, \lambda) = \sum_{i=1}^{N}\sum_{j=1}^{N} \gamma_t(i,j) \text{ for every } t \end{aligned}$$

Hence computing the forward paths gives an efficient algorithm to obtain the likelihood of observing O given the model λ.

Let i be in $\{1, ..., N\}$ and t be in $\{1, ..., T\}$. We define the "backward paths" $\beta_t(i)$ as:

$$\beta_T(i) = 1 \textit{ by convention,} \text{ and:}$$

$$\beta_t(i) = P(S_t = i\,, O_{t+1}...O_T \,|\, \lambda)$$

They can be computed recursively in a similar fashion to the $\alpha_t(i)$, but "backward", meaning, starting from $t = T-1$, as

$$\beta_t(i) = \sum_{j=1}^{N} a_{ij}.B_{t+1}(j)(O_{t+1}).\beta_{t+1}(j)$$

Indeed $\beta_t(i) = P(S_t = i\,, O_{t+1}...O_T \,|\, \lambda) = \sum_{j=1}^{N} a_{ij}.B_{t+1}(j)(O_{t+1}).P(S_{t+1} = j\,, O_{t+2}...O_T \,|\, \lambda)$.

Now, it follows trivially from the definition of $\alpha_t(i)$ and $\beta_t(i)$ that, for all t in $\{1, ..., T\}$ and all i in $\{1, ..., N\}$:

$$P(S_t = i\,, O \,|\, \lambda) = \alpha_t(i)\beta_t(i)$$

Hence the expression for $\gamma_t(i)$:

$$\gamma_t(i) = \frac{\alpha_t(i)\beta_t(i)}{P(O \,|\, \lambda)}$$

Similarly we can compute the $\gamma_t(i,j)$ from the $\alpha_t(i)$ and $\beta_t(i)$ as:

$$\begin{aligned} \gamma_t(i,j) &= P(S_t = i,\, S_{t+1} = j \,|\, O, \lambda) \\ &= \frac{P(S_t = i\,, S_{t+1} = j,\, O \,|\, \lambda)}{P(O \,|\, \lambda)} \\ &= \frac{\alpha_t(i).a_{ij}.B_{t+1}(j)(O_{t+1}).\beta_{t+1}(j)}{P(O \,|\, \lambda)} \end{aligned}$$

Also, notice that, for t in $\{1, ..., T-1\}$:

$$\gamma_t(i) = \sum_{j=1}^{N} \gamma_t(i,j)$$

This is the formula we used for computing $\gamma_t(i)$ from the $\gamma_t(i,j)$, with the exception of $t = T$.

The second step We know turn our attention to the maximisation of the log-likelihood estimate obtained. As mentionned earlier, the three terms appearing in the sum can be maximised separately. To prove all the three results above is simply done by equating the partial derivatives to 0. We do not detail the computation, but the two first results can be found in the literature (see for example, [30]).

a) The first term, $\sum_{i=1}^{N}\gamma_1(i)\log(\pi_i)$, is maximised for $\pi_i = \gamma_1(i)$.

b) The second for $a_{ij} = \frac{f_{ij}}{\sum_{k=1}^{N} f_{ik}}$ where $f_{ij} := \sum_{t=1}^{T-1}\gamma_t(i,j)$.

c) The third term, as noticed above, can be split for the purpose of maximisation into N distinct maximisation problems. The value or set of values which realise the maximum depends, of course, of the form assumed for the probability distribution functions $B_t(i)$. In the case of a binomial distribution of parameter n_t and p_i, we get:

$$\sum_{t=1}^{T}\gamma_t(i)\log(B_t(i)(O_t)) = \sum_{t=1}^{T}\gamma_t(i)(\log(\binom{n_t}{O_t}) + O_t.\log(p_t) + (n_t - x_t).\log(1-p_t))$$

whose maximum is attained at:

$$p_i = \frac{\sum_{t=1}^{T}\gamma_t(i).O_t}{\sum_{t=1}^{T}\gamma_t(i).n_t}$$

Notice that O_t is the number of defaults at time t, hence the number above makes sense; in particular p_i belongs to $[0,1]$. In the particular case where $n_t =: n$ is constant (as assumed for example when dealing with rates) meaning that at each step it is assumed that the number of re-introduced loans compensates exactly for the previous step defaults, we get the somehow simpler formula: $p_i = \frac{1}{n}\frac{\sum_{t=1}^{T}\gamma_t(i).O_t}{\sum_{t=1}^{T}\gamma_t(i)}$.

