THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITE DE TOULOUSE

Délivré par l'Université Toulouse Capitole

École doctorale : **Sciences Economiques-Toulouse School of Economics**

Présentée et soutenue par

DASTARAC Hugues

le 1er juillet 2020

Essays on Intermediation in Financial Markets

Discipline : **Sciences Economiques**

Unité de recherche : **TSE-R (UMR CNRS 5314 – INRA 1415)**

Directeur de thèse : Monsieur Bruno BIAIS, Professeur, HEC Paris

JURY

Rapporteurs Madame Sabrina BUTI, Professeure, Université Paris-Dauphine
Monsieur Jérôme DUGAST, Professeur associé, Université Paris-Dauphine

Suffragants Madame Fany Declerck, Professeure, Université Toulouse I Capitole

Essays on Intermediation in Financial Markets

Hugues Dastarac

PhD Thesis

Toulouse School of Economics
Université Toulouse I Capitole

Disclaimer

I gratefully acknowledge financial support from Banque de France during my PhD. The views expressed in this thesis are those of the author and do not necessarily reflect those of Banque de France.

Summary of the dissertation

The three chapters of this dissertation describe various aspects of the activity of intermediaries in financial markets. These intermediaries are often called *dealers*, referring to a legal statute in the US for companies whose main business is to trade in financial markets for their own account; or *market makers* by reference to their intermediation function, with some ambiguity about what types of activity market making covers. Dealers include the largest banks with activities in financial markets, and more recently some large asset management companies have set up important dealer subsidiaries.

In the first chapter I empirically investigate the trading activity of dealers in the US corporate bond market. To make terminology clear, I use the phrase *market making* in a narrow sense: buying or selling on customer demand, and re-selling or re-purchasing the same asset to other customers. Then I refer to non-market making trades as *proprietary trades*.[1] Using dealer transaction in investment grade corporate bonds and holding data, I show that on top of trading as market makers in a narrow sense dealers also trade on spreads between bonds that they likely perceive as mispricings, a form of proprietary trading.

First, I separate market making trades from proprietary trades in the following way. Market making trades are initiated by customers, so that customer buys are associated with price decreases and vice-versa, consistent with market making theories based on customer adverse selection and dealer inventory costs. By contrast, as in limits of arbitrage theories, proprietary trades are initiated by dealers, implying an opposite correlation between customer trades and price changes. To distinguish between the two, as the data do not indicate who initiate the trade, I separate transactions into two bins and assess the correlation in each bin. In the first bin (approximately half of the observations), customer buys are partially or completely offset by customer sells; in the second bin, all customer transactions are in the same direction. Using price impact regressions, I show that outside the 2007-2009 financial crisis, the first bin corresponds to market making, while the second corresponds to proprietary trading.

In a second step, I show that transactions in the second bin, and not in the first bin, correspond to dealers buying (selling) bonds that are cheap (expensive) with respect to Treasury bonds, and with respect to bonds of the same maturity. Such patterns are much stronger before the 2007-2009 crisis and are reduced afterwards. Trading on the corporate-Treasury spread may correct an actual mispricing, if the spread is partly explained by the lower liquidity of corporate bonds; however, for a given bond, the spread is also justified by a higher credit risk, whose fair value is notoriously hard to assess.

Consistently with exploitation of the corporate-Treasury bond spread, I show

[1] Strictly speaking, any outright purchase or sale by a dealer is trading for its own account, which includes market making transactions. I follow the terminology used in regulatory debates by labelling proprietary trading any trade for the dealer's account that is not market making.

that the largest dealers held increasingly large corporate bonds inventory, and a mirroring inventory of bonds borrowed and sold (*i.e.* sold short) in the years leading to the 2007-2009 crisis, while these positions have not been rebuilt after the crisis. Uncovering this strategy in particular sheds light on the role of dealers in shaping corporate bond credit spreads, whose aggregate levels has been shown to impact firms' funding costs and thus to forecast economic activity.

In the second chapter I theoretically investigate why dealers trade a broad class of derivative contracts - forwards and futures, and to some extent swaps - instead of trading the underlying asset directly. I propose a dynamic equilibrium model in which dealers trade excess inventories progressively among each other, because of imperfect competition. However, in the course of trading, some other investors can post unexpected orders for reasons unrelated to dealers' asset valuation, which creates a risk to which buying and selling dealers have opposite exposure: buyers fear that other investors buy at the same time as them, making the price increase, while sellers fear that other investors sell, for symmetric reasons. This creates gains from trading the risk created by investor supply or demand shocks, which can be implemented or more generally approximated, with forward or future contracts of maturity shorter than dealers' horizon.

These derivatives contract slow down and decrease the sharing of the underlying asset risk, implying a gross welfare loss for dealers; however, this loss is more than compensated by the benefit of a surer surplus of future transactions, *i.e.* a better management of dynamic trading.

On the positive side, the equilibrium with derivatives and longer inventory holdings by dealers is reminiscent of the pre-2007-2009 crisis situation; while the equilibrium without derivatives reminds of the post-crisis situation, with regulatory limits to leverage for larger dealer banks and dealers less willing to hold inventories for a long time. This model does not model default and financing constraints that real-world dealers face and that appear crucial in explaining amplification of financial crises.

In the third chapter, I explore an underlying reason why financial markets are fragmented. Market fragmentation underlies mispricings between assets as pointed out in the first chapter, and inventory imbalances between dealers that are the basis of derivative trading in the second chapter. I show that under imperfect competition and dynamic trading, dealers choose to open a parallel market to trade with customers in order to extract more trading rents, which makes customers worse off.

Résumé de la thèse

Les trois chapitres de cette thèse décrivent différents aspects de l'activité des intermédiaires sur les marchés financiers. Ces intermédiaires sont souvent appelés *dealers* (négociants), par référence à un statut légal américain pour les entreprises dont l'activité principale est l'achat et la vente pour compte propre sur les marchés financiers; ou bien *market makers* (teneurs de marché), avec une ambiguïté sur ce que la tenue de marché recouvre exactement. Parmi les *dealers*, on compte les plus grandes banques ayant des activités sur les marchés financiers, tandis que plus récemment de grands gestionnaires d'actifs ont créé des filiales sous statut de *dealer*.

Dans le premier chapitre, nos investigations se portent sur les activités de trading des dealers sur le marché américain des obligations d'entreprises. Pour préciser la terminologie, nous réservons le terme *tenue de marché* á des transactions exécutées á la demande des clients, une transaction opposée étant exécutée plus tard avec d'autres clients. Les autres transactions sont regroupées sous le terme de *trading pour compte propre*[2]. Utilisant des données de transactions par les dealers sur obligations de qualité de crédit *investment grade*, et des données de détentions, nous montrons qu'en plus d'agir comme teneur de marché dans un sens restreint - acheter ou vendre à la demande d'un client, et revendre ou racheter à d'autres clients - les *dealers* exploitent aussi des écarts de prix qu'ils jugent non justifiés par les fondamentaux des émetteurs, ce qui est une forme de trading pour compte propre (*proprietary trading*).

Dans un premier temps, nous séparons les transactions relevant de la tenue de marché de celles qui relèvent du trading pour compte propre de la manière suivante. Les transactions relevant de la tenue de marché sont initiées par les clients, de sorte que les achats par les clients sont associés à des augmentations de prix et réciproquement, ainsi qu'il est prédit par les théories de tenue de marché en présence d'antisélection des clients, ou de coûts d'inventaire chez les dealers. Au contraire, comme dans les théories de limites à l'arbitrage, les transactions pour compte propre sont initiées par les *dealers*, ce qui implique une corrélation opposée entre les transactions des clients et les variations de prix. Pour distinguer entre les deux, dans la mesure oú les données n'indiquent pas qui initie la transaction, nous séparons les transactions en deux paniers et mesurons la corrélation entre transactions des clients et changements de prix dans chaque panier. Dans le premier panier, qui représente approximativement la moitié des observations, les achats de clients sont partiellement ou totalement compensés par des ventes de clients; dans le second panier, toutes les transactions des clients sont dans la même direction. Nous utilisons des régressions d'impact sur les prix pour mesurer les corrélations dans chacun des paniers, et trouvons qu'en dehors de la crise de 2007-2009, le premier panier

[2]Le trading pour compte propre désigne à proprement parler toute transaction d'un dealer pour compte propre et non pour le compte d'un client (ce que ferait un courtier), de sorte que la tenue de marché relève du trading pour compte propre. Nous suivons néanmoins la terminologie utilisée dans les débats réglementaires, qui distinguent entre la tenue de marché et le reste, appelé trading pour compte propre.

correspond à la tenue de marché, le second au trading pour compte propre.

Dans un second temps, nous montrons que les transactions du second panier, et seulement celles-ci, correspondent à des achats (ventes) par les *dealers* d'obligations peu chères (chères) comparées aux obligations du Trésor américain, et comparées à d'autres obligations d'entreprises de même maturité. Le résultat est beaucoup plus fort avant la crise de 2007-2009, et réduit après la crise. L'exploitation de l'écart de taux entre les obligations d'entreprise et du Trésor peut corriger des aberrations, si cet écart est partiellement expliqué par l'illiquidité des obligations d'entreprises; cela étant, pour une obligation donnée, l'écart de taux est aussi justifié par un risque de crédit plus important, mais dont la juste valeur est notoirement difficile à établir.

De façon cohérente avec l'exploitation de l'écart de taux entre obligations du Trésor et d'entreprises, nous montrons que dans les années précédant la crise de 2007-2009, les plus gros *dealers* ont détenu des stocks d'obligations d'entreprise de plus en plus grands, et des stocks symétriques d'obligations du Trésor vendues à découvert (*i.e.* empruntées et vendues), tandis que ces positions n'ont pas été reconstituées après la crise. La mise en évidence de cette stratégie éclaire le rôle des *dealers* dans la formation des spreads de crédit des obligations d'entreprises, dont il a été montré que le niveau agrégé influe sur les coûts de financement des entreprises et ainsi prédit statistiquement le niveau d'activité économique.

Dans le deuxième chapitre, nous répondons théoriquement à la question suivante: pourquoi les *dealers* échangent des dérivés de type *forward* ou *future* (*i.e.* des contrats à terme) plutôt que d'échanger,directement l'actif sous-jacent. Nous proposons un modèle d'équilibre dynamique dans lequel les *dealers* s'achètent et se vendent leurs inventaires excédentaires ou déficitaires de façon *progressive*, en raison d'une concurrence imparfaite. Cependant, en cours d'achat ou de vente, d'autres investisseurs peuvent également acheter ou vendre de façon imprévue, ce qui fait monter ou descendre le prix: ces transactions représentent un risque pour les dealers, auquel ils sont exposés de façon opposée: les acheteurs craignent que les autres investisseurs achètent en même temps qu'eux, les vendeurs que les investisseurs vendent. Cette exposition opposée crée des gains à échanger le risque lié à ces achats ou ventes imprévues, ce qui peut être réalisé ou approximé par des contrats à terme de maturité plus courte que l'horizon de trading des dealers.

Ces produits dérivés ralentissent l'échange et diminuent la quantité d'actif sous-jacent échangée, ce qui implique une perte brute pour les *dealers*; cependant cette perte est plus que compensée par le bénéfice d'un surplus des transactions futures plus sûr, c'est-á-dire d'une meilleure gestion de l'échange dynamique.

Sur le plan descriptif, l'équilibre avec produits dérivés ressemble à la situation qui prévalait avant la crise de 2007-2009; tandis que l'équilibre sans dérivés ressemble à la situation post-crise, avec des réglementations limitant l'effet de levier des grandes banques, et des dealers moins enclins à porter longtemps des stocks. Ce modèle n'incorpore pas de possibilité de défaut ni de contraintes financières auxquelles les dealers sont soumis en réalité, et qui apparaissent cruciales pour expliquer l'amplification des crises.

Dans le troisième chapitre, nous explorons une raison possible pour laquelles les marchés financiers sont fragmentés. La question de la fragmentation est sous-jacente à celle des aberrations de prix évoquées dans le premier chapitre, ainsi que les équilibres d'inventaires qui sous-tendent l'utilisation de produits dérivés dans le deuxième chapitre. Nous montrons théoriquement qu'en concurrence imparfaite et avec plusieurs opportunités de trading dans le futur, les dealers choisissent d'ouvrir un marché parallèle, *i.e.* de fragmenter le marché, pour extraire de la rente de leur clients, ce qui nuit à ces clients.

Acknowledgements

I am indebted and extremely gratetul to Bruno Biais for his invaluable guidance and support. I am also very grateful to Thierry Foucault, whose contribution was also decisive at various steps of my work.

I would also like to thank Jean-Edouard Colliard, Denis Gromb, Stéphane Guibaud, Augustin Landier, Anders Trolle for many useful discussions, and Benoît Mojon, whose help was decisive at the beginning of my PhD. I am very grateful to Bruno Biais, Augustin Landier and David Thesmar for organizing my stay and welcoming me at MIT.

I am very grateful to Banque de France for financial support during my PhD.

I am deeply indebted to many friends as well, for their support, feedback and precious advice at various stages of my PhD and for always stimulating discussions. Among many others, I would like to thank, in alphabetical order, Jean Barthélemy, Valère Fourel, Mattia Girotti, Yves Le Yaouanq and Eric Mengus.

I would like to thank my parents. I also thank my parents-in-law in particular for welcoming us during the lockdown period, allowing me to finish this thesis in wonderful conditions.

Last but no least, I would like to thank my wife Anne just for her being and for her support, patience and love; and my children Armand and Thaïs for who they are, and who also took their part in supporting me during these years.

Contents

Chapter 1

Market Making and Proprietary Trading in the US Corporate Bond Market

Market Making and Proprietary Trading in the US Corporate Bond Market

Hugues Dastarac[1]

[1]I am deeply indebted to my PhD advisor Bruno Biais for his invaluable guidance, and very grateful to Thierry Foucault for pivotal discussions, to Jean Barthélemy, Jean-Edouard Colliard, Valère Fourel, Mattia Girotti, Denis Gromb, Alexander Guembel, Stéphane Guibaud, Ulrich Hege, Sophie Moinas, Benoit Nguyen, Angelo Ranaldo, Norman Schuerhoff, Anders Trolle, seminar participants at Banque de France, HEC Paris Finance PhD workshop, Toulouse School of Economics, HEC Lausanne and HEC Montréal for very helpful discussions and comments. All errors are mine. I gratefully acknowledge financial support from Banque de France. The views expressed here are those of the author and do not necessarily reflect those of Banque de France.

Abstract

I study broker-dealers' trading activity in the US corporate bond market. I find evidence of market making and of proprietary trading exploiting possible mispricings. Market making occurs when customers both buy and sell a bond in a day, which happens half of the time: as predicted by market making theories with adverse selection or inventory costs, prices go down (up) as customers sell (buy). Otherwise, evidence is in favor of broker-dealer initiated trades, *i.e.* proprietary trading: prices go up (down) when customers sell (buy). I test one aspect of proprietary trading predicted by theories of limits of arbitrage: dealers buy (sell) bonds that are relatively cheap (expensive) with respect to bonds of similar maturity, or with respect to Treasury bonds. These proprietary trading strategies are reduced after the crisis. Relatedly I show that before the 2007-2009 crisis, large broker-dealers borrowed and sold Treasury bonds in amounts similar to their corporate bond holding, but not after.

1.1 Introduction

Dealers are core intermediaries in financial markets. There is ambiguity about what they do, while knowing more about it would help understand their impact on asset prices,[1] their role in the 2007-2009 crisis and to assess subsequent regulation. They are often viewed as *market makers*: they buy or sell on customer demand, then revert the trade. But the macro-finance literature suggests that they trade actively, which is *proprietary trading*,[2] but remains elusive on the underlying trading activity. Is an empirical distinction between proprietary trading and market making possible? What are broker-dealers' trading strategies? What are the associated risks?

To answer these questions I empirically study dealer transactions in the US corporate bond market. I find evidence of both market making and proprietary trading. About half of bond×day observations contradict predictions of market making theories: prices tend to go up (down) when broker-dealers' customers sell (buy). Second, for these observations, broker-dealers buy (sell) more bonds that were cheap (expensive) compared with other bonds, as in limits of arbitrage theories. Third, I give suggestive evidence that proprietary traders' financing constraints generate risks for them, which materialized at the onset of the crisis in July 2007, with no obvious link to their prop trading strategies. Fourth, after the crisis, the proprietary trading strategies I documents are reduced.

I use customer-to-dealer transaction data for a large sample of liquid, investment grade bonds from FINRA's TRACE reporting system. On top of having macroeconomic relevance[3] the US corporate bond market also allows to focus on broker-dealer trades. In other markets like equity markets, isolating broker-dealer trades is generally impossible because the type of trader behind a given order is not disclosed.

To distinguish between market making and proprietary trading, I rely on testable predictions from economic theories. The first prediction is about who initiates the trade, implying opposite correlations between short-term price changes and customer trades. Under market making, the customer initiates the trade, implying that prices go down (up) when customers sell (buy).This is because a customer can be informed so his trade signals asset value,[4] or because risk averse market makers require a higher expected return to hold more risk.[5] By contrast, I understand proprietary trading broadly: the dealer compares prevailing prices with his/her assessment on the asset value, based on public and/or private information, and trades on it. Thus under dealer proprietary trading, dealers initiate the trade, and prices should go up when dealers buy, *i.e.* when customers sell, and conversely.

[1] See for instance Adrian, Etula, and Muir (2014), He, Kelly, and Manela (2017), Adrian and Shin (2009, 2010), Gilchrist and Zakrajsek (2012), Rapp (2016) and Siriwardane (2019).

[2] Proprietary trading can be viewed as an extended notion of market making: this paper is then about documenting new aspects of market making.

[3] Corporate-Treasury spreads forecast economic activity and recessions: *cf.* in particular Philippon (2009), Gilchrist and Zakrajsek (2012), Gilchrist and Mojon (2017), Lopez-Salido, Stein, and Zakrajsek (2017).

[4] Kyle (1985), Glosten and Milgrom (1985)

[5] Stoll (1978), Ho and Stoll (1981), Grossman and Miller (1988).

Then I test a second prediction from one theory of proprietary trading: that dealers trade on spreads in the cross-section of bonds that they think not justified by issuer's fundamentals. Then he/she buys the cheaper asset, and sells the more expensive asset and expects a profit. This is the situation described by theories of limits of arbitrage,[6] with proprietary traders being the analogs of arbitrageurs. I do not assess whether a low price is *too* low.

As first suggestive evidence of this form of dealer proprietary trading, figure 1.1 plots corporate bonds and Treasury bonds holding of Primary Dealers, *i.e.* large broker-dealers for which the New York Fed releases inventory data, and corporate-to-Treasury spreads indices for A and BBB bonds.[7] Primary Dealer trading activity is in my TRACE sample. Strikingly before the crisis, the large net corporate holdings were mirrored by negative net Treasury holdings, a fact not reported in the literature: it means Primary Dealers borrowed and sold Treasury bonds, *i.e.* held a short position. If this was proprietary trading, Treasury bond prices should go down and corporate bond prices should go up: this is what corporate spreads indicate. The BBB index even decreases as Primary Dealers holdings go up. After the crisis, the long/short position is not rebuilt, while spreads are higher and more volatile.

Thus I expect broker-dealers to do both proprietary trading and market making: to distinguish between the two, I find an intuitive criterion and check whether it holds in the data. The criterion is as follows. Suppose that on a given day for a given bond, order flow is a *partial* or *full roundtrip*: customers both buy and sell within a day. This may reveal that dealers have, say, bought bonds from customers but are unwilling to hold them so start to re-sell them within the day: this suggests market making. By contrast, suppose order flow is *one-way*: customers only buy or sell. This may be either because they they are making market but are not able to reverse the trade quickly, or because they are actually willing to hold the position because it enters a proprietary trading strategy.

To test the prediction about price changes and customer trades, I use price impact regressions, a standard tool in empirical market microstructure: I regress price changes on customer daily purchases net of daily sales, controlling for plausible determinants of these changes including stock return, long and short rate changes.

I find that the coefficient for one-way order flow is as predicted by limits of arbitrage theories, except during the crisis, while it is as consistent with market making on days with partial roundtrips. Running the price impact regression without distinguishing between one-way and partial roundtrips, the coefficient is consistent with market making, which suggests raw price impact regressions may mislead to a conclusion that broker-dealers are always passive market makers. Quantitatively, on average a $1 million customer net sale is associated with a price increase by 1.5 basis point. To explore heterogeneity within one-way observations, I add interactions

[6] *Cf.* Gromb and Vayanos (2002, 2010, 2018).

[7] Median spread in bonds from my sample with 4-6 years residual maturities. The spread is the log price difference between a fictitious risk-free price of a bond with the same cash flows discounted with the Treasury yield curve.