7.1.3 Obtention of confidence interval through parametric bootstrap

Parametric bootstrapping simply consists in simulating with the fitted model $\widehat{\Theta}$ a large number K of samples of the same size T and fitting each of the simulated sample to obtain parameter estimates $\widehat{\Theta_1}, \widehat{\Theta_2}, ..., \widehat{\Theta_K}$. The empirical distribution of these estimates $\widehat{\Theta_1}, \widehat{\Theta_2}, ..., \widehat{\Theta_K}$ then provides an indication about the model stability. In particular, an empirical variance-covariance matrix of all the model parameters can be computed in the usual way:

$$\widehat{Var}(\widehat{\Theta}) = \frac{1}{K-1}\sum_{i=1}^{K}(\widehat{\Theta_i} - \overline{\widehat{\Theta}})'.(\widehat{\Theta_i} - \overline{\widehat{\Theta}})$$

where $\overline{\widehat{\Theta}} := \frac{1}{K}\sum_{i=1}^{K}\widehat{\Theta_i}$ and $'$ denote the matrix transpose operator.

Notice that the initial guessed values for the estimate still have to be selected. As a rule, we decided to always select the same initial values than the one which led to the fitted model $\widehat{\Theta}$.

7.1.4 Obtention of h step forward forecasts from the model

Assume we want to obtain the h step forward forecast from the HMM model λ at time t. As we observed $O_1...O_t$, the probability of observing O_{t+h} in h periods from now, where O_{t+h} is, in this section, any element of the observation set $\{1, 2, ..., n_t\}$, is simply the quantity

$$P(O_{t+h} \mid O_1...O_t, \lambda) = \frac{P(O_{t+h}\,,\, O_1...O_t \mid \lambda)}{P(O_1...O_t \mid \lambda)}$$

Clearly, $P(O_1...O_t \,|\, \lambda) = \sum_{k=1}^{N} \alpha_t(k)$. For the numerator now, using in a similar way the formula for total probabilities, we can write:

$$
\begin{aligned}
P(O_{t+h}\,,\, O_1...O_t \,|\, \lambda) &= \sum_{i=1}^{N} P(O_{t+h},\, S_t = i,\, O_1...O_t \,|\, \lambda) \\
&= \sum_{i=1}^{N} \sum_{j=1}^{N} P(O_{t+h},\, S_{t+h} = j,\, S_t = i,\, O_1...O_t \,|\, \lambda) \\
&= \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_t(i).a_{ij}(h).B_{t+h}(j)(O_{t+h})
\end{aligned}
$$

where $a_{ij}(h) := P(S_{t+h} = j \,|\, S_t = i)$ is simply the (i,j) th coefficient of the matrix A^h (easy to check). Hence:

$$
P(O_{t+h} \,|\, O_1...O_t, \lambda) = \frac{\sum_{i=1}^{N}\sum_{j=1}^{N} \alpha_t(i).a_{ij}(h).B_{t+h}(j)(O_{t+h})}{\sum_{k=1}^{N} \alpha_t(k)} = \sum_{i=1}^{N} \sum_{j=1}^{N} \widetilde{\alpha}_t(i).a_{ij}(h).B_{t+h}(j)(O_{t+h})
$$

where $\widetilde{\alpha}_t(i) = \frac{\alpha_t(i)}{\sum_{k=1}^{N} \alpha_t(k)}$ are the rescaled forward paths.

In the case of one step forward forecasts for a two state HMM this becomes:

$$
\begin{aligned}
P(O_{t+1} \,|\, O_1...O_t, \lambda) \;=\;& \widetilde{\alpha}_t(1).a_{11}.B_{t+1}(j)(O_{t+1}) + \widetilde{\alpha}_t(1).a_{12}.B_{t+1}(2)(O_{t+1}) + \\
& \widetilde{\alpha}_t(2).a_{21}.B_{t+1}(1)(O_{t+1}) + \widetilde{\alpha}_t(2).a_{22}.B_{t+1}(2)(O_{t+1})
\end{aligned}
$$

Remark: in case we would like to obtain *out-of-sample forecasts* from a HMM model, notice that these forward forecasts are not strictly speaking out-of-sample forecasts when the parameter estimates of the fitted model have taken into account the observations after time t.

7.1.5 Assessing the goodness-of-fit of the model through the forward mid-pseudo residuals

Let $\widetilde{O}_{t+1}$ be the random variable indicating the number of defaults at time $t+1$. Hence, its realisation is O_{t+1}. From the above we can compute the empirical cumulative distribution function of $\widetilde{O}_{t+1}$ conditional on the model λ and on observing $O_1...O_t$: this is simply

$$
P(\widetilde{O}_{t+1} \le x \,|\, O_1...O_t, \lambda) = \sum_{k \le x} P(k \,|\, O_1...O_t, \lambda)
$$

where $P(k \,|\, O_1...O_t, \lambda)$ can be computed from the formula given in Section 7.1.4.