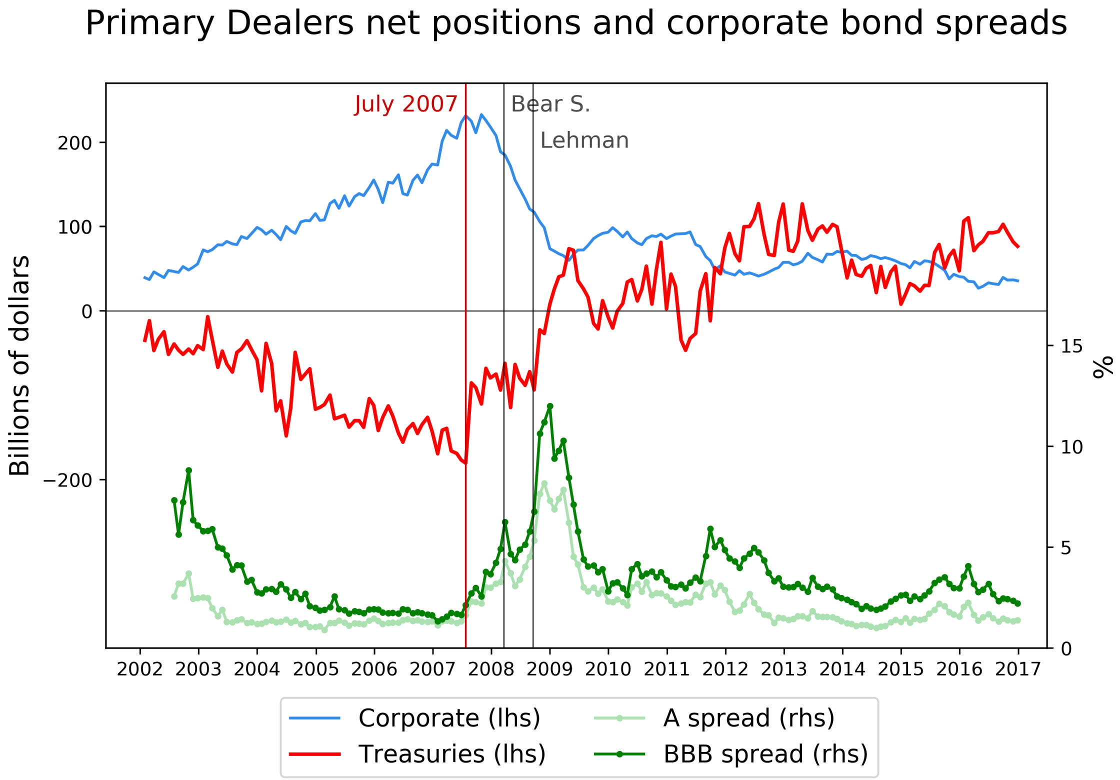

Figure 1.1: Primary Dealers holdings in corporate bonds and Treasury bonds (all maturities above 3 months), and corporate-Treasury spreads for A and BBB bonds with 4-6 years residual maturity. Holding data are from the New York Fed, spreads computed from my sample, all data are as of the last Wednesday of each month (holding data are published weekly as of Wednesday).

with dummies for the bond's initial or residual maturity, or its age, being less than 10 years: I find that proprietary trading is concentrated on bonds with initial or residual maturity less than 10 years, or on bonds younger than 10 years.

I look at the evolution through time. I interact one-way and partial rountrip with dummies for time periods: I find that one-way order flow has coefficients consistent with proprietary trading both before and after the crisis; before the crisis, customer sales by $1 million come with an increase by 5 basis points, while it reduces to 1.7 basis point after the Dodd-Frank Act. During the crisis, the one-way coefficient is consistent with market making and higher than the partial roundtrip coefficient, suggesting that dealers were not willing to hold the bonds they bought (or hold the inventory deficit) for long, but expected to be unable to re-sell (or repurchase) them quickly.

I also discuss endogeneity concerns. In particular to address the concern that there are common drivers of order flow and price changes, I control for risk-free rate changes and credit risk changes through stock return, two main sources of corporate bond price change and likely of customer willingness to trade.

Then I test the second prediction, by regressing customer order flow on lagged measures of bond cheapness, controlling for lagged order flow, lagged price changes and market factors. The cheapness measures are spreads between baskets of bonds.

I first test a measure in which a bond's spread to an equivalent Treasury is compared with the median spread in a basket of bonds with similar maturity, credit rating and callability (presence or absence of an embedded call). This measure captures idiosyncratic component of the bond spread. I further split one-way and roundtrip order flow based on bond initial or residual maturity and bond age being more or less than 10 years. I find that the measure is significant for one-way order flow and in the expected direction for bonds with initial or residual maturity less than 10 years, and weaker but significant results for bond age. The estimates imply that an increase by one percentage point in the measure is associated with dealer purchases being higher by 12%, and conversely.

I add three measures of cheapness in the regression, which are differences in median corporate-to-Treasury spreads of different baskets of bonds. The first additional measure captures a bond's cheapness relative to bonds with similar maturity irrespective of their credit risk and callability, having controlled for idiosyncratic components. The second additional measure captures arbitrage between bonds of different maturities, having controlled for its credit risk and callability. The third additional measure captures corporate bond cheapness with respect to Treasuries: the former may be more expensive than the latter even after controlling for credit risk, because of liquidity or other services they provide. It may be on average profitable to sell Treasury bonds to buy corporate bonds. I call this measure the Treasury convenience yield component. Primary Dealers appear to exploit it on figure 1.1.

I find significant effects on the credit risk/callability component, and on the Treasury convenience yield component. When I further distinguish by age or maturity, I find unchanged effect for the idiosyncratic component, and highly significant effects for the credit risk/callability and Treasury convenience yield components for bonds with age or maturity below 10 years. The maturity component does not appear exploited by dealers. For roundtrip order flow, coefficients are zero or small in absolute value, consistently with price impact regressions. Quantitatively, I find that an increase in the credit risk/callability measure by one percentage point for a bond with maturity lower than 10 years is associated with higher dealer purchases by 24%. For the Treasury convenience yield component, the effect is 15%. Breaking down by subperiods, the corresponding coefficients imply effects of both measures higher than 60% before the crisis and below 20% after the crisis, in line with figure 1.1.

These results suggest that 1) broker-dealers do proprietary trading on top of market making and 2) my measures distinguish the former from the latter, which is useful for existing regulation and 3) proprietary trading was weaker after the crisis.

Regarding the proprietary trading risks, I give suggestive evidence from the 2007-2009 crisis that a crucial risk is not obviously linked with the underlying strategy, but comes from margin requirements or regulatory capital requirements. Figure 1.1 shows that Primary Dealers' net Treasury position shrank by half in July 2007, months before major crisis events. If concerns were about the corporate-Treasury position, lenders would probably have imposed scaling down of both legs: this is not the case. Instead, I show that the short position cut occured a few weeks after 1) the historical volatility of daily return on Treasuries increased with respect to the previous 18 months and 2) Primary Dealers' CDSs, a proxy for their perceived default risk, abruptly rose. By contrast corporate bond volatility did not vary much and the corporate bond position was not cut. Primary Dealers also faced financing constraints: I show that the long and short position were funded independently through short-term repurchase agreements, which imply haircuts. Another hypothesis, with equivalent effect, is that regulatory capital requirements, computed with statistical models for large dealers. This suggests that Primary Dealers faced financing constraints that increasing in asset volatility and their own default risk:[8] a tightening of the short side occured in July 2007.

Literature review. Several papers question the assumption that broker-dealers are always passive, as Choi and Huh (2019): I further show that broker-dealers initiate trades also to buy relatively cheap and sell expensive bonds. An (2019) shows that broker-dealers initiate trades to build a wide menu of bonds to match buyers' preferences. Effects described in his and my papers are fully compatible; empirically, he focuses on reversals within 15 minutes, while I focus on reversals within a day. Other papers hint at broker-dealer proprietary trading but do not test the hypothesis. Adrian et al. (2017) show that before the crisis, bond market liquidity was positively related with broker-dealer leverage, but negatively after the crisis: I exhibit long-short positions by Primary Dealers to explain this fact. In FX markets, Du et al. (2018) show that mispricings are stronger for positions appearing on quarterly financial statements. In equity markets, (Brogaard et al. 2014, 2019)

[8] Which may come from other activities than prop trading, *e.g.* mortgage-backed securities.

show that high-frequency traders (HFT), which are also dealers, initiate trades.

This paper is also connected to the literature assessing post-crisis regulation. Many papers focus on broker-dealer market making activity, showing that dealers tend to hold position for shorter periods of time (Bessembinder, Jacobsen, Maxwell, and Venkataraman 2018, Schultz 2017); Duffie (2018) and Saar, Jian, Yang, and Zhu (2019) suggest this is not necessarily a bad thing. Bao et al. (2018) and Dick-Nielsen and Rossi (2018) show that abnormal returns surrounding bond downgrades to speculative grades have increased after the crisis. Dick-Nielsen and Rossi (2018) also mentions decreased Primary Dealers' corporate bond inventories: I connect this pattern with the Treasury short position. Overall my contribution with respect to these papers is to separate proprietary trading from market making.

Goldstein and Hotchkiss (2020) show that for the *least* traded bonds, in about 60% of the cases, inventory holding is less than one day. While I look at *most* traded bonds for which proprietary trading is more likely, the fact they uncover is the basis of my identification. Some papers look at dealer networks (Di Maggio et al. 2017, Friewald and Nagler 2019, Li and Schuerhoff 2019). Overall I confirm that a large fraction of broker-dealer trading activity is market making, but I show that broker-dealer also provide liquidity through prop trading.

Price impact regressions have seldom been run on the US corporate bond market.[9] Rapp (2016) runs such regressions with a different focus and finds results consistent with market making. I find qualitatively similar results on average, but I isolate a subset of proprietary trading transactions.

This paper also relates to the macro-finance literature on broker-dealers, which does not explore broker-dealer trading strategies. Adrian and Shin (2009, 2010) show that broker-dealers' leverage ratio is positively correlated with asset prices, and in particular with the market price of risk. Gilchrist and Zakrajsek (2012) show that negative shocks to Primary Dealers' equity is translated in higher corporate bond/Treasury bond yield spread. Adrian, Etula, and Muir (2014) and He, Kelly, and Manela (2017) show that Primary Dealer leverage is an important factor explaining asset prices.

The paper is divided as follows. Section 1.2 develops the empirical hypotheses. Section 1.3 gives institutional background and presents the main dataset with summary statistics. Section 1.4 presents the price impact regressions results. Section 1.5 studies the relationship between customer order flow and bond cheapness measures. Section 1.6 suggests that the end of Primary Dealers' long-short strategy shown by figure 1.1 was caused by a tightening of the margin requirement on the short Treasury position. Section 1.7 discusses the implications for financial regulation and for safe asset production by the private sector. Section 2.8 concludes.

[9] Possibly because pre-trade quotes are not available. In other markets price impact regressions are standard: in the stock market, seminal papers include Glosten and Harris (1988), Hasbrouck (1991), Madhavan and Smidt (1993), Huang and Stoll (2015). In FX market, cf. Lyons (1995), Evans and Lyons (2002). *Cf.* Collin-Dufresne, Junge, and Trolle (2018) in the index CDS market.

1.2 Empirical hypotheses

1.2.1 Theoretical predictions

Market making and proprietary trading appear complementary activities that are hard to separate, leading practioners and some academics (*e.g.* Duffie 2012) to make no difference between the two. Indeed both activities are in the end about buying low from investors eager to sell, selling high to investors eager to buy: both are thus liquidity provision.

In this paper I choose to make a difference between the two, based on clear theoretical predictions and because it sheds new light on what broker-dealers do. For convenience I borrow the terminology from the Volcker rule, which does *not* imply any stance on the optimality of the Volcker rule, which is neither the only way nor necessarily the best way to regulate proprietary trading: such assessment is beyond the scope of this paper.

Price changes and order flow. Theories of market making and proprietary trading give opposite predictions regarding the correlation between short-term price changes and customer trades. Theories of market making predict that customer sales (purchases) are associated with price decreases (increases). This is because a customer can be informed and his trade may signal asset value (Kyle 1985, Glosten and Milgrom 1985) or because risk averse market makers require a higher expected return to hold more risk (Stoll 1978, Ho and Stoll 1981, Grossman and Miller 1988).

By contrast, under dealer proprietary trading prices should go up when dealers buy, i.e. customers sell. I understand dealer proprietary trading as dealers comparing prevailing asset prices to their private or public information, and then trade on it. Dealer information can be private, in which case the dealer plays the role of the informed trader within market making theories. Dealer information can also be public, as when there is a spread between two publicly listed assets that appears not justified by fundamentals, but which persists because of some market frictions. In this case dealers are expected to buy the cheaper asset and sell the more expensive asset: except that the strategy possibly involves risk over the strategy's terminal payoff, this is exactly the situation described by theories of limits of arbitrage (Gromb and Vayanos 2002, 2010, 2018) with proprietary traders being the analogs of arbitrageurs.

Drivers of transactions. Theories of limits of arbitrage therefore give a second prediction: dealers should buy (sell) more a bond i that trade at low (high) price with respect to other bonds, given bond i's and other bonds' respective characteristics. In this paper I test this prediction giving up on whether a low price is indeed *too* low.

However, proprietary traders may have other strategies that those predicted by theories of limits of arbitrage: dealers may be informed on the issuer's credit risk or on order flow. I do not test it in this paper however.

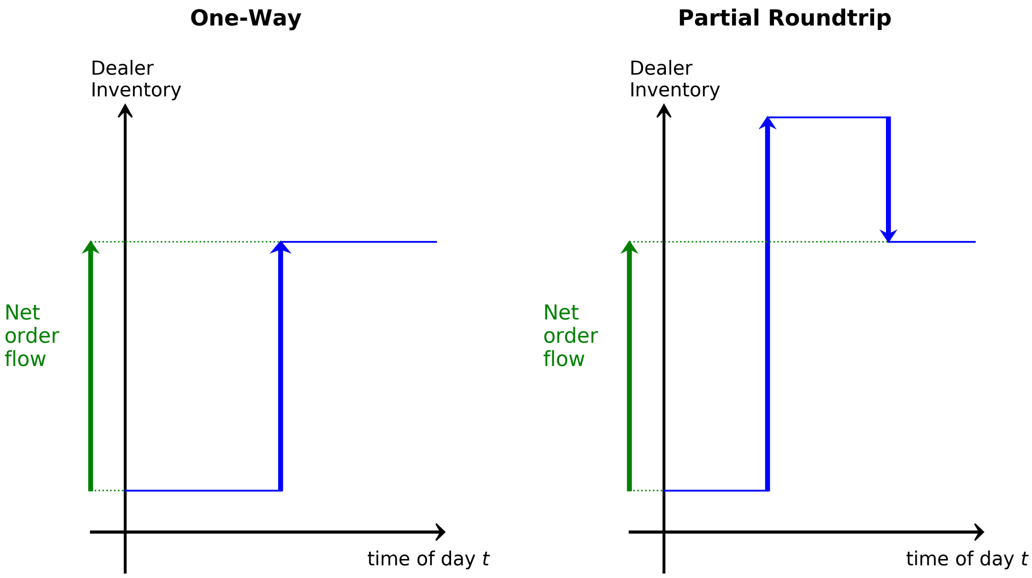

Figure 1.2: **One-Way *vs.* Partial Roundtrip order flow.** This figure illustrates how some days, customer net order flow may result from transactions all in the same direction (only customers buys or only customer sales) as in the left panel, or in both directions as in the right panel. Partial Roundtrip order flow is likely to correspond to market making, because it suggests dealers were unwilling to hold bonds sold by customers for long; One-way order flow is likely to correspond to dealer proprietary trading because it suggests dealers were willing to hold all bonds sold by customers.

1.2.2 One-way and Roundtrip days

I expect broker-dealers to do both market making and proprietary trading, which should correspond to different subsets of transactions.

To identify these subsets I separate days where large order flow goes only one way – only customer buys or only customer sells – from days where large order flow results from partial or full roundtrip – a customer buy is partially or fully offset by a customer sell – as illustrated by Figure 1.2.

I expect days with partial Roundtrip order flow to be associated with market making. A market maker expects to re-sell (repurchase) a security he has bought (sold) at some horizon. Partial Roundtrip order flow suggests that dealers were not willing to hold inventory deviations from their target overnight, and that they were able to revert part of customer initial transactions to get closer to the target.

This hypothesis is consistent with evidence by Goldstein and Hotchkiss (2020) on the least liquid US corporate bonds: dealers tend to revert transactions within a day in 58.2% of the roundtrips they study. In a related way, Li and Schuerhoff (2019) show in the US municipal bond market that when some customers demand immediacy, dealers match them more directly with other customers, meaning that a customer-to-dealer is more quickly associated with other customer-to-dealer trades in the other direction. Thus if my hypothesis is true, this means that customers initiate the trades and dealers accommodate, meaning these trades correspond to market making by dealers.

By contrast, I expect days with One-way order flow to be associated with proprietary trading. In principle One-way order flow could reflect one of two events. Either dealers are willing to hold the bonds they have purchased from customers, because it enters a proprietary trading strategy. Or dealers are unwilling to hold the position overnight, but they are not able to start to revert their position because none of their customers is willing to buy on short notice: in this case it corresponds to market making. Thus I expect the first motive to dominate.

Do I miss roundtrips in the interdealer market? Given that I only take customer-to-dealer trades into account to assess whether there is a partial roundtrip, I may miss situation where dealer A trades with a customer and reverts the trade in the interdealer market.

This causes no problem for my purposes: if dealer A reverts a trade with dealer B and one sees no reversal with another customer, it means that dealer B was willing to keep the position, for proprietary trading purposes if my hypothesis is true. From the perspective of the broker-dealer sector as a whole, it means that one dealer is willing to enter proprietary trading.

1.2.3 Hypotheses

Crossing the theoretical predictions from subsection 1.2.1 with the hypothesis that One-way order flow is associated with prop trading, and Roundtrip order flow with market making, I formulate the following testable hypotheses.

Hypothesis 1. *On days with one-way order flow in a given bond, price increases (decreases) are correlated with customer net sales (buys).*

On days with partial roundtrip order flow in a given bond, price increases (decreases) are associated with customer net buys (sales).

Hypothesis 2 (Limits of arbitrage)**.** *On days with one-way order flow, dealers tend to buy (sell) bonds that are cheap (expensive) compared with other bonds.*

In this paper I understand “cheap” and “expensive” in a broad and agnostic way. Broad, because bonds with very different risk level or maturity, for instance, can be said to be cheap relative to one another after adjustment for risk or maturity: one bond can have a risk premium deemed “too large” compared with the other. Agnostic, because I am not interested in whether a spread between two bonds is justified from a theoretical point of view or not: I am simply interested in the positive fact that dealers take some spreads into account or not.

1.3 Data

1.3.1 Background: broker-dealers, and the US corporate bond market

Broker-dealers. Broker-dealers are institutions that trade a lot in financial markets: they include banks' activities in financial market, such as Barclays Capital, Citigroup Global Markets, Goldman Sachs & Co. or JP Morgan Securities; in the past decade, other large players have emerged in the asset management industry, such as BlackRock, Citadel.[10],[11] More precisely, a *dealer* is a person or company "engaged in the business of buying and selling securities for his own account" (Section 3(a)(5)(A) of the Securities Exchange Act of 1934) as a regular business.[12] A *broker* is "any person engaged in the business of effecting transactions in securities for the account of others". A *broker-dealer* is thus a person or company that acts as a broker and/or as a dealer. In this paper I am interested only in the dealer activity of broker-dealers.

Dealers are often viewed as market makers. But the SEC definition is broader, as the following are typical examples of dealer activities:

- Market making: "a person who holds himself out as being willing to buy and sell a particular security on a continuous basis;"
- Proprietary trading or arbitrage: "a person who runs a matched book of repurchase agreements" (in appendix 1.12 I describe how long-short positions are implemented through repurchased agreements).
- Securitization: "a person who issues or originates securities that he also buys and sells".

A central point of this paper is to show that broker-dealers follow strategies consistent with proprietary trading.

Broker-dealers have to register with the SEC, and have to join a "Self-Regulatory Organization" (SRO) such as the FINRA and national securities exchanges: a SRO is a professional association that assist the SEC in regulating the activities of broker-dealers. Becoming a member of FINRA is mandatory for dealers who trade outside exchanges, such as in the US corporate bond market.[13]

The US corporate bond market The US corporate bond market is over-the-counter: investors do not trade through exchanges, but with individual broker-dealers. It is thus unlike equity markets where trading occurs through limit-order

[10] Through the company BlackRock Execution Services, BlackRock Investments LLC, Citadel Securities respectively

[11] A much more complete list of broker-dealers can be found on FINRA's website: https://www.finra.org/about/firms-we-regulate.

[12] Most information in this paragraph comes from Security and Exchange Commission's website: https://www.sec.gov/reportspubs/investor-publications/divisionsmarketregbdguidehtm.html

[13] https://www.sec.gov/reportspubs/investor-publications/divisionsmarketregbdguidehtm.html#III.

book markets. Most often, customers use telephone and messaging systems to request quotes from dealers, or dealers may contact investors to make them offers. The content of the conversations between dealers and their customers is not available in the US corporate bond market: in particular, quotes offered by dealers, which may be customer-specific, are not disclosed.

Some platforms have emerged to allow investors to request quotes from several dealers at once (Hendershott and Madhavan 2015), but this does not change the fact that investors can choose to trade with a dealer A and not with another dealer B even if dealer B would be willing to trade.

In July 2002, dealers were requested to report all their transactions in quasi real time in the TRACE system, which is managed by FINRA. These reports would be instantaneously released to market participants to give them *ex post* market transparency.

Broker-dealer reporting of transactions. FINRA requires all its members, that is all broker-dealers who trade in the US corporate bond market, to report all their transactions, both as broker and as dealer, in the corporate bond market through TRACE system.

Therefore all transactions in the US corporate bond market that involve a broker-dealer are in TRACE. However, it is not required by broker-dealers to enter the type of activity the transaction is involved in such as market making or proprietary trading, or the party who initiated the trade.

1.3.2 Data and sample selection

Data I use US corporate bonds transaction data from FINRA's enhanced TRACE engine, which I retrieved through WRDS. The sample runs from July 1st, 2002 to December 31st, 2014. Each transaction report in my version of the dataset contains a bond identifier (CUSIP), the date and time of the transaction, the transaction price, the transaction size in terms of par value traded, whether the reporting dealer was a buyer or a seller, and whether the trading counterparty was another dealer or a customer. I clean the data with usual procedures described in appendix 1.9.

Daily Treasury yield curves come from Gurkaynak, Sack, and Wright (2007) and updates by the Fed. Other financial indicators (LIBOR, TYVIX, ...) come from various usual providers.

Sample selection I select customer-to-dealer transactions for dollar-denominated corporate bonds with characteristics available in Mergent FISD database, with issuer common stock to match with using WRDS CRSP-TRACE linking suite. I drop bonds with principal value different from $1000 as these are usually non-standard, whose payoff depend on an index (with "-linked" in their name), and with issue size less than $10 million, as all these bonds are likely to be very illiquid. I keep bonds with embedded options (call, put).

I focus on most actively traded bonds: I keep bonds for which there are customer-to-dealer trades at least 75% of its relevant business days, *i.e.* between first trade and last trade, like Bao, Pan, and Wang (2011). I drop transactions for bonds with residual maturities less than one year, as the trading patterns are special. I drop trades that occur until 7 days after the bond's offering date, as these are likely related to the primary market. I drop bonds that have less than 50 observations, to exclude bonds that are traded only a few days and then disappear.

Finally, I keep observations for investment-grade bonds. Ratings are by S&P, Fitch and Moody's, accessed through Mergent. When ratings from several agencies are available, I retain the worst nonmissing one. If a bond is not rated, then I consider it with a worse rating than any other rating.

This procedures leaves 3080 unique bonds corresponding to 546 issuers as identified by Mergent FISD.

1.3.3 Variables definitions

I study daily order flow in parallel with daily price changes. For bond i and day t, I retain the price $p_{i,t}$ of the last transaction. I also keep the size $q_{i,t}$ of this transaction, with the convention that $q_{i,t} > 0$ if it is a customer buy, $q_{i,t} < 0$ if it is a customer sale.

Order flow I define customer order flow for bond i on day t as the sum of the sizes of customer large buys minus the sum of the sizes of customer large sells. Large transactions are those for which there is at least \$100,000 of par value traded at once: these trades are generally considered as of institutional size and comprise 97% of trading volume in my sample. Formally, denoting $q^1_{i,t}, q^2_{i,t}, ..., q^n_{i,t}, ...$ the 1st, 2nd, ..., nth large transaction in bond i and day t, the order flow is

$$OF_{i,t} = \sum_n q^n_{i,t} \ \mathbb{1}_{|q^n_{i,t}|>\$100,000}$$

It includes the last transaction of day t if this last transaction is large. As shown in the summary statistics subsection 1.3.4, the distribution of order flow has fat tails. To avoid large observations driving the regression estimates, I use the following:

$$\widetilde{OF}_{i,t} = \text{sign}\left(OF_{i,t}\right) \times \log_{10} |OF_{i,t}|$$

which allow to reduce the tails of the order flow distribution while keeping track of the sign of order flow.

I will also need to distinguish between One-way and Partial Roundtrip order flow, and sometimes with a further split between bond maturity being more or less than a cutoff. Thus I define

$$\widetilde{OF}^{OneWay}_{i,t} = \begin{cases} \widetilde{OF}_{i,t} & \text{if One-Way order flow} \\ 0 & \text{otherwise.} \end{cases}$$

and I define $\widetilde{OF}_{i,t}^{Roundtrip}$ for Partial Roundtrip order flow in a similar way. I also define in an analogous way

$$\widetilde{OF}_{i,t}^{OneWay,M\leq 10y} = \begin{cases} \widetilde{OF}_{i,t} & \text{if One-Way order flow and maturity } \leq 10y \\ 0 & \text{otherwise.} \end{cases}$$

and similarly for One-way order flow and bond maturity above 10 years, and Partial Roundtrip with bond maturities above and below 10 years.

Bond spreads. I compute bond spreads in a way similar to Gilchrist and Zakrajsek (2012): for each bond i, I compute a risk-free price as the sum of its theoretical cash-flows (coupons + principal at maturity, maturity being the theoretical maturity for callable bonds) each discounted by the fitted Treasury yield curve of Gurkaynak et al. (2007), and adjusted for accrued interest. The spread is the log of the ratio of the risk free price to the observed price.

Subperiods In some sections of the paper I test the predictions over several subperiods, defined as follows:

1. *Opaque*, from TRACE inception (first observation on July 1st, 2002) to February 7th, 2005. During this period, TRACE data were released for a limited number of bonds only, so that post-trade transparency was limited for other bonds for which data were not released to market participants. On February 8th all transactions were disclosed. One expects a different dealer behavior during this period.

2. *Pre-Crisis*, from February 8th, 2005 to June 30th, 2007.

3. *Crisis*, from July 1st, 2007 to April 30, 2009. The crisis dates (June 30th, 2007 to April 30, 2009) are borrowed from Bessembinder et al. (2018). Making the crisis start in July 2007 is consistent with the increase in bond spread and Primary Dealers' short Treasury position cut in July 2007.

4. *Post-Crisis*, from May 1st, 2009 to July 20th, 2010. This period ends the day before the Dodd-Frank Act was passed. The Dodd-Frank Act in particular contained the Volcker rule, although the implementation details were not written at this time. However, Bessembinder et al. (2018) show large investment banks announced they shut down their proprietary trading desks, suggesting anticipatory effects of the rule.

5. *Dodd-Frank*: from July 21st 2010 to the end of the sample, corresponding to the period when the Dodd-Frank Act was voted.

	Mean	Std. Dev.	p5	p10	p25	p50	p75	p90	p95
Issue size ($mn)	1,026	793	268	350	500	800	1,250	2,000	2,500
Initial maturity	10.6	8.5	3	5	5	10	10	30	30
Maturity $\leq$ 10y	78.9%	-	-	-	-	-	-	-	-
Callable	57.8%	-	-	-	-	-	-	-	-

Table 1.1: Distribution of bond characteristics across the 3,080 bonds in the sample: issue size in millions of dollats, maturity at issuance, and for the last two lines the fraction of bonds with initial maturity less than 10 years, and of bonds with an embedded call giving the bond issuer the opportunity to redeem his bond before maturity.

	Nr observations	% Callable	% One-way
With large trades	1,464,701	52.1	50.4
All	2,220,248	53.6	33.0

Table 1.2: Percentage of observations with callable bonds; percentage of observations with One-way order flow (only large customer buys or only large customer sells), for the subsample with large trades (1st line) and for all observations (2nd line).

1.3.4 Summary statistics

Whole sample

Bond characteristics. Table 1.1 provides a few summary statistics for the sample of bonds. The first line indicates an average issue size of one billion dollars, above the median which stands at 800 million: issues are in general large. At issuance, bonds have an average maturity of 10.6 years, with the median and the 75th percentile being at 10 years. As suggested by the table, most bonds have initial maturity 10 years (31.7%), 5 years (27.0%) or 30 years (9.4%), and 78.9% of bonds have initial maturity less than or equal to 10 years. Finally, more than half of the bonds in the sample have an embedded call as shown by the last line.

Transaction-level. Figure 1.3 gives the share of total volume that transactions of given size represent over the whole sample. For instance transactions between $1 million and $5 million represent close to 40% of total trading volume in my sample. The distribution does not change much across my subperiods. Large transactions, with par value traded above $100,000, represent 97% of total trading volume.

Daily observations. This paragraph gives summary statistics about bond $\times$ day observations. Table 1.2 shows that out of the 2.2 million observations, 1.5 million have large transactions. Slightly more than half observations are with callable bonds, in line with the proportion of callable bonds shown in table 1.1.

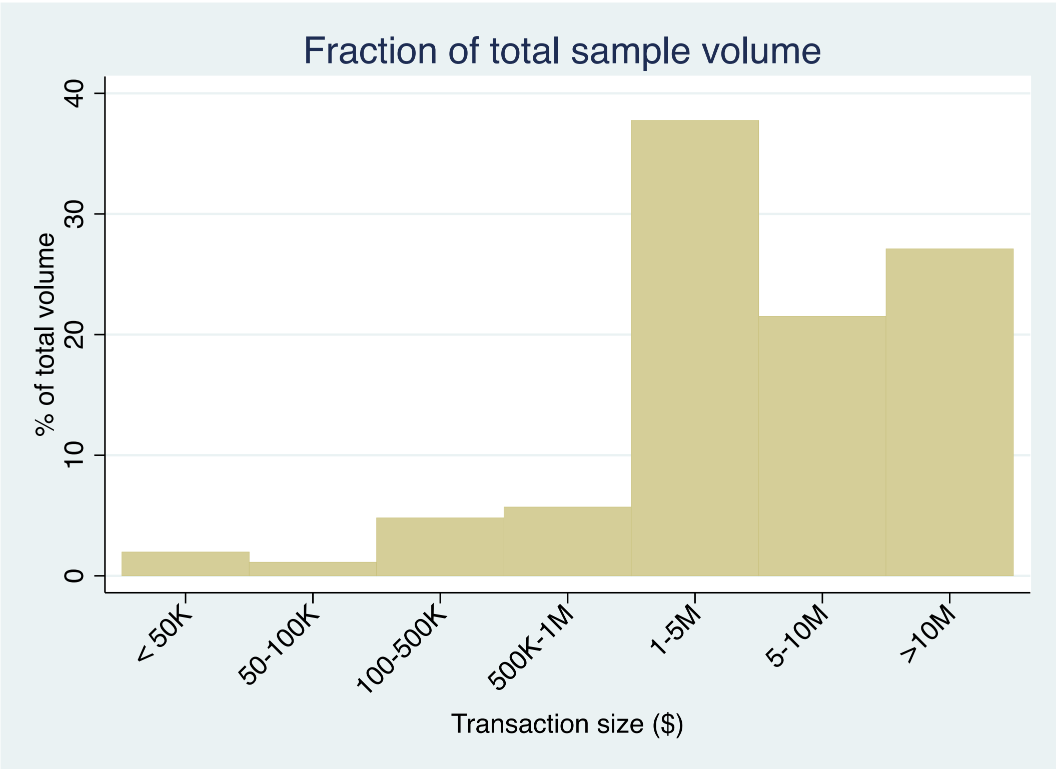

Figure 1.3: Share of total trading volume of transactions in various size buckets.

	Mean	Std. dev.	p5	p10	p25	p50	p75	p90	p95
$OF_{i,t}$ (\$ million)	0	11.2	-8.7	-4.2	-.7	0	.9	4	7.9
$OF_{i,t}$ if One-way	.1	7.1	-5.7	-2.9	-.6	.2	.8	3	5.8
$\Delta \log p_{i,t} \times 100$	0	1.85	-2.36	-1.51	-.53	0	.54	1.53	2.38
Years to maturity	7.8	8.1	1.4	1.8	2.9	5	8.5	23.7	27.7
Age (years)	3.3	2.7	.4	.7	1.4	2.6	4.5	6.9	8.2
# trades (all sizes)	7.5	13.4	1	1	2	4	8	15	22
# trades (large)	1.7	2.6	0	0	0	1	2	4	6

Table 1.3: Sample distribution of order flow $OF_{i,t}$, order flow conditional on it being One-way, of residual maturity and age *conditional on having large transactions* ($>$ \$100,000) in bond i on day t. The distribution of log price changes $\Delta \log p_{i,t}$ and of the number of trades are *unconditional*. Order flow is positive when customers are net buyers.

Table 1.3 shows the distribution of several variables. The distribution of order flow $OF_{i,t}$ is symmetric, with a mean equal to the median at 0. The distribution has fat tails, with more than 20% of the observations having absolute value above \$4 million: this motivates the use of the signed logarithm of order flow $\widetilde{OF}_{i,t}$ in the regressions, to compress the distribution. Conditional on order flow being one-way, order flow is slightly skewed towards customer buys with positive mean and median. Quantiles of one-way order flow are slightly smaller in absolute value than for the unconditional distribution, but the order of magnitude remains the same.

The distribution of log price changes if also symmetric around zero. Log price changes are multiplied by 100, so that they are expressed in percentage points. The volatility of these daily price changes is suprisingly large, but mitigated in the longer run by reversals.

The mean of residual maturity when bonds are traded is lower than the mean bond maturity at issuance, which is partly mechanical due to the fact that residual maturity decreases through time. Interestingly, the distribution of bond ages suggests that most observations are for relatively young bonds, with a mean and median close to 3 years, while the 95th percentile is at 8.2 years, suggesting bond older than 10 years are seldom traded.

Finally, the last two lines show the distribution of the number of transactions per bond $\times$ day observation. The mean of 7.5 is low compared to equity markets for instance; the lower median at 4 transactions suggests the mean is driven by a few observations with many transactions. Over all observations, the number of large transactions (more than \$100,000 traded) is even lower, with a median at 1 transaction and a mean at 1.7 transaction.

Subperiod	One-Way	One-Way $M \leq 10$y	OneWay $M > 10$y
Opaque	47	38	9
Pre-Crisis	56	44	11
Crisis	49	38	11
Post-Crisis	44	34	10
Dodd-Frank	51	39	13

Table 1.4: Percentage of One-way order flow and further breakdown by initial maturity M, conditional on observations having large transactions.

Evolution through time

Table 1.4 plots the percentage of One-way order flow across observations by subperiod, with a further split by initial maturity. I use the further split by maturity in the order flow regressions in section 1.5, to better identify proprietary trading.

There is no striking change from before to after the crisis: the proportion of One-way order flow is slightly higher during the Pre-Crisis and Dodd-Frank periods, as is the proportion for One-way order flow and maturity below 10 years.

1.4 Price changes and order flow

In this section I test hypothesis 1 on the correlation between price changes and customer order flow. To do this I regress daily log price changes on customer order flow, controlling for other plausible determinants of price changes. These regressions allow to determine whether broker-dealers act only as market makers, or if they also follow other strategies. I also discuss potential endogeneity concerns in subsection 1.4.3

1.4.1 Specifications

Here I present three different specifications: one with order flow with no distinction between One-way and Partial Rountrip, one with the distinction, and one with the distinction and an interaction with a dummy for the post-Lehman crisis. I present estimation results in the next subsection.

Baseline

Here I present a benchmark specification where I do not distinguish between One-way and Roundtrip order flow. A first limit to measuring price impact is that pre-trade quotes are not available. These would allow to separate actual quote movements, which are what I am interested in, from one-shot order processing costs that are charged without impact on subsequent prices. Inspired by Foucault, Pagano, and Roell (2014) and as in Rapp (2016) I circumvent the issue raised by the absence of pre-trade quotes issue by aggregating transactions at daily level, so that the price impact component of the spread can be isolated from other order processing costs and rents.

Thus I compute the log price difference between the last transaction of day t and the last transaction of day $t-1$, and regress it on large customer order flow, controlling for small order flow, measures of order processing costs and various controls. Thus I estimate the following equation:

$$\Delta \log p_{i,t} = \alpha + \beta\ \widetilde{OF}_{i,t} + \gamma' X^{(p)}_{i,t} + \epsilon_{0,i,t} \tag{1.4.1}$$

Market making theories predict $\beta > 0$. As reviewed in subsection 1.2.1, under dealer inventory costs and/or customer private information about bond value, one expects customer sales (resp. purchases) to be associated with price decreases (resp. increases): a customer sell either signals bad news about the asset value, or imposes more risk on the dealer's balance sheet - both of which leading the dealer to trade at a lower price.

$X^{(p)}_{i,t}$ is a vector of controls that contains the following variables. First, I control for order processing costs: each price $p_{i,t-1}, p_{i,t}$ is the price of an actual transaction that can be a customer buy at the ask price, or a customer sell at the bid price lower than the ask price. Thus I include the directions $d_{i,t-1}, d_{i,t}$ of the the corresponding transactions, equal to $+1$ for customer buys and -1 for customer sells. The empirical microstructure literature in the US corporate bond market also suggest that order processing costs may decrease with the size of the order because larger customers get better terms: to the direction of the last transaction I add their signed log sizes, $\text{sign}(q_{i,t-1}) \log_{10} |q_{i,t-1}|$ and $\text{sign}(q_{i,t}) \log_{10} |q_{i,t}|$, with the log again to prevent the largest observations to drive the results.

I also control for the bond issuer's stock return, for changes in the 10 years US Treasury yield, changes in the 3 months LIBOR, changes in rating, TYVIX (an implied volatility index for 10 years Treasury futures) and changes in TYVIX. All changes are daily, using closing prices.

Log price changes are $\Delta \log p_{i,t} = \log(p_{i,t}/p_{i,t-1})$. Therefore $\Delta \log p_{i,t} = 0.01$ means the price has changed by 1 percent. To express price changes in percentage points, in the regressions I replace $\Delta \log p_{i,t}$ by $\Delta \log p_{i,t} \times 100$ as in table 1.3.

Both trades and price changes may be driven by information related to the bond issuer. Therefore I cluster standard errors by bond issuer. In the appendix I also compute standard errors clustered by bond issuer, by maturity and by calendar month.

One-way *vs.* Roundtrip

Now I distinguish between One-way and Partial Roundtrip order flow: I replace the order flow measure by a one-way order flow measure and a partial roundtrip order flow measure:

$$\Delta \log p_{i,t} = \alpha + \beta_1\ \widetilde{OF}^{OneWay}_{i,t} + \beta_2\ \widetilde{OF}^{Roundtrip}_{i,t} + \gamma X^{(p)}_{1,i,t} + \epsilon_{1,i,t} \tag{1.4.2}$$

$\widetilde{OF}_{i,t}^{OneWay}$ equals $\widetilde{OF}_{i,t}$ if order flow in bond i was One-way on day t; it equals zero otherwise. Similarly $\widetilde{OF}_{i,t}^{Roundtrip}$ equals $\widetilde{OF}_{i,t}$ if order flow in bond i on day t was a partial roundtrip, and zero otherwise. Hypothesis 1 formally reads

$$\beta_1 < 0 \quad \text{and} \quad \beta_2 > 0.$$

Splitting order flow in the previous way is equivalent to putting an interaction of order flow with the dummy for order flow being One-way: I include this dummy in the vector of controls $X_{1,i,t}^{(p)}$, which otherwise contains the same controls as in equation (1.4.1). Again $\log p_{i,t}$ is multiplied by 100 to express it in percentage points. I cluster standard errors by bond issuer similarly to regression (1.4.1).

One-way *vs.* Roundtrip outside post-Lehman crisis

The financial crisis period has been special regarding prop trading: after Lehman Brothers' failure on September 15th, 2008, markets were reportedly highly illiquid, which may have induced dealers to stop proprietary trading. Therefore One-way order flow during this period may reflect only market making; in addition, market making costs may have been very high: this would tend to bias the coefficient β_1 up in equation 1.4.2.