The one-step ahead forward pseudo-residual at time t is then, by definition:

$$
P(\widetilde{O}_{t+1} \le O_{t+1} \,|\, O_1...O_t, \lambda)
$$

By applying Lemma 2 at each different time t, for t in $\{1, ..., T\}$, we get that these forward pseudo-residuals are all uniformely distributed. This provides a way to assess the goodness-of-fit of the model: if the data-generating process is indeed equal or close to the fitted model, we should observe that the forward pseudo-residuals follow closely a uniform distribution. This can be check visually using histogram and (uniform) quantile-to-quantile plot.

A drawback of the use of a uniform distribution is that it is difficult to determine outliers, as low or high values are repectively mapped close to 0 or 1. Hence we further transform these uniformly distributed quantiles into standard normal ones using Lemma 1: if Φ is the cdf of a standard normal law, then the

$$
z_t = \Phi^{-1}(P(\widetilde{O}_{t+1} \le O_{t+1} \,|\, O_1...O_t, \lambda))
$$

are realisations of a standard normal law.

Because the probability distribution of the number of default is discrete, we make a last adjustement by considering the average of the uniform upper and lower quantiles before applying Φ^{-1}. Hence, we get:

$$z_t = \Phi^{-1}(\frac{P(X < O_{t+1} \,|\, O_1...O_t, \lambda) + P(X \leq O_{t+1} \,|\, O_1...O_t, \lambda)}{2})$$

These are called (normal) mid-pseudo residuals, and normality tests can be performed directly on the z_t.

7.1.6 Non-homogeneous hidden Markov chains

Markov chains can be made dependent on covariates $X(t)$, which can result in non-stationary processes for the coefficients of the transition probability matrix A. The algorithms described in the previous sections are still valid when one

- replaces the a_{ij} by $a_{ij}(t)$, where $a_{ij}(t)$ is the (i,j) th coefficient of the transition probability matrix at time t, denoted by $A(t)$,

- specifies a dependence structure betwen $A(t)$ and the covariates $X(t)$. For example, the loggistic transformation with $X(t)$ representing a single covariate gives the model:

$$\begin{aligned} a_{ii}(t) &= \frac{1}{\exp(X(t).\Psi_{ij} + \eta_{ij})} \\ a_{ij}(t) &= 1 - a_{ii}(t) \text{ for } i \neq j \end{aligned}$$

with Ψ_{ij} and η_{ij} parameters to estimate.

- replaces the second step of Section 7.1.2 by numerical maximisation, as no closed-formula is known.

Université Panthéon-Assas

Résumé :

La thèse étudie plusieurs problématiques centrales et actuelles de la finance moderne : la rationalité limitée des agents et leurs biais comportementaux vis-à-vis des valeurs nominales, le problème de la juste évaluation du prix des actions, la refonte du paysage de l'industrie post-négociation en Europe suite à l'introduction du projet de l'Eurosystème Target-2 Securities, ainsi que les modèles de défaut et les méthodes d'estimation des cycles de défaut pour un secteur donné. Les techniques employées sont variées: enquêtes sur données individuelles, économétrie, théorie des jeux, théorie des graphes, simulations de Monte-Carlo, chaînes de Markov cachées.

Concernant l'illusion monétaire, les résultats confirment la robustesse des résultats d'études précédentes tout en dévoilant de nouvelles perspectives de recherche, par exemple tenter d'expliquer la disparité des réponses selon les caractéristiques individuelles des répondants, en particulier leur formation universitaire. L'étude du modèle de la Fed montre que la relation de long terme entre taux nominal des obligations d'Etat et rendement des actions n'est ni robuste, ni utile à la prédiction sur des horizons temporels réduits. L'étude sur Target 2 Securities a été confirmée par les faits. Enfin, le modèle d'estimation des défauts à partir de chaînes de Markov cachées fait preuve de bonnes performances dans un contexte européen, malgré la relative rareté des données pour sa calibration.

Descripteurs :

illusion monétaire, modèle de la Fed, modèles de défauts, théorie des jeux, théorie des graphes, simulations de Monte-Carlo, chaînes de Markov cachées

Title and Abstract :

The thesis studies various themes that are central to modern finance : economic agents rationality and behavioural biases with respect to nominal values, the problem of asset fundamental valuation, the changing landscape of the European post-trade industry catalysed by the Eurosystem project Target 2 Securities, and models of defaults and methods to estimate defaults cycles for a given sector. Techniques employed vary: studies on individual data, econometrics, game theory, graph theory, Monte-Carlo simulations and hidden Markov chains.

Concerning monetary illusion, results confirm those of previous study while emphasizing new areas for investigation concerning the interplay of individual characteristics, such as university education, and money illusion. The study of the Fed model shows that the long term relationship assumed between nominal government bond yield and dividend yield is neither robust, nor useful for reduced time horizons. The default model based on hidden Markov chains estimation gives satisfactory results in a European context, and this besides the relative scarcity of data used for its calibration.

Keywords :

Monetary illusion, Fed model, game theory, graph theory, Monte-Carlo simulations, hidden Markov chains