Therefore I interact both measures of order flow with the dummy $Lehman_t$ that equals 1 between September 15th, 2008 and April 30th, 2009, *i.e.* for 7.5 months out of the 12.5 years of my sample. The equation becomes:

$$\begin{aligned}\Delta \log p_{i,t} = \alpha &+ \beta_1 \widetilde{OF}_{i,t}^{OneWay} + \beta_2 \widetilde{OF}_{i,t}^{Roundtrip} \\ &+ \beta_3 \widetilde{OF}_{i,t}^{OneWay} \times Lehman_t + \beta_4 \widetilde{OF}_{i,t}^{Roundtrip} \times Lehman_t \\ &+ \beta_0 Lehman_t + \gamma X_{2,i,t}^{(p)} + \epsilon_{2,i,t} \end{aligned} \tag{1.4.3}$$

The vector of controls $X_{2,i,t}^{(p)}$ includes the dummy for one-way order flow, the $Lehman_t$ dummy and the interaction between the two. Otherwise it contains the same controls as in specifications 1.4.1 and 1.4.2.

1.4.2 Results

Table 1.5 presents the estimation results for equations 1.4.1, 1.4.2 and 1.4.2.

The first column reports the estimation result for the baseline regression 1.4.1. The coefficient is positive and highly significant, consistently with market making: running the price impact regression without distinction thus legitimates the idea that broker-dealers are pure market makers. The estimates imply that customer net buys by 1 million are on average associated with a price increase by $\log_{10}(1,000,000) \times .0033 = .02$ percentage points, that is, 2 basis points. It may appear small, but this is not my point: I distinguish between theories that imply positive or negative coefficients, with potentially broader, systemic implications.

Table 1.5: Regression of daily log price changes on customer order flow and controls. $\widetilde{OF}_{i,t}$ is the sign of order flow times the logarithm of the absolute value of order flow. Order flow is the sum of customer large buys minus the sum of customer large sells. A customer buy or sell is large its size is above $100,000. $\widetilde{OF}_{i,t}^{OneWay}$ equals $\widetilde{OF}_{i,t}$ if order flow in bond i on day t is one-way (only customer buys or only customer sells), and zero otherwise. $\widetilde{OF}_{i,t}^{Roundtrip}$ equals $\widetilde{OF}_{i,t}$ if order flow is not One-way, *i.e.* (partial) roundtrip, and zero otherwise. Controls are issuer stock return, changes in 10 years Treasury yield, changes in 3-months LIBOR, TYVIX, an implied volatility index for Treasury futures, and changes in TYVIX. In the second and third column, a dummy $OneWay_{i,t}$for one-way order flow is included. $Lehman_t$ is a dummy that equals 1 if t is between September 15th, 2008 and April 30th, 2009. In the third column, the interaction $OneWay_{i,t} \times Lehman_t$ is included.

	Baseline	One-way *vs.* Roundtrip	One-way *vs.* Roundtrip Lehman interaction
$\widetilde{OF}_{i,t}$	0.0033***		
	(10.68)		
$\widetilde{OF}_{i,t}^{OneWay}$		0.0008*	**-0.0025*** **
		(2.05)	**(-6.12)**
$\widetilde{OF}_{i,t}^{Roundtrip}$		0.0056***	0.0040***
		(13.67)	(11.65)
Interaction with $Lehman_t$	N	N	Y
Constant and controls	Y	Y	Y
R^2	0.24	0.24	0.24
N	2,220,248	2,220,248	2,220,248

Standard errors clustered by bond issuer
* (p<0.05), ** (p<0.01), *** (p<0.001)

The second column reports the results for One-way and Partial Roundtrip order flow separately. The coefficient for One-way order flow is 7 times smaller that the coefficient for partial Roundtrip order flow, and the difference is statistically significant (not shown); it is however positive and just significant at 5% level. The difference between the two coefficients is already striking.

The third column reports the estimation results for the main effects of One-way and Roundtrip order flow, *i.e.* for these measures outside the crisis post-Lehman. The coefficient for one-way order flow is now negative and highly significant, while the other is significantly positive: this validates hypothesis 1.

1.4.3 Endogeneity concerns and robustness checks

One may be concerned about endogeneity when interpreting correlations in price impact regressions as reflecting one class of theories or the other. Endogeneity is certainly there, but I claim that 1) it does not raise significant concerns for my purpose and 2) I already control for plausible sources of endogeneity.

Direction of causality

A first concern is about the direction of causality: one does not know whether prices decrease because dealers sell (i.e. customers sell), or whether customers buy because they see a low price. One would view the second case as a sign dealer proprietary trading. One may view the first direction as consistent with market making behavior, with dealers decreasing their quotes to liquidate inventory they have just purchased.

I view the market making interpretation as misled. It is true that a market maker willing to re-sell inventory he has just bought is quoting low price to induce customer purchases. However this does not imply that the market maker is systematically willing to *decrease* his quotes to induce customer purchases: indeed such statement implies the market maker would systematically buy at a high price and sell at a low price, and thus make a systematic market making loss that would drive him out of business.

In fact, a profitable pure market maker anticipates that the re-sell price will have to be low as well when he buys, thus buys at a lower price. This predicts that customer buys are associated with price increases.

Therefore observing a correlation between customer buys (sales) and price decrease (increase) does not require any stand on the direction of causality to conclude that it corresponds to proprietary trading. However, one may worry that common drivers of customer net trades and price changes drive the results. I review these concerns in what follows.

Common drivers for price changes and customer trades

One may also be concerned that there are common drivers for price changes and customer trades: for instance public bad news may induce both prices to decrease and customers to sell. In this case the correlation does not seem to reflect market making effects such as adverse selection and dealer inventory costs.

I see the reverse correlation as less problematic for my interpretation. If bad news induce both price drops and dealer sells or customer buys, it means that dealers are more willing to sell than their customers, who thus provide liquidity to dealers.

In any case I control for plausible common drivers of order flow and price changes, which I review now.

Credit and risk-free rate risks. I address this endogeneity concern by controlling for proxies for public information. Relevant information should be primarily about issuer's credit quality and interest rates. Regarding credit risk, I add control for stock return: informed traders should also trade in the stock market, as bond credit risk and stocks valuation are both about the issuer's asset side valuation. Regarding interest rate movements, I control for changes in the 10 years Treasury yield, and for the 3-months LIBOR. The LIBOR captures both short-term interest rate movements and changes in credit risk concerns in the banking sector; including changes in the 3-months Treasury bill instead does not change the results.

Cheapness measures One may also worry that cheapness measures forecast both dealer purchases and price increases. I address this concern in appendix 1.10.1, in which I show that estimates of order flow are unaffected.

Predicted order flow I also assess in appendix 1.10.2 whether other predictors of order flow are driving the results: previous price changes, lags of order flow and lagged stock return and interest rate changes. To do this I compute a predicted and an unexpected component of order flow, separately for One-way and Partial Rountrip order flow, as described in the appendix. I am interested in the unexpected component. Again the estimates are unchanged.

Robustness to more conservative standard errors

My results are robust to additional layers of clustering: one may be concerned that order flow is correlated with price changes within a maturity bucket, or within a calendar month. In appendix 1.10.3 I re-run the price impact regression by computing standard errors with various multiway clustering, following the methodology by Cameron et al. (2011), and including the cheapness measures as in subsection 1.4.3 to check full robustness. The estimates remain significant at 5% level even with three-way clustering by bond issuer, maturity and calendar month.

1.4.4 Proprietary trading: refinements

Table 1.5 shows that outside the crisis, One-way order flow corresponds on average to proprietary trading, while it does not during the crisis. One may wonder whether such average effect hide some heterogeneity: sometimes dealers buy and do not sell and conversely because they are willing to hold the position, sometimes it may be because they were not able to find a counterparty sufficiently quickly.

Here I re-run regressions of price changes over One-way and Partial Roundtrip order flow, by further distinguishing by three different criteria: bond age, bond maturity at issuance (initial maturity) and bond residual maturity at the time of trade. I split the sample with according to whether the criterion is below or above 10 years. The generic specification for criterion $Z = Age$, $InitialMaturity$, $ResidualMaturity$ is thus

$$\begin{aligned}\Delta \log p_{i,t} = \alpha &+ \beta_1 \widetilde{OF}_{i,t}^{OneWay,Z\leq 10y} + +\beta_2 \widetilde{OF}_{i,t}^{OneWay,Z>10y} \\ &+ \beta_3 \widetilde{OF}_{i,t}^{Roundtrip,Z\leq 10y} + \beta_4 \widetilde{OF}_{i,t}^{Roundtrip,Z>10y} \\ &+ \sum_k LehmanInteractionTerms_{k,i,t} + Lehman_t + \gamma X^{(p)}_{2,i,t} + \epsilon_{2,i,t}\end{aligned} \tag{1.4.4}$$

where the $LehmanInteractionTerms_{k,i,t}$ are interaction terms of each of the four measures of order flow with the dummy $Lehman_t$.

Table 1.6 shows the results, which are clear: proprietary trading is concentrated on younger bonds (less than 10 years old), on bonds with initial maturity 10 years, and on bonds with residual maturity 10 years. Otherwise the coefficients are consistent with market making.

Overall it seems that older bonds, or bonds with longer maturity (initial or residual), one-way order flow corresponds to market making and dealers' inability to revert inventory quickly towards the target. These bonds are likely less liquid, so the results are also in line with the results by Goldstein and Hotchkiss (2020) who find on their sample of illiquid bonds that 42% of customer trades are *not* reversed within one day.

1.4.5 Evolution through time

I ran the previous regressions over the whole sample, covering plausibly very different environments from before the crisis to after. Here I add interaction of one-way and partial roundtrip order flows with dummies for the subperiods defined in section 1.3.

Formally, I estimate the following, where $Period_t^k$ is a dummy for one of the above periods *Opaque*, *Bear Stearns to Lehman*, *Lehman to end of crisis*, *Post-Crisis*, *Post Dodd-Frank*, the base level being the Pre-Crisis period (the base level choice does

Table 1.6: Regression of price changes on One-way and Partial Roundtrip order flow, further distinguishing order flow on whether a criterion $Z_{i,t}$ is below or above 10 years. Criteria Z are bond age (time since issuance), bond maturity at issuance (initial maturity), bond residual maturity at the time of trade. Each column reports the estimation with a different criterion. Controls are issuer stock return, changes in 10 years Treasury yield, changes in 3-months LIBOR, TYVIX, an implied volatility index for Treasury futures, and changes in TYVIX. In the second and third column, a dummy $OneWay_{i,t}$for one-way order flow is included. $Lehman_t$ is a dummy that equals 1 if t is between September 15th, 2008 and April 30th, 2009. Dummy main effects are included, as well as relevant interactions terms between dummies.

	Age	Initial Maturity	Residual Maturity
$\widetilde{OF}_{i,t}^{OneWay,Z\leq 10y}$	-0.0029***	-0.0050***	-0.0045***
	(-6.72)	(-9.83)	(-10.17)
$\widetilde{OF}_{i,t}^{OneWay,Z>10y}$	0.0109***	0.0066***	0.0093***
	(5.32)	(6.87)	(7.50)
$\widetilde{OF}_{i,t}^{Roundtrip,Z\leq 10y}$	0.0039***	0.0029***	0.0031***
	(11.21)	(8.94)	(9.98)
$\widetilde{OF}_{i,t}^{Roundtrip,Z>10y}$	0.0093**	0.0085***	0.0090***
	(3.28)	(10.97)	(9.13)
Interaction with $Lehman_t$	Y	Y	Y
Constant and controls	Y	Y	Y
R^2	0.24	0.24	0.24
N	2,220,248	2,220,248	2,220,248

Standard errors clustered by bond issuer

* (p<0.05), ** (p<0.01), *** (p<0.001)

not affect the results):

$$\begin{aligned}\Delta \log p_{i,t} = \alpha &+ \beta_1 \widetilde{OF}_{i,t}^{OneWay} + \beta_2 \widetilde{OF}_{i,t}^{Roundtrip} \\ &+ \sum_k \left(\beta_{k,OW} \widetilde{OF}_{i,t}^{OneWay} \times Period_t^k + \beta_{k,R} \widetilde{OF}_{i,t}^{Roundtrip} \times Period_t^k \right) \\ &+ \sum_k \beta_{k,0} Period_t^k + \gamma X_{2,i,t}^{(p)} + \epsilon_{2,i,t} \end{aligned} \tag{1.4.5}$$

I plot the total effects of this regression for each period, for One-way and Partial Roundtrip coefficients separately, in figure 1.4. I scale the coefficients so that they correspond to the price change in basis points associated with net customer purchases by \$1 million. The original coefficients are in table 1.13 in the appendix.

Proprietary trading, as shown by negative coefficients, is present before the Crisis in the Opaque and Pre-Crisis periods, and also after the crisis in the Dodd-Frank period. The coefficient on One-way order flow is positive during the Crisis and strikingly higher than the coefficient on Partial Roundtrip order flow. This likely comes from the fact that one-way order flow corresponds to positions dealers are not willing to hold, but cannot start to offload within a day: as they probably expect to hold these positions for longer, they entail higher price impact, for instance associated with higher inventory holding costs.

It is also striking that coefficients before the crisis are much strongly more negative than after the Dodd-Frank Act was passed: this suggests that post-crisis regulation, the Volcker rule in particular, decreased the strength of proprietary trading.

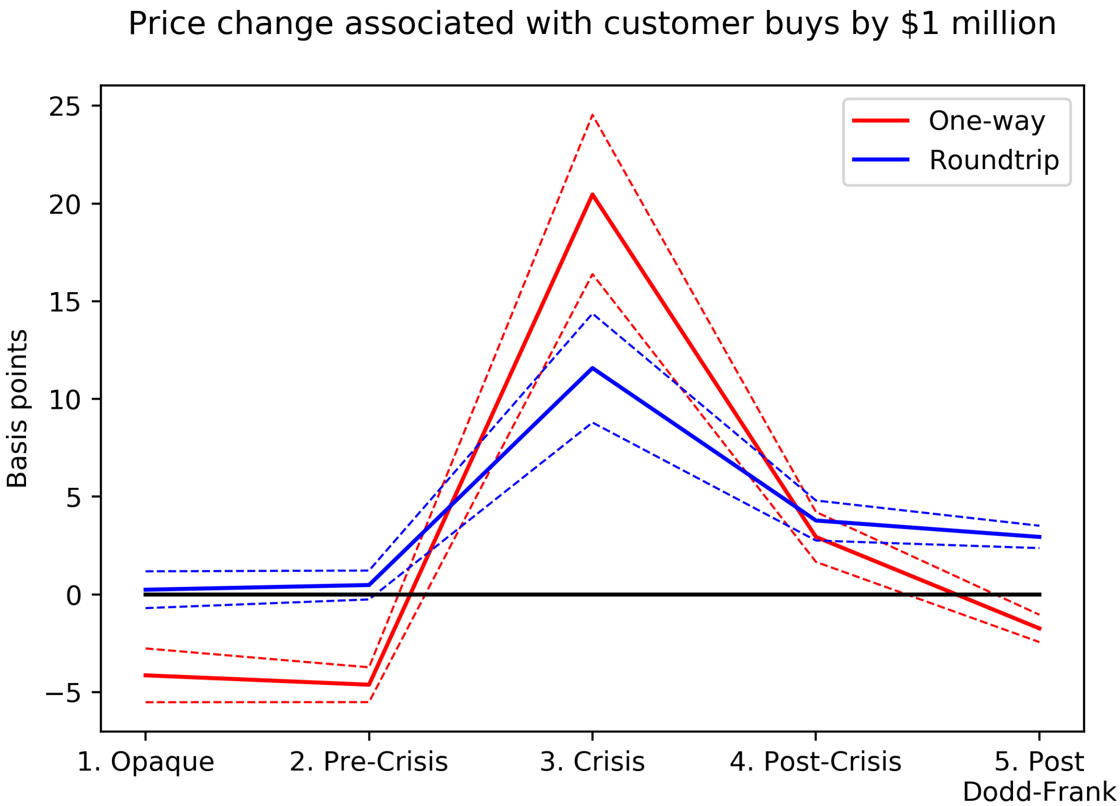

Figure 1.4: Regression-implied log price changes associated with customer net purchases by $1 million across sample subperiods, for One-way (red) and Roundtrip(blue) order flow. Dashed lines are the 95% confidence intervals for each coefficient. The Opaque period when not all TRACE transactions were disclosed to market participants starts on July 1st, 2002 and stops on February 7th, 2005. The Pre-Crisis period starts on February 8th, 2005 and stops on June 30th, 2007. The Xrisis period goes from July 1st, 2007 to April 30th, 2009 (a conventional date used in other papers). The Post-Crisis period goes from May 1st, 2009 to July 20, 2010. The Post Dodd-Frank goes from July 21st (Dodd-Frank Act voted) to the end of the sample.

1.5 Order flow and lagged cheapness

In this section I test hypothesis 2. If dealers tend to buy (sell) bonds that were cheap (expensive) the day before, it reveals prop trading activities by dealers.[14] I first focus on a simple of bond cheapness. I introduce individual corporate-to-Treasury spreads, *i.e.* the spread between the bond's yield and the yield of a fictitious bond with the same contractual cash flows discounted with the Treasury yield curve. Then I compare this spread to the median spread of all bonds with the same maturity, credit risk and callability features: this is a proxy for an idiosyncratic component of the bond's individual spread. I find that broker-dealers indeed tend to purchase (sell) bonds with maturity below (above) 10 years.

Then I add three spread components related to credit risk and callability given residual maturity, to maturity and to differences between corporate bonds as a whole and Treasury bonds. The four measures are a model-free decomposition of the spread: it only looks at relative price differences between various subsets of bonds. The analysis shows that broker-dealers arbitrage corporate bonds within maturities (thus across ratings/callability) and as an asset class with respect to Treasuries.

There are potentially many proprietary trading opportunities in the corporate bond market and in related markets, and I do not try to be exhaustive: documenting that broker-dealers exploit at least some price differentials is enough for my purposes. Furthermore, I proxy cheapness with various price differentials, that may or may not be justified from a normative viewpoint: I am only interested in the fact that broker-dealers react to these differentials, which likely shows that dealers perceive them as mispricings.

1.5.1 A narrow measure of relative cheapness

The measure

I compute spreads as indicated in section 1.3: they are the log difference between the observed clean bond price and a fictitious "risk-free" bond price. The latter price is the price of a fictitious bond that has the same contractual coupons and principal, but with these cash flows discounted with the Treasury yield curve.

There are two spurious correlations I have to avoid. First, by regressing customer order flow on *contemporaneous* last spread of the day, one may simply capture the price impact of market making trades: customer sales (purchases) would be associated with price decrease (increase) and misleadingly be interpreted as dealers purchasing cheap bonds. To limit this and other potential endogeneity concerns, I regress order flow on *lagged* spread.

Second, the raw spread may capture mechanical effects related to the bid-ask bounce. The problem is as follows. If the last transaction of day t was a customer

[14] Although this reveals nothing about their leverage.

sale, it was executed at a lower price than if it was a customer buy just because of a positive bid-ask spread. I may therefore capture spurious explanatory power of day $t-1$ bond spread: if 1) a customer buy at the end of day $t-1$ is correlated with customer sale at the end of day $t-1$, and 2) a customer buy at the end of day t is correlated with overall customer buys during day t, then the forecasting power of the bond spread simply captures some mechanical features of customer order flow, and not dealer proprietary trading. Averaging over past business days solves the problem.

Thus to address both concerns I consider the average $y_{i,t-1}$ of the last bond spreads of the seven previous business days $t-7$ to $t-1$ in cheapness measures. The choice of 7 days is conservative, while results do not materially change if I do not average at all over business days.

I first consider a simple strategy that is close to arbitrage: bonds with similar time until maturity, with similar credit rating and with or without an embedded call can be viewed as close substitutes, and from this perspective should have very similar prices.[15] Therefore I group bonds by integer part of years to maturity, by credit rating category (AAA/AA, A, BBB) and by callability. The difference between a bond's average spread $y_{i,t}$ and the median in its group $y_{i,t}^{sim}$ is the cheapness measure that captures this quasi arbitrage, and is a proxy for the idiosyncratic component of the bond's spread to Treasuries:

$$Idiosync_{i,t} = y_{i,t} - y_{i,t}^{sim}$$

This measure is positive if bond i is cheap, *i.e.* its spread is higher than the median in the basket of similar bonds.

Baseline specification

Again I distinguish between One-way and Partial Rountrip order flow. To do it I regress order flow on my measure of cheapness interacted with a dummy $Roundtrip_{i,t}$ equal to 1 if order flow is a partial roundtrip on day t for bond i.

$$\begin{aligned}\widetilde{OF}_{i,t} = \alpha + \nu_1\ Idiosync_{i,t-1}^{OneWay} + \nu_2\ Idiosync_{i,t-1}^{Roundtrip} \\ + \beta' X_{1,i,t-1}^{q} + \eta_{i,t} \qquad (1.5.1)\end{aligned}$$

Dealers tend to purchase bonds that are cheap with respect to bonds that are similar to it on One-way days if $\nu_1 < 0$. The same happens on Partial Roundtrip days if $\nu_2 < 0$.

The regression includes controls for 10 lags of log order flow, for 10 lags of past price changes, and the vector $X_{1,i,t-1}^{q}$ contains 3 lags of issuer stock return, lagged

[15] Assets with similar expected payoff and variance may have different prices because of different covariances with the market portfolio. For investment grade bonds, this should come from the credit risk component, and I do not expect the covariance between credit risk to play such a big role with respect to other components - interest rate risk, systematic component of credit spreads, ...

10 years Treasury yield change and lagged changes in 3 months LIBOR. I exclude contemporaneous controls to avoid an endogeneity concern: the cheapness measure may predict both order flow and the contemporaneous controls such as stock return and rate changes. It also contains the dummy $OneWay_{i,t}$ for the order flow on day t in bond i being one-way. I estimate equation 1.5.1 through OLS and cluster standard errors by issuer as for price impact regressions.

Refined specification with age or maturity criteria

I regress the signed logarithm of order flow on derived cheapness measures: for instance $Idiosync_{i,t}^{OneWay,Maturity\leq 10y}$ equals $Idiosync_{i,t}$ if bond i has residual maturity less than 10 years on day t, and day t order flow is One-way in bond i; it equals zero otherwise. Similarly, for Roundtrip order flow and for residual maturities above 10 years. Order flow can also be split according to an age, or an initial maturity criterion.

I thus estimate the following equation for each criterion Z being bond age, bond initial maturity and bond residual maturity:

$$\begin{aligned}\widetilde{OF}_{i,t} = \alpha &+ \nu_1\ Idiosync_{i,t-1}^{OneWay,Z\leq 10y} + \nu_2\ Idiosync_{i,t-1}^{OneWay,Z>10y} \\ &+ \nu_3\ Idiosync_{i,t-1}^{Roundtrip,Z\leq 10y} + \nu_4\ Idiosync_{i,t-1}^{Roundtrip,Z>10y} \\ &+ \beta' X_{Z,i,t-1}^{q} + \eta_{i,t} \end{aligned} \tag{1.5.2}$$

Bond lagged cheapness is associated with dealer purchases for criterion Z less than 10 years and One-way order flow if $\nu_1 < 0$: then a higher spread (cheaper bond) is associated with more customer sales ($\widetilde{OF}_{i,t} < 0$), *i.e.* dealer purchases. A similar reasoning applies to ν_2, ν_3, ν_4. The vector $X_{Z,i,t-1}^{q}$ includes the same controls as as for regression 1.5.1 and adds a dummy for the criterion Z being lower than 10 years, and its interaction term with the dummy $OneWay_{i,t}$ for order flow being One-way on day t.

Results

Table 1.7 shows the results. To save space I hide coefficients for Roundtrip order flow, which are never significant.

The coefficient for $Idiosync_{i,t-1}^{OneWay}$ is not significant in the first column, although it is negative as expected. In the second column, I show the estimates of regression 1.5.2 for the age criterion: the coefficient is negative (weakly) significant for bonds older than 10 years, and negative insignificant for bonds younger than 10 years.

The maturity criteria work better: the coefficient is negative significant for shorter maturity bonds, whether initial (third column) or residual (fourth column). Spreads are computed as log price differences multiplied by 100. The estimate for the initial maturity criterion implies that a one point increase in the $Idiosync_{i,t-1}$ measure is associated with a .0550 decrease in log customer purchases, meaning that customer purchases vary by $10^{-0.0550} - 1 = -12\%$, *i.e.* dealer purchases increase by

Table 1.7: Order flow regressed on lagged measure of cheapness $CheapInSim$, equal to the bond spread to a fictitious Treasury bond with the same cash flows and discounted with the Treasury yield curve of the day, minus the median of these spreads in the basket of bonds with the same credit rating, maturity and callability. Controls include the main effects for the dummy $OneWay_{i,t}$ for one-way order flow, and in the 2nd, 3rd and 4th columns the main effect for the criterion $Crit$ (bond age, initial maturity and residual maturity) being less than 10 years, and its interaction with $OneWay_{i,t}$. Additional controls are 10 lags of order flow, 10 lags of bond price changes, 3 lags of stock return and lagged 10 years Treasury yield changes. To save space only the coefficient for $Idiosync$ for One-way order flow are shown. Coefficients for Roundtrip order flow are never significant.

Z	None	Age	InitMaturity	ResidMaturity
$Idiosync_{i,t-1}^{OneWay}$	-0.0114			
	(-1.54)			
$Idiosync_{i,t-1}^{OneWay,Age\leq 10y}$		-0.0100	**-0.0550*****	**-0.0541*****
		(-1.29)	**(-4.06)**	**(-4.26)**
$Idiosync_{i,t-1}^{OneWay,Age>10y}$		**-0.0209***	0.0095*	0.0121*
		(-2.06)	(2.03)	(2.55)
Roundtrip order flow	Y	Y	Y	Y
Constant and controls	Y	Y	Y	Y
R^2	0.00	0.00	0.00	0.00
N	2,316,162	2,316,162	2,316,162	2,316,162

Standard errors clustered by bond issuer

* (p<0.05), ** (p<0.01), *** (p<0.001)

12%. The estimate is similar for the residual maturity criterion. For longer maturity bonds (by both measures) the estimate is slightly positive significant, suggesting customer proprietary trading.

1.5.2 More general proprietary trading strategies

In this section I investigate additional arbitrage strategies by decomposing the bond spread in four model-free components including that of section 1.5.1. The decomposition brings stronger evidence of dealer proprietary trading and isolates three relevant broker-dealer trading strategies:

- between similar bonds as in subsection 1.5.2
- between bonds of similar maturities, irrespective of their credit rating/maturities
- between corporate bonds and Treasury bonds, which is consistent with figure 1.1.

A model-free spread decomposition

A given bond can enter proprietary trading strategies because of one or several of its characteristics. Broker-dealers could arbitrage between highly similar bonds in terms of maturity, credit rating and embedded options; or between bonds that have similar maturities, irrespective of their credit rating and embedded options; or between baskets of bonds of similar maturities; or between corporate bonds and Treasury bonds as asset classes. To capture possibly perceived trading opportunities along each of these dimensions, for each bond i I decompose its spread at the end of day t (averaged over the past 7 business days up to t) as the sum of four bond cheapness measures:

$$y_{i,t} = \underbrace{(y_{i,t} - y_{i,t}^{sim})}_{\mathbf{Idiosync_{i,t}}} + \underbrace{(y_{i,t}^{sim} - y_{i,t}^{\tau})}_{\mathbf{CreditCall_{i,t}}} + \underbrace{(y_{i,t}^{\tau} - y_{i,t}^{1-7})}_{\mathbf{Maturity_{i,t}}} + \underbrace{y_{i,t}^{10}}_{\mathbf{TreasuryConv_{i,t}}} \tag{1.5.3}$$

where:

- $y_{i,t}^{sim}$ is the median of $y_{i,t}$ in the group of bonds with the same rating category (AAA/AA, A, BBB), the same residual maturity rounded to the year and the same callability (presence or absence of an embedded call) as bond i;
- $y_{i,t}^{\tau}$ is the median of $y_{i,t}$ in the group of bonds with the same residual maturity τ (rounded to the year) as bond i;
- $y_{i,t}^{10}$ is the median spread of bonds with residual maturity less than 10 years if it is the case for bond i, and the median spread for bonds with residual maturity more than 10 years otherwise.

Each term between parenthesis in (1.5.3) is a spread differential that captures one of the strategies described above, in isolation from the others.

To visualize the components of spread from equation 1.5.3, I consider a "spread curve", by analogy with a yield curve for Treasury bonds: as illustrative examples, Figure 1.5 plots bond spreads as a function of residual maturity, for randomly chosen dates. Each component in equation 1.5.3 reflects an aspect of figure 1.5:

- $Idiosync_{i,t}$ - *Scatter thickness within rating category/callability*:[16] a bond is considered cheap (expensive) with respect to similar bonds as assessed through credit risk, maturity, and the presence of embedded call or not[17]); it thus capture a bond idiosyncratic component of its yield;
- $CreditCall_{i,t}$ - *Scatter thickness across ratings/callability*: bonds with similar credit risk maturity and callability are considered overall cheap (expensive) with bonds of similar maturity τ; it thus jointly captures credit risk and callability components of the spread, controlling for maturity;
- $Maturity_{i,t}$ - *Scatter slope*: bonds of a given maturity τ are considered cheap (expensive) with respect to bonds of another maturity τ';
- $TreasConv_{i,t}$ - *Scatter level*: bonds within a broad maturity bucket are considered cheap (expensive) with respect to Treasuries. It captures safe asset demand arbitrage.

I regress order flow on each of the terms in parenthesis in (1.5.3) separately for the three criteria $Z = Age, InitMaturity, ResidMaturity$ below and above 10 years, interacting cheapness measures with the dummy $OneWay_{i,t}$ used in the price impact regressions, and controlling lags of order flow, lags of price changes and other lagged market factor changes:

$$\widetilde{OF}_{i,t} = \alpha + \sum_x \big(\nu_1^x Idiosync_{i,t-1}^x + \nu_2^x CreditCall_{i,t-1}^x + \nu_3^x Maturity_{i,t-1}^x + \nu_4^x TreasConv_{i,t-1}^x\big) + \beta' X_{i,t-1}^q + \eta_{i,t} \tag{1.5.4}$$

with

$$x = (OneWay, Z \leq 10y), (OneWay, Z > 10y), (Roundtrip, Z \leq 10y), (Roundtrip, Z > 10y).$$

For consistency I stick to the previous convention that customer buys correspond to positive order flow, and vice-versa. If dealers are also arbitrageurs, then we expect them to buy cheap bonds and sell expensive bonds to customers.

[16] Callability is not distinguished on Figure 1.5.

[17] Callable corporate bonds, i.e. with an option to the issuer to redeem the bond before maturity, are highly frequent. The option to call the bond *e.g.* to reissue at lower rate should be priced and incorporated into the bond spread.

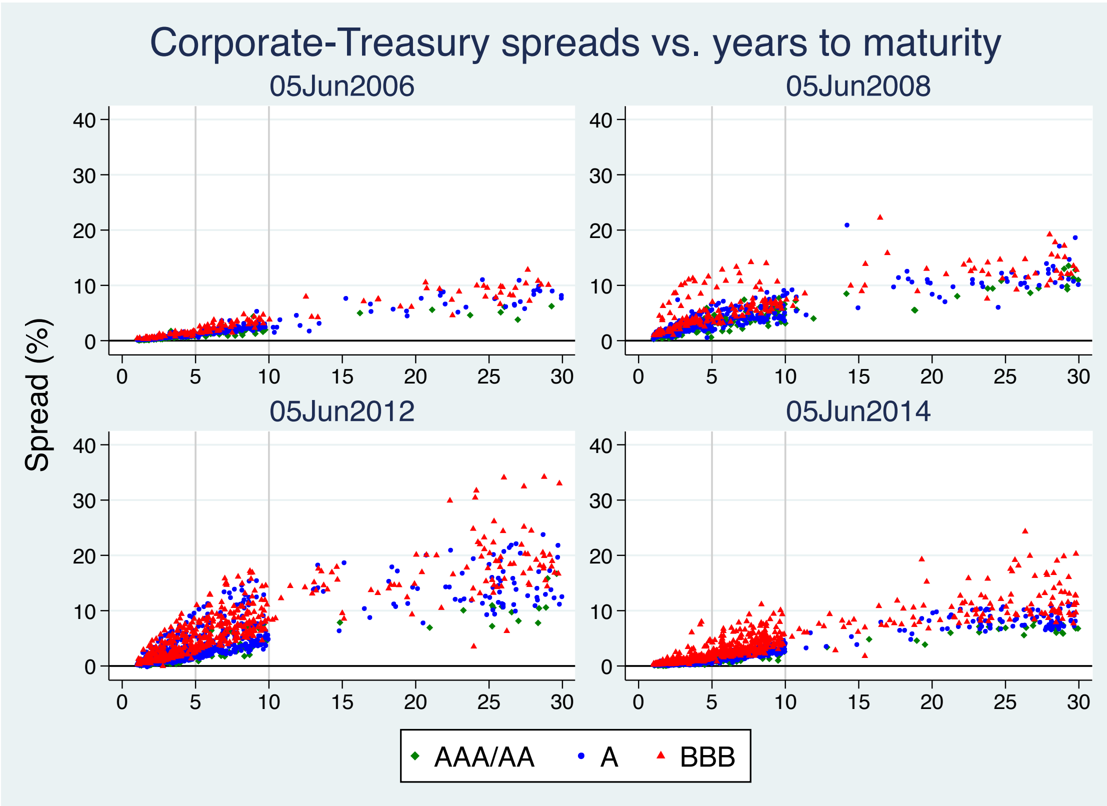

Figure 1.5: *Visualizing potential proprietary trading opportunities.* Bond spreads as a function of residual maturity for four randomly chosen dates, by rating category (bonds with maturity more than 30 years not shown). These graphs illustrate four trading opportunities that could be perceived by broker-dealers or their customers. The first two ones show up as the scatter's thickness: bonds within the same rating category and callability (not distinguished) may have spreads considered too different (upper right and lower left panels); and bonds within the same maturity (irrespective of rating/callability) may have different spread. The third one is related to the slope of the scatter: bonds of different maturities may trade at spreads considered too different. The fourth one relates to the level of the scatter: corporate bonds as an asset class may be considered cheap with respect to Treasuries. Strategies involve selling bonds considered expensive and buying bonds considered cheap.

Dealer proprietary trading again implies $\nu_k^x < 0$ for at least one measure; I also expect the coefficient for one-way order flow and criterion $Z \leq 10$ years to be the largest in absolute terms and the most significant. For instance bonds that are cheap with respect to similar ones are those with high spread with respect to the median in its similarity group ($y_{i,t-1} - y_{i,t-1}^{sim} > 0$) and are expected under dealer arbitrage to be associated with more customer sales: thus we expect a negative coefficient ν_1.

Results

Table 1.8 shows the estimation results for One-way order flow, while table 1.14 in the appendix shows them for Partial Roundtrip order flow.

The first column in table 1.8 gives a negative significant coefficient for the *CreditCall* component of the spread and a highly significant for the *TreasConv* component: dealers appear to manage bonds by maturity bucket irrespective of rating and callability; similarly they buy corporate bonds as long as they are overall cheap with respect to Treasury bonds. In table 1.14, in the first column only the *TreasConv* component is negative significant, to a weak level. The second column, where observations are split according to bond age, gives similar results.

The results are more striking when observations are split according to bond initial maturity or residual maturity: the *CheapInSim* measure becomes highly significant, as the *CreditCall* and *TreasConv* components. For shorter maturities, dealers appear to buy (sell) more bonds that are cheap (expensive) with respect to similar bonds in terms of rating, maturity and callability, with respect to bonds of the same maturity, and with respect to Treasury bonds. The *CreditCall* component appears especially strong: the coefficient implies that a one percentage point increase in this measure is associated with an increase by $-(10^{-.1211} - 1) = 24\%$ in dealer purchases.

1.5.3 Evolution through time and plausible impact of regulation

In this section I run regression (1.5.4), splitting by age and maturity being below or above 10 years on the same subperiods as for price impact regressions.

Tables 1.9 gives the results with the split over initial maturities (results are very similar with residual maturity). To save space and focus on proprietary trading behavior I show only coefficients for one-way order flow and bonds with initial maturity $M \leq 10$ years. I show the other coefficients in table 1.15 in appendix.

The evolution across periods is clear: before the crisis, the coefficients are larger in absolute value than during the Post-Dodd-Frank period, especially for the *CreditCall* and *TreasConv* measures that are the most relevant. If the *CreditCall* measure for a shorter maturity bond (2nd line) is one percentage point higher, dealers tend to buy it more by $10^{-.4542} - 1 = 65\%$ during the Pre-Crisis period (second column), while it falls to 15% after the crisis (sixth column). For the *TreasConv* measure, the effects of a one percentage point increase are 62% more dealer purchase before

Table 1.8: Order flow regressed on four lagged measures of cheapness. $Idiosync$ is the difference between the bond's spread $y_{i,t}$ to an equivalent Treasury bond minus the median $y_{i,t}^{sim}$ of these spreads in the basket of bonds with the same credit rating, maturity and callability. $CreditCall_{i,t}$ is equal to $y_{i,t}^{sim}$ minus the median spread for all bonds with the same maturity $y_{i,t}^{\tau}$. $Maturity_{i,t}$ is equal to $y_{i,t}^{\tau}$ minus the median $y_{i,t}^{10}$, the median spread of all bonds that have residual maturity below 10 years if it is the case for bond i, or above 10 years otherwise. Controls include the dummy $OneWay_{i,t}$ for one-way order flow, and where applicable the dummy for the criterion Z (bond age, initial maturity and residual maturity) being less than 10 years, and its interaction with $OneWay_{i,t}$. Additional controls are 10 lags of order flow, 10 lags of bond price changes, 3 lags of stock return and lagged 10 years Treasury yield changes. To save space **only the coefficients for One-way order flow are shown**. Coefficients for Roundtrip order flow are shown in table 1.14.

	NoCrit	Age	InitMaturity	ResidMaturity
$Idiosync_{i,t-1}^{OneWay}$	-0.0127			
	(-1.83)			
$Idiosync_{i,t-1}^{OneWay,Z\leq 10y}$		-0.0116	**-0.0496*****	**-0.0492*****
		(-1.62)	**(-4.58)**	**(-4.82)**
$Idiosync_{i,t-1}^{OneWay,Z>10y}$		-0.0265	0.0095*	0.0132**
		(-1.96)	(2.04)	(2.77)
$CreditCall_{i,t-1}^{OneWay}$	**-0.0176***			
	(-2.08)			
$CreditCall_{i,t-1}^{OneWay,Z\leq 10y}$		**-0.0206***	**-0.1211*****	**-0.1215*****
		(-2.19)	**(-11.46)**	**(-12.36)**
$CreditCall_{i,t-1}^{OneWay,Z>10y}$		-0.0100	0.0119*	0.0179**
		(-0.71)	(2.11)	(3.13)
$Maturity_{i,t-1}^{OneWay}$	0.0156***			
	(3.55)			
$Maturity_{i,t-1}^{OneWay,Z\leq 10y}$		0.0181***	0.0321***	0.0270***
		(3.75)	(3.91)	(3.53)
$Maturity_{i,t-1}^{OneWay,Z>10y}$		-0.0055	0.0082	0.0088
		(-0.51)	(1.59)	(1.63)
$TreasConv_{i,t-1}^{OneWay}$	**-0.0443*****			
	(-9.12)			
$TreasConv_{i,t-1}^{OneWay,Z\leq 10y}$		**-0.0449*****	**-0.0733*****	**-0.0664*****
		(-8.31)	**(-5.77)**	**(-5.84)**
$TreasConv_{i,t-1}^{OneWay,Z>10y}$		-0.0017	**-0.0155***	**-0.0587*****
		(-0.18)	**(-2.38)**	**(-5.86)**
Constant and controls	Y	Y	Y	Y
R^2	0.00	0.00	0.00	0.00
N	2,316,162	2,316,162	2,316,162	2,316,162

Standard errors clustered by bond issuer

* (p<0.05), ** (p<0.01), *** (p<0.001)

Table 1.9: Order flow regressed on four lagged components of bond spread $Idiosync_{i,t-1}$, $CreditCall_{i,t-1}$, $Maturity_{i,t-1}$ and $TreasConv_{i,t-1}$ (see table 1.8 for details on the components), broken down by One-way / partial Rountrip order flow and bond initial maturity M being below or above 10 years. Controls include the dummy $OneWay_{i,t}$ for one-way order flow, and where applicable the dummy for bond initial maturity M being less than 10 years, and its interaction with $OneWay_{i,t}$. Additional controls are 10 lags of order flow, 10 lags of bond price changes, 3 lags of stock return and lagged 10 years Treasury yield changes. To save space **only the coefficients for one-way order flow and bonds of maturity $M \leq 10$ years are shown**. Other coefficients are in table 1.15

	Opaque	Pre-Crisis	Crisis	Post-Crisis	Dodd-Frank
$Idiosync_{i,t-1}^{OneWay,M\leq 10y}$	-0.0886**	-0.0528	-0.1188***	-0.0243	-0.0029
	(-3.02)	(-0.67)	(-7.79)	(-1.61)	(-0.24)
$CreditCall_{i,t-1}^{OneWay,M\leq 10y}$	-0.1310***	-0.4542***	-0.1315***	-0.0865***	-0.0717***
	(-3.94)	(-6.79)	(-7.67)	(-3.38)	(-3.49)
$Maturity_{i,t-1}^{OneWay,M\leq 10y}$	0.0391	-0.0134	0.0440*	0.0511*	0.0235*
	(0.88)	(-0.26)	(2.24)	(2.16)	(2.36)
$TreasConv_{i,t-1}^{OneWay,M\leq 10y}$	-0.1493**	-0.4195***	-0.1140***	-0.1039***	-0.0830***
	(-3.16)	(-6.70)	(-5.16)	(-4.04)	(-4.51)
R^2	0.00	0.01	0.01	0.00	0.00
N	327,389	328,629	276,026	242,140	1,141,978

Standard errors clustered by bond issuer.
* ($p<0.05$), ** ($p<0.01$), *** ($p<0.001$)

the crisis, and 17% after the crisis.

This is likely to be related to post-crisis regulation, what was indeed intended by the Volcker rule: it seems that prop trading was reduced. The Volcker rule was fully enforced in 2015, after the end of my sample. As widely noticed in the literature, it may have had effects well before 2015: it was announced as the Dodd-Frank act was voted on July 21st, 2010, and Bessembinder et al. (2018) noticed that many large investment banks shut their proprietary trading desks as early as 2011-2012.

1.5.4 Consistency with evidence on Primary Dealers

In order flow regression results, the measure $TreasConv$ comes with a high and highly significant coefficient: broker-dealers thus appear to purchase more corporate bonds to the extent they are cheap with respect to Treasury bonds, possibly even after adjustment for risk. Primary Dealers trading activity is included in my sample, as Primary Dealers are registered as broker-dealers. This is consistent with evidence from Primary Dealers in figure 1.1: Primary Dealers accumulated bond inventories, while they borrowed Treasury bonds to sell them; figure 1.1 also suggests this compressed corporate bond spreads.

In addition, the coefficient on the measure $TreasConv$ sharply decreases at the onset with the crisis and remains stable until after the crisis: this is again consistent with figure 1.1, as Primary Dealers short Treasury position was cut at the onset of the crisis, and did not reconstruct their long corporate, short Treasury position. The coefficient on $TreasConv$ is not zero after the crisis however: this simply suggests that broker-dealers other than Primary Dealers are exploiting corporate-to-Treasury spreads. This also suggests that Primary Dealers had a comparative advantage in this strategy, either because of their Primary Dealer status (they underwrite and make markets for Treasury bonds) or to the bank status many Primary Dealers have, or for another reason.

1.6 Why proprietary trading stopped in July 2007: the plausible role of margin constraints and capital requirements

As shown by figure 1.1, Primary Dealers were net long in corporate bonds, and net short in Treasury bonds before the crisis. However, in July 2007 the Treasury position was cut by half, leaving Primary Dealers with an unhedged interest rate exposure on their corporate bond holdings. Primary Dealers started shrinking their corporate bond inventory only months later, in January 2008.

In the following I explore plausible causes for this. First I give suggestive evidence that Primary Dealers were facing separate financing constraints for each leg of the strategy, with two origins. Second, I give suggestive evidence that Primary Dealers' short Treasury position shrinkage in July 2007 was related to a tightening of the financing constraint on the Treasury side. Overall, the evidence appears consistent with the assumption made in Gromb and Vayanos (2002) on arbitrageurs' financing contraints.

1.6.1 Financing constraints #1: repos and reverse repos

A proprietary trader's long and short positions are implemented as follows: the long positions are implemented with repurchase agreements (repos) and the short by reverse repos - Primary Dealers lend cash to their counterparties so that they get the desired collateral, and sell this collateral. I give more details in appendix 1.12 and what a proprietary traders' balance sheet looks like (figure 1.11).

Repo financing of corporate bonds. Together with corporate bond net outright position (solid thin blue line), Figure 1.6 plots Primary Dealers' net financing positions for corporate bonds (solid thick blue line). Net financing are net funds received with an opposite transfer of corporate bonds for financing purposes: repos and reverse repos, security lending and borrowing, collateralized borrowing and loans, *etc.*, that we generically label "net repo position" for convenience. It is positive if PD are net cash borrowers, *i.e.* net security lenders.

It also plots the position in such contracts that are easily "runnable" (dashed red line): with overnight maturity or on a continuing basis, *i.e.* with no specific maturity but that can be ended on demand. The difference between the solid thick line (all repos) and the dashed line (overnight repos) is thus net amount of repos with maturity of at least two days. The match of both curves with corporate bond net outright position is good, suggesting that most of Primary Dealers' corporate bond position was indeed financed with runnable repos, and that corporate bonds were mainly used as collateral for the outright position purpose.

Drop in net reverse repos financing of Treasury bond. Figure 1.7 plots the same graph for Treasury positions. The thick line shows that Primary Dealers are structurally net Treasury borrowers (they lend funds and receive Treasury securities as collateral), and that they borrowed more securities than they sold short as comparison with the outright position shows. The securities borrowed that were not sold were likely to be kept as collateral for loans provided by Primary Dealers to their customers. After the crisis, Primary Dealers were still net Treasury borrowers, while they held Treasury bonds outright.

The July 2007 cut in the short position was associated with a cut of similar magnitude in the total net repo position. The *overnight* net repo position (dashed line) exhibits a similar break of comparable magnitude; it also increases a few months ahead (end of 2006) without a similar pattern in total position: this suggests that security lenders gradually shortened the maturity of their loans before they started to run on the security lending contracts.

Therefore it is likely that Primary Dealers' financing constraint on the reverse repo position became more binding in July 2007, which imposed a reduction in the

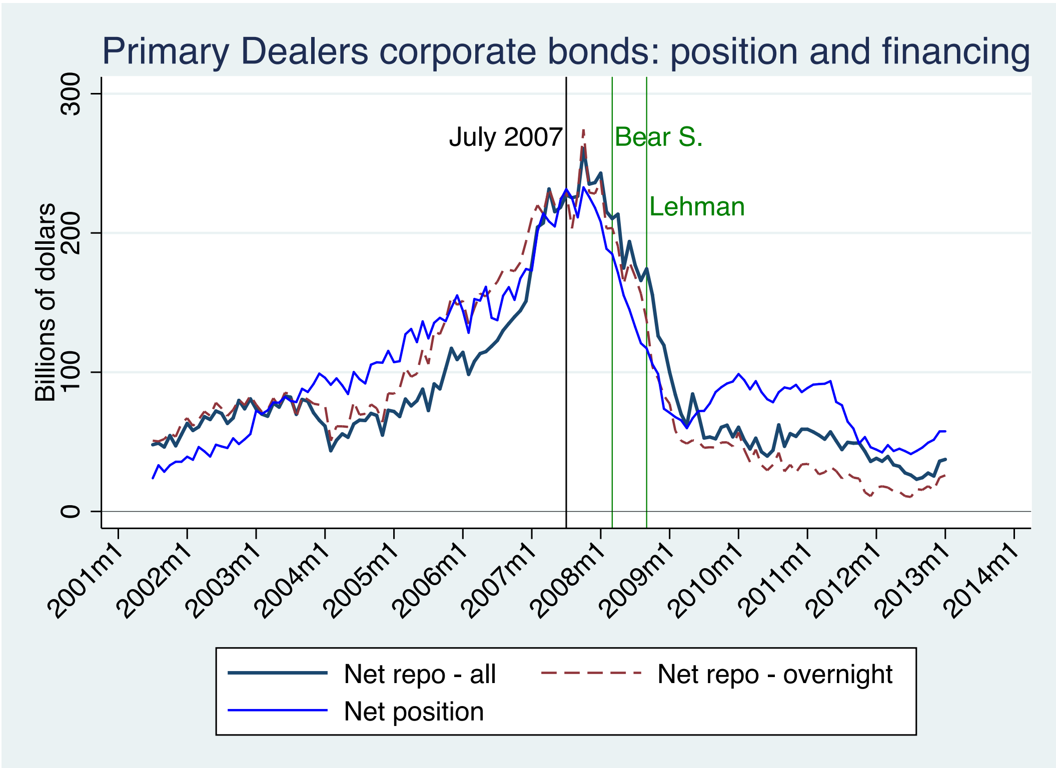

Figure 1.6: **Repo financing of Primary Dealers' corporate bond inventories.** This graph plots Primary Dealers corporate bond inventory (solid, thin blue line), together with net borrowing collateralized by corporate bonds (thick, solid black line) and net borrowing with overnight maturity (dashed red line). The good fit between the three curves before the crisis suggests that most corporate bond inventories were funded through very short-term repos.

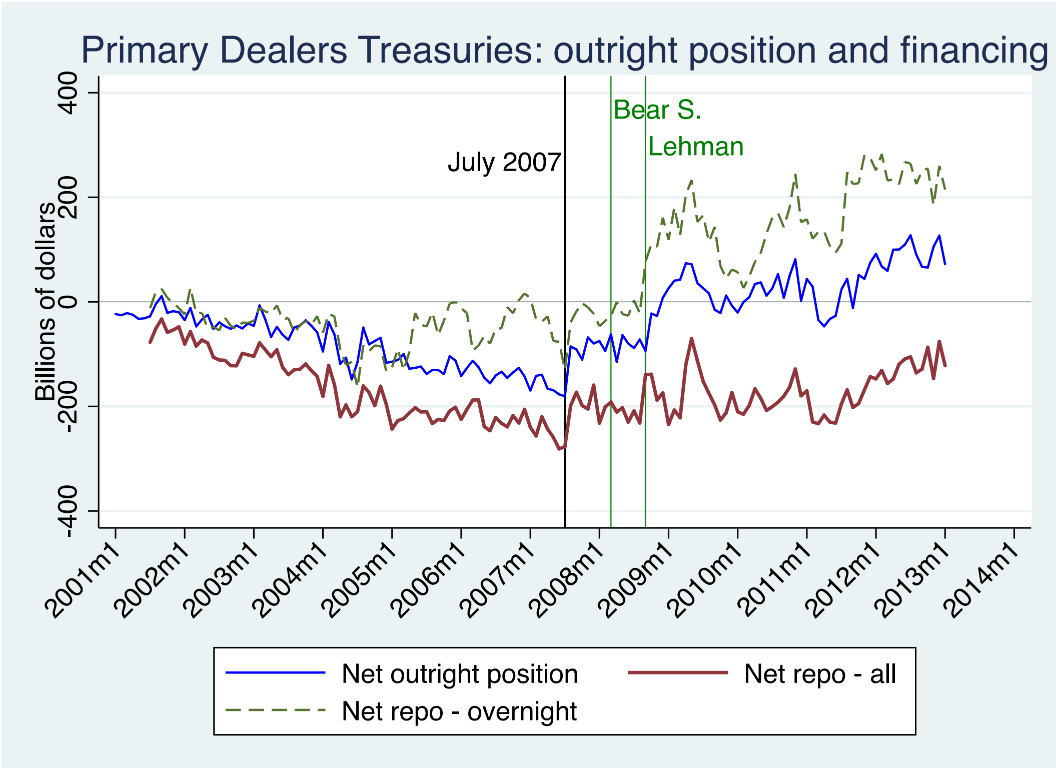

Figure 1.7: **Reverse repo funding of Treasuries and run.** This graph plots Primary Dealers net inventories in Treasury securities (excluding inflation-protected and T-Bills), together with net collateralized borrowing with Treasuries as collateral (thick, solid brown line) and net collateralized borrowing collateralized by Treasuries that are overnight (dashed green line)

Treasury short position. However, similar constraints stem from regulatory capital requirements, as exposed in the next subsection.

1.6.2 Financing constraints #2: regulatory capital requirements

Broker-dealers also face capital requirements set by the Securities Exchange Act of 1934 and implemented by their regulator, the Security and Exchange Commission.[18] The most important points for my purpose here are the following:

- these capital requirements have to be met at all times,[19] and the broker-dealer has to notify the SEC or the FINRA immediately when it is approaching the limit. In practice, it seems that daily mark-to-market is the lowest frequency of computation admitted by FINRA.[20]
- Broker-dealer equity, net of haircuts on securities long and on short positions, is higher than a fraction of dealer indebtedness
- The haircuts can be computed using dealers' own internal statistical models

1.6.3 Financing constraints at the onset of the 2007-2009 crisis: suggestive evidence

The July 2007 drop in the short Treasury position may have been caused by a tightening of haircuts on the Treasury borrowing contracts, a supply effect. By contrast, it could also be Primary Dealers who reduced their demand for Treasury borrowing because they were willing to reduce the short position. I am not able to unambiguously test one hypothesis against the other because I am not aware of dataset on haircuts over the period, but several elements point to Primary Dealers undergoing a tightening of haircuts.

First, after July 2007 Primary Dealers had a new exposure to interest rate risk on half of their corporate bond inventories, as shown on figure 1.1: they subsequently liquidated these inventories, suggesting that this new exposure was not desired. Lehman's collapse in September 2008 and a further drop in the Treasury short position came just the long-short position became balanced.

Second, I give suggestive evidence that the position cut was associated with an increase in haircut on Treasury reverse repos. Haircuts for Primary Dealers are likely to increase with the underlying asset volatility (as a proxy for the risk of price decrease for repo, or increase for reverse repo) and with Primary Dealers' default risk.

[18] Capital requirements are defined in Rule 15c3-1 of the Securities Exchange Act of 1934, commonly referred to as the "Net Capital Rule" dates back to 1975.

[19] Including intraday

[20] https://www.finra.org/rules-guidance/key-topics/portfolio-margin/faq

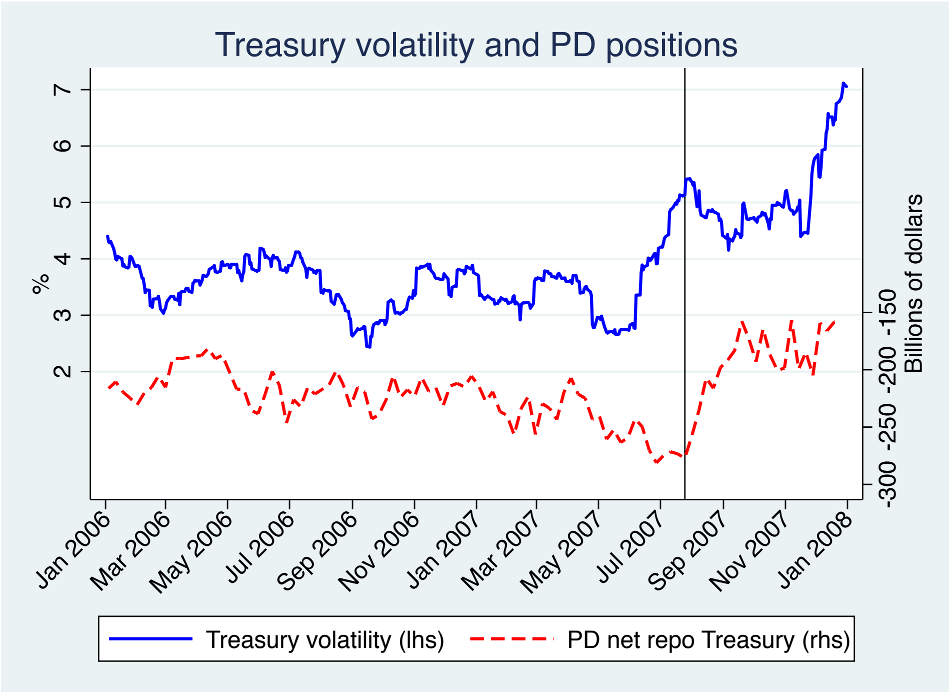

Figure 1.8: This graph plots the 2-months rolling-window standard deviation of 10 years Treasury yield daily changes and Primary Dealers' (PDs') net reverse repo position in US Treasuries, *i.e.* minus the amount of Treasuries borrowed by Primary Dealers. The black vertical line is on July 25th, 2007 when PDs' outright position in Treasuries began to shrink in absolute terms: it coincides with a decrease in the PD Treasury borrowing. The volatility of the 10 years yield increased by half in June 2007.

Regarding asset volatilities, figures 1.8 and 1.9 plot the two-month rolling-window volatility of daily changes in the 10 years Treasury yield, and on the cross-sectional median two-month rolling-window volatility of daily returns on a portfolio of corporate bonds. The portfolio retains bonds within my sample whose maturity is between 4 and 6 years and whose rating is at least AA. This relatively tight subsample avoids volatility related to slope and thickness of the spread curve, but the results carry over when the portfolio is broadened.

These figures show that the volatility of 10 years bond clearly increased in June 2007, from a recent history average level between 3 and 4% to about 5.5%, and increase by roughly 50%. By contrast, the historical volatility of corporate bonds did not appear to rise much with respect to the levels observed in the previous 18 months.

However, earlier in the sample the volatility of 10 years Treasury bonds was higher, up to 6%. This suggests that asset volatility is not the only driver of haircuts. Another potential determinant is PD default risk, as argued by Copeland et al. (2010) from data starting in March 2008. Figure 1.10 plots an equal-weighted index of 7 Primary Dealers CDS (the major investment banks with US parents), as a proxy

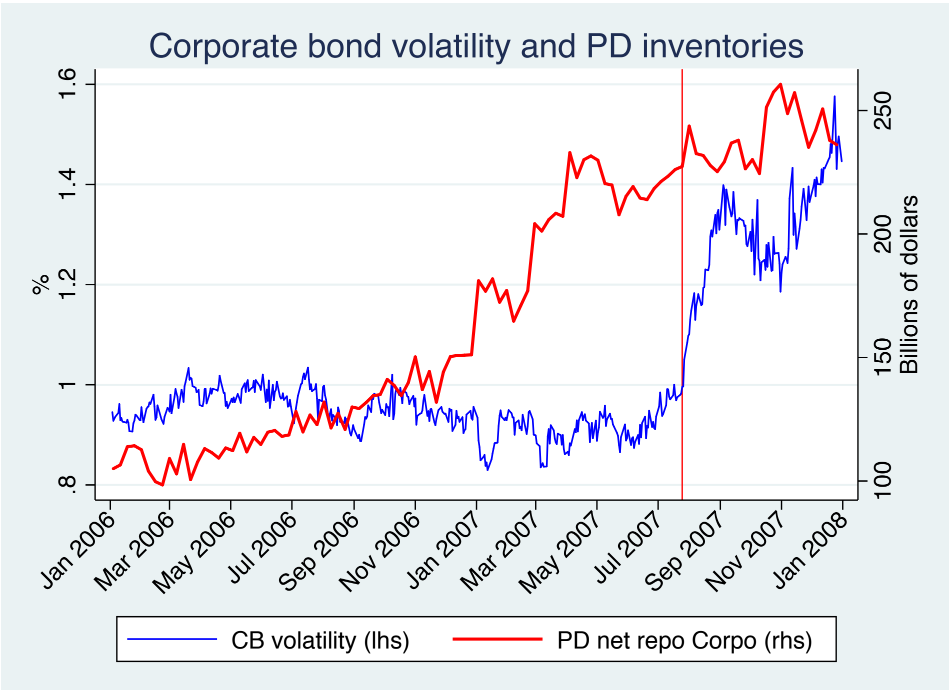

Figure 1.9: 2 month rolling-window volatility of corporate bonds (returns on the median price of a portfolio investment grade, maturities between 4 and 6 years) and Primary Dealers' net position in repo contracts involving corporate bonds. The vertical line marks 25th July 2007 when PDs started shrinking the net short Treasury position. The graph shows basically no connection between the two even during the weeks before July 2007. It also shows that corporate bond volatility rose right after the short Treasury position cut.

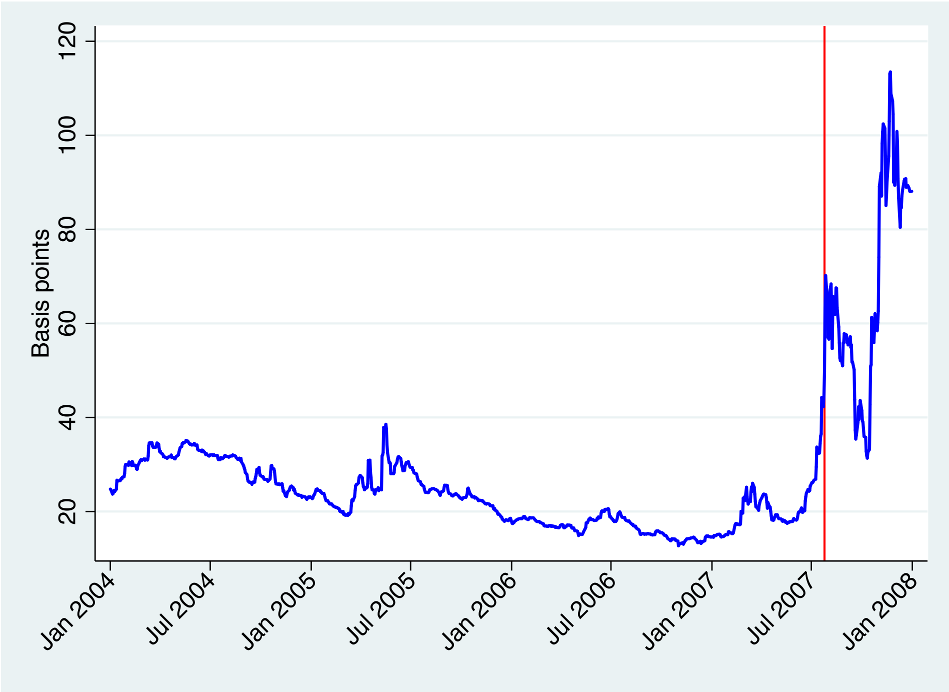

Figure 1.10: CDS index for 7 Primary Dealers with US parent. This graph plots and equal-weighted five years CDS index for all Primary Dealers with available CDS in US dollars, for senior unsecured debt except Bear Stearns (subordinated debt). These include the major US investment banks. Individual CDS movements are broadly parallel. The vertical line is on July 25th, 2007, when Primary Dealers' aggregate short Treasury position began to shrink.

for default risk (individual movements are broadly parallel): CDS spreads increased sharply in July 2007.

However, default risk alone cannot explain haircut increases, as these should also have impacted repos involving corporate bonds and thus simultaneous shrinkage of both the long corporate position and the short Treasury position.

Overall this suggests that lender may begin to start revising haircuts when default risk becomes more of a concern. Default risk may arise because of other Primary Dealer activities than the proprietary trading strategy I document.

1.7 Implications for financial regulation

The results in my paper show that dealers are not only market makers, but also proprietary traders who exploit price spreads they think not justified by fundamentals. The Volcker rule bans the latter for bank-affiliated dealers, while in principle still allowing the former.

Proprietary trading is not *per se* bad. By uncovering some proprietary trading strategies, I do not mean to say what regulation is optimal, if regulation is needed. Proprietary trading as described by theories of limits of arbitrage is to some extent liquidity provision, and bears many similarities with market making as described by above theories. In both cases, the market is fragmented, and market makers or prop traders bridge the gap between investors eager to sell and investors eager to buy, which is a priori improving social welfare of all market participants.[21]

The macro-finance literature suggests that when broker-dealers buy, this compresses risk premia (Adrian and Shin 2009, Adrian et al. 2014). Again this does not necessarily means that they take excessive risk: if markets are fragmented before broker-dealers enter, risk-sharing is limited and asset risk premia are high. When broker-dealers enter, they may improve risk sharing, so that all investors are more willing to take risk and compress equilibrium risk premia: in this case this is socially optimal.

"Risk-shifting" and "margin constraint" problems Indeed, Volcker (2010) asserts that proprietary trading is socially useful even if risky. The assumption underlying the Volcker rule is that broker-dealers that are in large banking groups have access to public bailouts or liquidity backstops by the Fed: thus they do not internalize the downside risk of their strategies, leading them to take excessive risk. In what follows I label this assumption the "risk-shifting problem".

There are two interpretations of the risk-shifting problem. The first one relates to proprietary trading strategies themselves, as they are inherently too risky: for instance, dealers could bet on spreads that reflect differences in issuers' fundamentals with 50% probability. If it is not the case, dealers win, if indeed fundamentals matters, depositors lose.

A second interpretation of the risk-shifting problem is that it stems from a margin constraint problem: even if they correct true mispricings, dealers are subject to financing constraints that have to be met at interim dates before the strategy pays off. A risk is that these constraints bind for a reason unrelated to the proprietary trading strategy, forcing liquidation at fire sales prices and thus causing losses. In section 1.6 I suggest this is an important driver of the 2007-2009 financial crisis. Fire sales losses may be the risk that is not internalized by bank-affiliated dealers because they expect to pass it to depositors or taxpayers.

However, as I show in the next paragraph, a growing literature points to the role of the pecuniary externality that these financing constraints generate.

The Volcker rule only partially addresses the margin constraint problem. The margin constraint problem is likely to arise even outside the banking sector. Several papers have noted that a pecuniary externality stemming from these margin constraints, together with competition between dealers, leads to ex ante too large

[21] Although Hart (1975) shows that welfare improvement is warranted only if market makers or prop traders *fully* complete the markets.

dealer positions and thus larger losses.[22]

In the risk-shifting case, bank-affiliated dealers take excessive risk because the perspective of bank public bailout make them insensitive to losses: the externality is on bank's creditors. Therefore it is largely unclear that the Volcker rule in its principle has prevented the financial system from proprietary trading-related risks.

If proprietary trading needs regulation, then regulation should address the margin constraint problem. This is beyond the scope of this paper.

1.8 Conclusion

In this paper I show that not all dealer trading in the US corporate bond market is market making in the sense of standing ready to buy or sell on customer demand. Dealers also trade actively to exploit spreads in the cross-section of bonds that may or may not be justified by fundamentals. Shedding light on their behavior contributes to a better understanding of the findings of the intermediary asset pricing literature, and to the understanding of crises.

The intermediary asset pricing literature has found that broker-dealer leverage and/or asset growth is an explanatory factor of asset prices, but it remains elusive on what this leverage consists in. This paper suggests that long-short strategies that tend to compress spreads are an important determinant of dealer leverage.

Many papers have found that the level of corporate bond spreads predicts future economic activity. I have shown that dealers are responsive mainly to the spread between Treasury bonds and corporate bonds, both from transaction and holding data, and that this responsiveness is correlated with the level of spreads in the aggregate. Therefore dealers may have a central role in shaping asset prices and investment.

In addition, proprietary trading appears to come with constraints that began to bind at the onset of the crisis, likely forcing dealer to liquidate their positions at fire sales prices and triggering the corporate bond market crisis. This problem has at best been partially addressed by existing regulation such as the Volcker rule: therefore such regulation may not preclude forced liquidation even with tight regulation of the banking sector.

[22] *Cf. e.g.* Gromb and Vayanos (2002), Lorenzoni (2008), Brunnermeier and Pedersen (2008), Stein (2012)

1.9 Appendix 1: Data

Cleaning. For each bond, I drop transactions within 7 days before and 7 days after the bond's offering date, because primary market transaction are very specific.

TRACE includes transaction information with prices, quantity, the direction of the trade (buy or sell) and whether the counterparty to the reporting dealer is a customer or another dealer. I clean the data following Dick-Nielsen (2014) to remove explicit reporting errors, when-issued transactions and special trading under special circumstances. I remove interdealer transactions.

I remove transactions with price lower than 80% or higher than 120% of the median price of the day, or with price lower than \$1 or higher than \$500. I remove transactions with amounts lower than the par value of one bond, or higher than the offered amount for the bond are removed.

Stock prices and returns. I retain bonds for which a stock price is available (average of end-of-day bid and ask), using the WRDS Bond-CRSP linking suite at PERMCO (company) level for CRSP. I use common stocks (CRSP share code's first digit equal to 1); whenever there are several common stock classes for a single company, I use the average common stock prices weighted by the number of shares outstanding. I do not adjust stock returns are not adjusted for dividend payment or annoucement: dividend decisions may have an impact on the company's perceived credit risk.

1.10 More on testing hypothesis 1

1.10.1 Regression with cheapness measures

To address the concern that cheapness measures defined in equation (1.5.3) may forecast both dealer purchases (sales) and price decreases (increases), I add these measures in regression (1.4.3), interacting them with the $Lehman_t$ dummy.

$$
\begin{aligned}
\Delta \log p_{i,t} = \alpha &+ \beta_1 \widetilde{OF}_{i,t}^{OneWay} + \beta_2 \widetilde{OF}_{i,t}^{Roundtrip} \\
&+ \beta_3 \widetilde{OF}_{i,t}^{OneWay} \times Lehman_t + \beta_4 \widetilde{OF}_{i,t}^{Roundtrip} \times Lehman_t \\
&+ \sum_x \big(\nu_1^x Idiosync_{i,t-1}^x + \nu_2^x CreditCall_{i,t-1}^x \\
&\qquad + \nu_3^x Maturity_{i,t-1}^x + \nu_4^x TreasConv_{i,t-1}^x \big) \\
&+ \sum_k LehmanInteractionTerms_{i,t} \\
&+ \gamma X_{2,i,t}^{(p)} + \epsilon_{3,i,t}
\end{aligned} \tag{1.10.1}
$$

where the $LehmanInteractionTerms$ are the cheapness measures $Idiosync^{OneWay}_{i,t-1}$, $Idiosync^{Roundtrip}_{i,t-1}$, $CreditCall^{OneWay}_{i,t-1}$, *etc.* times the dummy $Lehman_t$, and

$$x = OneWay, Roundtrip$$

I also include the $Lehman_t$, $OneWay_{i,t}$ dummies and their interaction in the vector of controls $X^{(p)}_{3,i,t}$.

Table 1.10 shows that the coefficient on One-way order flow is unaffected by the presence of the cheapness measures: it stands at 0.0024, while it was at 0.0025 in table 1.5

1.10.2 Regression with predicted order flow

I also assess whether other predictors or order flow are driving the results: previous price changes, lags of order flow and lagged stock return and interest rate changes. To do this I compute a predicted and an unexpected component of order flow, separately for One-way and Partial Rountrip order flow, as follows:

$$\widetilde{OF}^{x}_{i,t} = \underbrace{\alpha_x + \sum_{k=1}^{10} \rho_k \widetilde{OF}_{i,t-k} + \sum_{k=1}^{10} \mu_k \Delta \log p_{i,t-k} + \gamma X_{1,i,t-1}}_{\widehat{\widetilde{OF}}^{x}_{i,t}} + \eta^{x}_{i,t}$$

where the vector $X_{1,i,t-1}$ includes lagged stock return and interest rate changes.

Then I regress price changes on the predicted and unpredicted components of order flow, still with the interaction with the $Lehman_t$ dummy as in the main regression:

$$\begin{aligned}
\Delta \log p_{i,t} = \alpha &+ \beta_1 \widehat{\widetilde{OF}}^{OneWay}_{i,t} + \beta_2 \widehat{\widetilde{OF}}^{Roundtrip}_{i,t} \\
&+ \beta_3 \widetilde{OF}^{OneWay}_{i,t} \times Lehman_t + \beta_4 \widetilde{OF}^{Roundtrip}_{i,t} \times Lehman_t \\
&+ \beta_0 Lehman_t + \gamma X^{(p)}_{2,i,t} + \epsilon_{2,i,t}
\end{aligned}$$

The regressor of interest are the unexpected components of order flow $\eta^{OneWay}_{i,t}$ and $\eta^{Roundtrip}_{i,t}$.

Table 1.11 presents the results. The coefficient on $\eta^{OneWay}_{i,t}$ and $\eta^{Roundtrip}_{i,t}$ are very close to the estimates of One-Way and Partial Roundtrip order flow in table 1.5, which were at -0.0025 and 0.0040 respectively.

1.10.3 Robustness to multiway clustering

I re-run regression 1.10.1, computing standard errors in more conservative ways: I allow clustering by the number of years to bond maturity, and by calendar month, following the methodology by Cameron et al. (2011). One concern with the maturity clustering is that the number of clusters along this dimension is close to 30, which may not be sufficient. In any case the results are robust, as shown by table 1.12.

Table 1.10: Regression of daily log price changes on customer order flow, cheapness measures and controls. $\widetilde{OF}_{i,t}$ is the sign of order flow times the logarithm of the absolute value of order flow. Order flow is the sum of customer large (above \$100,000) buys minus the sum of customer large sells. $\widetilde{OF}_{i,t}^{OneWay}$ equals $\widetilde{OF}_{i,t}$ if order flow in bond i on day t is one-way (only customer buys or only customer sells), and zero otherwise. $\widetilde{OF}_{i,t}^{Roundtrip}$ equals $\widetilde{OF}_{i,t}$ if order flow is not One-way, *i.e.* (partial) roundtrip, and zero otherwise. Cheapness measures are defined in equation (1.5.3). Controls are issuer stock return, changes in 10 years Treasury yield, changes in 3-months LIBOR, TYVIX, an implied volatility index for Treasury futures, and changes in TYVIX. In the second and third column, a dummy $OneWay_{i,t}$ for one-way order flow is included. $Lehman_t$ is a dummy that equals 1 if t is between September 15th, 2008 and April 30th, 2009. The interaction $OneWay_{i,t} \times Lehman_t$ is included.

	$\Delta \log p_{i,t}$
$\widetilde{OF}_{i,t}^{OneWay}$	**-0.0024*****
	(-6.10)
$\widetilde{OF}_{i,t}^{Roundtrip}$	0.0041***
	(11.82)
$CheapInSim_{i,t-1}^{OneWay}$	0.0014
	(1.37)
$CreditCall_{i,t-1}^{OneWay}$	0.0023
	(1.79)
$Maturity_{i,t-1}^{OneWay}$	-0.0009
	(-0.88)
$TreasConv_{i,t-1}^{OneWay}$	0.0045***
	(5.83)
$CheapInSim_{i,t-1}^{Roundtrip}$	0.0040**
	(2.69)
$CreditCall_{i,t-1}^{Roundtrip}$	0.0019*
	(2.52)
$Maturity_{i,t-1}^{Roundtrip}$	0.0014
	(1.79)
$TreasConv_{i,t-1}^{Roundtrip}$	0.0022***
	(4.37)
Interaction with $Lehman_t$	Y
Constant and controls	Y
R^2	0.24
N	2,203,723

Standard errors clustered by bond issuer

* (p<0.05), ** (p<0.01), *** (p<0.001)

Table 1.11: Regression of daily log price changes on customer order flow, cheapness measures and controls. $\widetilde{OF}_{i,t}$ is the sign of order flow times the logarithm of the absolute value of order flow. Order flow is the sum of customer large (above $100,000) buys minus the sum of customer large sells. $\widetilde{OF}_{i,t}^{OneWay}$ equals $\widetilde{OF}_{i,t}$ if order flow in bond i on day t is one-way (only customer buys or only customer sells), and zero otherwise. $\widetilde{OF}_{i,t}^{Roundtrip}$ equals $\widetilde{OF}_{i,t}$ if order flow is not One-way, *i.e.* (partial) roundtrip, and zero otherwise. Controls are issuer stock return, changes in 10 years Treasury yield, changes in 3-months LIBOR, TYVIX, an implied volatility index for Treasury futures, and changes in TYVIX. In the second and third column, a dummy $OneWay_{i,t}$ for one-way order flow is included. $Lehman_t$ is a dummy that equals 1 if t is between September 15th, 2008 and April 30th, 2009. The interaction $OneWay_{i,t} \times Lehman_t$ is included.

	$\Delta \log p_{i,t}$
$\eta_{i,t}^{OneWay}$	-0.0027***
	(-6.47)
$\eta_{i,t}^{Roundtrip}$	0.0037***
	(11.01)
$\widetilde{\widetilde{OF}}_{i,t}^{OneWay}$	-1.5375***
	(-7.17)
$\widetilde{\widetilde{OF}}_{i,t}^{Roundtrip}$	3.4071***
	(7.50)
R^2	0.24
N	2,212,269

Standard errors clustered by bond issuer
* (p<0.05), ** (p<0.01), *** (p<0.001)

Table 1.12: Regression of daily log price changes on customer order flow and controls, clustering standard errors by the variable indicated in column header. $\widetilde{OF}_{i,t}$ is the sign of order flow times the logarithm of the absolute value of order flow. Order flow is the sum of customer large buys minus the sum of customer large sells. A customer buy or sell is large its size is above \$100,000. $\widetilde{OF}_{i,t}^{OneWay}$ equals $\widetilde{OF}_{i,t}$ if order flow in bond i on day t is one-way (only customer buys or only customer sells), and zero otherwise. $\widetilde{OF}_{i,t}^{Roundtrip}$ equals $\widetilde{OF}_{i,t}$ if order flow is not One-way, *i.e.* (partial) roundtrip, and zero otherwise. Controls are bond cheapness measures, issuer stock return, changes in 10 years Treasury yield, changes in 3-months LIBOR, TYVIX, an implied volatility index for Treasury futures, and changes in TYVIX. In the second and third column, a dummy $OneWay_{i,t}$for one-way order flow is included. $Lehman_t$ is a dummy that equals 1 if t is between September 15th, 2008 and April 30th, 2009. The interaction $OneWay_{i,t} \times Lehman_t$ is included.

	Issuer/Month	Issuer/Mat	Issuer/Month/Mat
$\widetilde{OF}_{i,t}^{OneWay}$	-0.0024**	-0.0024***	-0.0024*
	(-2.73)	(-4.36)	(-2.50)
$\widetilde{OF}_{i,t}^{Roundtrip}$	0.0041***	0.0041***	0.0041***
	(9.19)	(6.78)	(5.44)
Cheapness measures	Y	Y	Y
Interaction with $Lehman_t$	Y	Y	Y
Constant and controls	Y	Y	Y
R^2	0.24	0.24	0.24
N	2,203,723	2,203,723	2,203,723

Standard errors clustered by bond issuer

* (p<0.05), ** (p<0.01), *** (p<0.001)

1.10.4 Estimation table for regression through time

Table 1.13 shows the estimation results of regression 1.4.5. I plot the coefficients for each subperiods in figure 1.4.

Table 1.13: Regression of daily log price changes on customer order flow and controls, interacted with subperiod dummies. Coefficients are plotted on figure 1.4 $\widetilde{OF}_{i,t}$ is the sign of order flow times the logarithm of the absolute value of order flow. Order flow is the sum of customer large buys minus the sum of customer large sells. A customer buy or sell is large its size is above $100,000. $\widetilde{OF}_{i,t}^{OneWay}$ equals $\widetilde{OF}_{i,t}$ if order flow in bond i on day t is one-way (only customer buys or only customer sells), and zero otherwise. $\widetilde{OF}_{i,t}^{Roundtrip}$ equals $\widetilde{OF}_{i,t}$ if order flow is not One-way, *i.e.* (partial) roundtrip, and zero otherwise. Controls are issuer stock return, changes in 10 years Treasury yield, changes in 3-months LIBOR, TYVIX, an implied volatility index for Treasury futures, and changes in TYVIX. In the second and third column, a dummy $OneWay_{i,t}$for one-way order flow is included. $Lehman_t$ is a dummy that equals 1 if t is between September 15th, 2008 and April 30th, 2009.

Opaque		
$\widetilde{OF}_{i,t}^{OneWay}$		-0.0069***
		(-6.02)
$\widetilde{OF}_{i,t}^{Roundtrip}$		0.0004
		(0.51)
Pre-Crisis		
$\widetilde{OF}_{i,t}^{OneWay}$		-0.0077***
		(-10.38)
$\widetilde{OF}_{i,t}^{Roundtrip}$		0.0008
		(1.30)
Crisis		
$\widetilde{OF}_{i,t}^{OneWay}$		0.0341***
		(10.04)
$\widetilde{OF}_{i,t}^{Roundtrip}$		0.0193***
		(8.31)
Post-Crisis		
$\widetilde{OF}_{i,t}^{OneWay}$		0.0049***
		(4.60)
$\widetilde{OF}_{i,t}^{Roundtrip}$		0.0063***
		(7.38)
Post-Dodd-Frank		
$\widetilde{OF}_{i,t}^{OneWay}$		-0.0029***
		(-4.98)
$\widetilde{OF}_{i,t}^{Roundtrip}$		0.0049***
		(10.28)***
R^2	53	0.24
N		2,220,248

Clustered by bond issuer

* (p<0.05), ** (p<0.01), *** (p<0.001)

1.11 Appendix 2: More on testing Hypothesis 2

1.11.1 General proprietary trading strategies: coefficients for Roundtrip order flow

Table 1.14 complements table 1.8 by showing coefficients for the four components of the bond spread for Partial Roundtrip order flow. The coefficients are less significant than for One-Way order flow, and in any case of smaller magnitude.

1.11.2 Evolution through time

Initial maturity: other regression coefficients

Table 1.15 complements table 1.9 by giving the coefficients on One-way order flow and initial maturities more than 10 years, and for Partial Roundtrip order flow. Coefficients are not not often significant and in any case of lower magnitude than for one-way order flow and shorter maturities.

Results with residual maturity

Table 1.16 and 1.17 show the results of running regression 1.5.2 by subperiod with a breakdown by bond *residual* maturity.

The results are qualitatively similar to the results with the split by initial maturity. The coefficient for the $TreasConv$ measure for one-way order flow, bonds with maturities below 10 years, is however weaker, although still strongly negative significant. The evolution through time is still in the direction of decreased proprietary trading after the crisis.

Table 1.14: Order flow regressed on four lagged measures of cheapness. *Idiosync* is the difference between the bond's spread $y_{i,t}$ to an equivalent Treasury bond minus the median $y_{i,t}^{sim}$ of these spreads in the basket of bonds with the same credit rating, maturity and callability. $CreditCall_{i,t}$ is equal to $y_{i,t}^{sim}$ minus the median spread for all bonds with the same maturity $y_{i,t}^{\tau}$. $Maturity_{i,t}$ is equal to $y_{i,t}^{\tau}$ minus the median $y_{i,t}^{10}$, the median spread of all bonds that have residual maturity below 10 years if it is the case for bond i, or above 10 years otherwise. Controls include the dummy $OneWay_{i,t}$ for one-way order flow, and where applicable the dummy for the criterion $Crit$ (bond age, initial maturity and residual maturity) being less than 10 years, and its interaction with $OneWay_{i,t}$. Additional controls are 10 lags of order flow, 10 lags of bond price changes, 3 lags of stock return and lagged 10 years Treasury yield changes. To save space **only the coefficients for Roundtrip order flow are shown**. Coefficients for One-way order flow are shown in table 1.8.

	NoCrit	Age	InitMaturity	ResidMaturity
$Idiosync_{i,t-1}^{Roundtrip}$	0.0014			
	(1.20)			
$Idiosync_{i,t-1}^{Roundtrip,Z\leq 10y}$		0.0016	0.0030	0.0029
		(1.24)	(1.41)	(0.28)
$Idiosync_{i,t-1}^{Roundtrip,Z>10y}$		0.0037	0.0007	0.0003
		(1.55)	(0.60)	(1.46)
$CreditCall_{i,t-1}^{Roundtrip}$	-0.0019			
	(-1.58)			
$CreditCall_{i,t-1}^{Roundtrip,Z\leq 10y}$		-0.0025	**-0.0137*****	**-0.0117*****
		(-1.78)	**(-3.52)**	**(-3.44)**
$CreditCall_{i,t-1}^{Roundtrip,Z>10y}$		0.0023	0.0005	0.0002
		(0.97)	(0.43)	(0.16)
$Maturity_{i,t-1}^{Roundtrip}$	0.0004			
	(0.33)			
$Maturity_{i,t-1}^{Roundtrip,Z\leq 10y}$		0.0006	0.0080**	0.0069**
		(0.51)	(2.74)	(2.84)
$Maturity_{i,t-1}^{Roundtrip,Z>10y}$		0.0020	0.0003	0.0003
		(0.90)	(0.30)	(0.29)
$TreasConv_{i,t-1}^{Roundtrip}$	**-0.0033***			
	(-2.36)			
$TreasConv_{i,t-1}^{Roundtrip,Z\leq 10y}$		**-0.0034***	0.0027	0.0020
		(-2.21)	(0.77)	(0.68)
$TreasConv_{i,t-1}^{Roundtrip,Z>10y}$		0.0016	0.0016	0.0023
		(1.13)	(1.30)	(1.20)
Constant and controls	Y	Y	Y	Y
R^2	0.00	0.00	0.00	0.00
N	2,316,162	2,316,162	2,316,162	2,316,162

Standard errors clustered by bond issuer

* (p<0.05), ** (p<0.01), *** (p<0.001)

Table 1.15: Order flow regressed on four lagged components of bond spread $Idiosync_{i,t-1}$, $CreditCall_{i,t-1}$, $Maturity_{i,t-1}$ and $TreasConv_{i,t-1}$ (see table 1.8 for details on the components), broken down by One-way / partial Rountrip order flow and bond initial maturity M being below or above 10 years. Controls include the dummy $OneWay_{i,t}$ for one-way order flow, and where applicable the dummy for bond initial maturity M being less than 10 years, and its interaction with $OneWay_{i,t}$. Additional controls are 10 lags of order flow, 10 lags of bond price changes, 3 lags of stock return and lagged 10 years Treasury yield changes. To save space **only the coefficients for one-way order flow and bonds of maturity $M > 10$ years, and for Partial Roundtrip order flow, are shown**.

	Opaque	Pre-Crisis	Crisis	Post-Crisis	Dodd-Frank
$Idiosync^{OneWay,M>10y}_{i,t-1}$	-0.0190	0.0122*	-0.0011	0.0065	0.0223***
	(-1.92)	(2.18)	(-0.11)	(0.54)	(3.34)
$CreditCall^{OneWay,M>10y}_{i,t-1}$	0.0015	0.0221	-0.0046	0.0199	0.0211**
	(0.19)	(1.52)	(-0.59)	(1.51)	(2.73)
$Maturity^{OneWay,M>10y}_{i,t-1}$	-0.0250**	0.0320***	0.0111	-0.0004	0.0074
	(-2.97)	(3.53)	(1.11)	(-0.03)	(0.94)
$TreasConv^{OneWay,M>10y}_{i,t-1}$	-0.0063	-0.0314	-0.0582***	0.0345*	0.0017
	(-0.30)	(-1.30)	(-6.03)	(2.57)	(0.22)
$Idiosync^{Roundtrip,M\leq10y}_{i,t-1}$	0.0018	-0.0040	-0.0045	0.0076	0.0135*
	(0.15)	(-0.43)	(-1.25)	(1.76)	(2.52)
$CreditCall^{Roundtrip,M\leq10y}_{i,t-1}$	-0.0338**	-0.0802**	-0.0100	0.0034	-0.0024
	(-2.79)	(-3.06)	(-1.79)	(0.33)	(-0.36)
$Maturity^{Roundtrip,M\leq10y}_{i,t-1}$	0.0324*	0.0077	0.0050	0.0088	0.0052
	(2.12)	(0.39)	(0.76)	(1.00)	(1.55)
$TreasConv^{Roundtrip,M\leq10y}_{i,t-1}$	-0.0209	-0.0042	0.0067	-0.0092	-0.0014
	(-1.16)	(-0.20)	(1.13)	(-0.85)	(-0.27)
$Idiosync^{Roundtrip,M>10y}_{i,t-1}$	-0.0060	0.0004	0.0015	0.0042*	0.0012
	(-1.88)	(0.45)	(1.06)	(2.09)	(0.39)
$CreditCall^{Roundtrip,M>10y}_{i,t-1}$	0.0020	-0.0003	0.0018	0.0073*	-0.0030
	(1.11)	(-0.12)	(0.83)	(2.01)	(-1.21)
$Maturity^{Roundtrip,M>10y}_{i,t-1}$	-0.0040	-0.0011	0.0015	0.0084**	0.0000
	(-1.09)	(-0.69)	(0.68)	(2.80)	(0.02)
$TreasConv^{Roundtrip,M>10y}_{i,t-1}$	0.0104*	0.0023	0.0014	-0.0065	0.0022
	(2.06)	(0.46)	(0.60)	(-1.25)	(1.19)
R^2	0.00	0.01	0.01	0.00	0.00
N	327,389	328,629	276,026	242,140	1,141,978

Standard errors clustered by bond issuer.
* (p<0.05), ** (p<0.01), *** (p<0.001)

Table 1.16: Order flow regressed on four lagged components of bond spread $Idiosync_{i,t-1}$, $CreditCall_{i,t-1}$, $Maturity_{i,t-1}$ and $TreasConv_{i,t-1}$ (see table 1.8 for details on the components), broken down by One-way / partial Rountrip order flow and bond **residual maturity** T being below or above 10 years. Controls include the dummy $OneWay_{i,t}$ for one-way order flow, and where applicable the dummy for bond residual maturity T being less than 10 years, and its interaction with $OneWay_{i,t}$. Additional controls are 10 lags of order flow, 10 lags of bond price changes, 3 lags of stock return and lagged 10 years Treasury yield changes. To save space **only the coefficients for One-way order flow are shown.**

	Opaque	Pre-Crisis	Crisis	Post-Crisis	Dodd-Frank
$Idiosync_{i,t-1}^{OneWay,T\leq 10y}$	-0.0843*	-0.0497	-0.1171***	-0.0259	-0.0005
	(-2.60)	(-0.73)	(-8.00)	(-1.92)	(-0.04)
$CreditCall_{i,t-1}^{OneWay,T\leq 10y}$	-0.1459***	-0.4025***	-0.1299***	-0.0957***	-0.0711***
	(-4.05)	(-6.03)	(-8.12)	(-3.83)	(-3.52)
$Maturity_{i,t-1}^{OneWay,T\leq 10y}$	0.0316	-0.0025	0.0293	0.0363	0.0228*
	(0.77)	(-0.05)	(1.70)	(1.67)	(2.37)
$TreasConv_{i,t-1}^{OneWay,T\leq 10y}$	-0.0534	-0.1784***	-0.1072***	-0.0994***	-0.0724***
	(-1.57)	(-3.58)	(-5.07)	(-4.12)	(-3.93)
$Idiosync_{i,t-1}^{OneWay,T>10y}$	-0.0231*	0.0136*	0.0066	0.0096	0.0217**
	(-2.22)	(2.27)	(0.62)	(0.79)	(3.26)
$CreditCall_{i,t-1}^{OneWay,T>10y}$	0.0081	0.0243	0.0017	0.0245	0.0251**
	(1.09)	(1.68)	(0.22)	(1.73)	(3.19)
$Maturity_{i,t-1}^{OneWay,T>10y}$	-0.0201**	0.0332***	0.0139	0.0006	0.0026
	(-2.61)	(3.65)	(1.31)	(0.04)	(0.35)
$TreasConv_{i,t-1}^{OneWay,T>10y}$	-0.0835**	-0.0684	-0.0765***	0.0687*	-0.1135***
	(-2.61)	(-1.01)	(-5.69)	(2.23)	(-7.20)
R^2	0.00	0.01	0.01	0.00	0.00
N	327,389	328,629	276,026	242,140	1,141,978

Clustered by firm, month, years to maturity

* (p<0.05), ** (p<0.01), *** (p<0.001)

Table 1.17: Order flow regressed on four lagged components of bond spread $Idiosync_{i,t-1}$, $CreditCall_{i,t-1}$, $Maturity_{i,t-1}$ and $TreasConv_{i,t-1}$ (see table 1.8 for details on the components), broken down by One-way / partial Rountrip order flow and bond **residual maturity** T being below or above 10 years. Controls include the dummy $OneWay_{i,t}$ for one-way order flow, and where applicable the dummy for bond residual maturity T being less than 10 years, and its interaction with $OneWay_{i,t}$. Additional controls are 10 lags of order flow, 10 lags of bond price changes, 3 lags of stock return and lagged 10 years Treasury yield changes. To save space **only the coefficients for Partial Roundtrip order flow are shown.**

	Opaque	Pre-Crisis	Crisis	Post-Crisis	Dodd-Frank
$Idiosync_{i,t-1}^{Roundtrip,T\leq 10y}$	-0.0004	-0.0048	-0.0034	0.0080	0.0125**
	(-0.04)	(-0.73)	(-0.99)	(1.93)	(2.64)
$CreditCall_{i,t-1}^{Roundtrip,T\leq 10y}$	-0.0294**	-0.0522**	-0.0094	0.0065	-0.0001
	(-2.71)	(-2.72)	(-1.82)	(0.70)	(-0.02)
$Maturity_{i,t-1}^{Roundtrip,T\leq 10y}$	0.0289*	0.0023	0.0013	0.0094	0.0053
	(2.14)	(0.15)	(0.25)	(1.26)	(1.86)
$TreasConv_{i,t-1}^{Roundtrip,T\leq 10y}$	-0.0072	-0.0090	0.0027	-0.0100	0.0009
	(-0.58)	(-0.66)	(0.54)	(-1.12)	(0.23)
$CheapInSim_{i,t-1}^{Roundtrip,T>10y}$	-0.0063	0.0000	0.0013	0.0035*	0.0005
	(-1.95)	(0.01)	(0.85)	(2.11)	(0.17)
$CreditCall_{i,t-1}^{Roundtrip,T>10y}$	0.0020	-0.0014	0.0017	0.0059	-0.0041
	(1.16)	(-0.61)	(0.72)	(1.72)	(-1.53)
$Maturity_{i,t-1}^{Roundtrip,T>10y}$	-0.0040	-0.0016	0.0012	0.0075*	0.0003
	(-1.09)	(-1.02)	(0.51)	(2.43)	(0.14)
$TreasConv_{i,t-1}^{Roundtrip,T>10y}$	0.0023	0.0246	0.0067	-0.0090	-0.0010
	(0.24)	(1.66)	(1.65)	(-1.05)	(-0.20)
R^2	0.00	0.01	0.01	0.00	0.00
N	327,389	328,629	276,026	242,140	1,141,978

Clustered by firm, month, years to maturity

* (p<0.05), ** (p<0.01), *** (p<0.001)

1.12 Appendix 3: Repos, reverse repos, long and short positions

Repos or collateralized loans are ways for an investor with limited capital to fund the acquisition of an asset: the investor borrows cash up to some horizon and gives the asset as collateral to the lender. In general the lender does not finance the full amount of the purchase, so that the investor has to complete the purchase amount with his capital. This ensures that if the borrower fails, the lender can sell the collateral and recover a high fraction of his loan even if the asset price has decreased. The ratio between the amount of investor's own capital required to fund the asset and the market value of the asset is called the *haircut*. Haircuts are thus financing constraints.

Reverse repos go in the opposite direction: instead of borrowing cash, the investor borrows a security and gives cash as collateral to the security lender. Similarly to repos, the security lender may require an amount of cash collateral that is higher than the market value of the asset, so that the investor has to find cash, typically from his own funds, to finance the cash collateral. For reverse repos I call haircut the ratio of the value of the cash collateral that is funded on the investor's own funds, to the market value of the security when the loan is made.

Repos and reverse repos are mirror images of each other: a repo for the investor is a reverse repo from the perspective of his (cash) lender. Positive haircuts from the perspective of one agent correspond to negative from the perspective of the other agent. Therefore the definition of haircuts depends on which agent one considers: here I define haircuts from the perspective of Primary Dealers.

In a reverse repo, the investor can sell the security he has borrowed as long as he comes back with an identical security at the expiration of the contract. In the interim period, the security borrowed and sold becomes a liability. The investor who sells the security makes a loss if he repurchases the asset at a higher price, and gains if he repurchases at a lower price. Reverse repos are thus used to implement short-selling.

In the definition I adopt here, arbitrage involves long and short position in correlated assets. A long position, consists in buying and holding an asset for some period of time and being exposed to the risk of a low payoff. A short position is the short-selling of a security as described above, so that the investor carrying a short position is exposed to the risk of a high payoff. When securities on the long and short legs of the strategy comove, the risks associated with each leg are partially or completely offset.

Implementing a long-short strategy naturally leads to using a combination of a repo to fund the long leg and a reverse repo to fund the short leg, as illustrated on figure 1.11. An investor with limited capital is willing to borrow a security: he has to finance the cash collateral, which is a loan to the security lender. To do this he benefits from an unsecured loan L, presumably from a clearing agent who knows his positions and thus his ability to reimburse well. Then the investor receives the security borrowed, for a lower value than the loan he grants reflecting the haircut. By selling the security borrowed, the investor gets cash, while he still owes the security at expiration of the security lending contract which now appear as a liability. With the cash he could purchase the asset on the long side of the strategy: then the unsecured lender may require to have it as collateral for the initial loan L, which would improve its terms; or he could equivalently redeem the loan L with the cash, and enter a separate repo to implement the long side. Both imply a repo and a reverse repo. In the following subsection I show that Primary Dealers appeared to implement their strategies in this way.

Such implementation implies distinct financial contraints for each leg of the long-short strategy, as each leg involves potentially different haircuts.

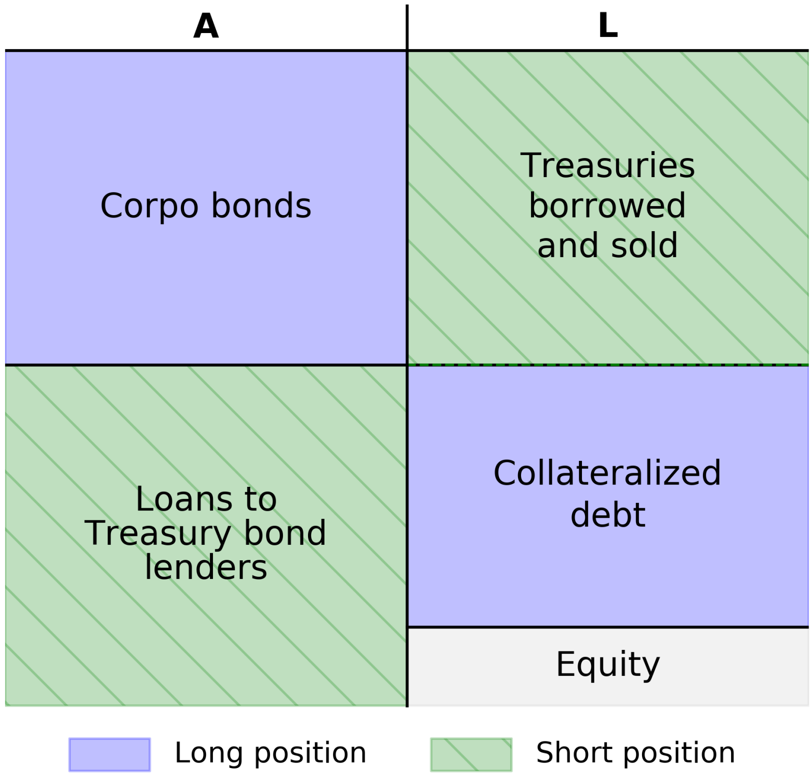

Figure 1.11: **Simplified balance sheet of an arbitrageur.** The long position (blue, without hatches), in corporate bonds for illustrative purpose, is funded through collateralized debt or repo contracts. The short position (green, with hatches), in Treasuries for illustration, is funded through a reverse repo: Treasury securities are borrowed from security lenders, the arbitrageur granting a loan to the security lender as collateral.

Chapter 2

Imperfect Competition, Dynamic Trading and Forward Contracts

Abstract

I study why financial institutions trade forward and future contracts on assets they could buy or sell directly. I provide a dynamic trading model in which this occurs because of 1) imperfect competition and 2) uncertainty about future customer trades. Under imperfect competition, risk averse traders realize gains from trading inventory imbalances slowly, leaving them differentially exposed to a supply shock: sellers fear customers sells that depress prices, buyers fear customer buys that increase prices. Opposite exposure to the supply shock implies gains from trading the risk through forward contracts: in equilibrium sellers of the asset sell forwards to buyers, and risk sharing in the asset is slowed down. The cost of slower risk sharing is compensated by the benefit of more certain future trading surplus. Traders are more willing to create inventory imbalance with forward contracts, leading to tighter spreads and/or higher trading volume in fragmented markets.

2.1 Introduction

Forward, futures and swap contracts are pervasive in fixed income, commodities and currency markets. These derivative contracts allow traders to buy or sell at a future date an underlying asset at a pre-agreed price.[1] In some cases, these contracts arise because the underlying asset does not yet exist at the time of the contract, while traders want to hedge against the price risk.[2] But in many cases, these contracts arise with an underlying asset that already exists: financial institutions often hedge holdings of a tradeable security with corresponding forwards or swaps, or purchase forwards or swaps instead of purchasing the underlying[3]. To offload the risk, they could sell their holdings directly and invest the proceeds in riskless cash. Why do traders trade forwards when the underlying asset is available for trade? What are the effects on inventory holdings and traders' welfare?

To answer these questions, in this paper I provide a dynamic trading model where forward contracts endogenously emerge. Forward contracts' role is to insure against supply or demand shocks on the underlying asset that impact the price in the course of trading: under imperfect competition, trading is slow and buyers of the asset fear a high price in the future, sellers fear a low price. In equilibrium sellers of the underlying asset sell forwards to buyers and delay selling of the underlying. Underlying risk sharing is impeded, a cost that is more than compensated for traders by the benefit of better management of dynamic trading. Under perfect competition, sellers sell all their inventories immediately and forwards are not traded.

I study a model in which risk averse traders, called *dealers* for concreteness,[4] can trade a risky asset at two dates 0 and 1. Dealers differ by their initial inventories of the asset, and have identical preferences and information. At each trading date, dealers meet in a centralized market. At date 1, some exogenous customers post an inelastic quantity that is unknown at date 0 and independent from the asset payoff. Date 1 price is low when customers sell, and vice versa.

To show why forward contracts endogenously emerge under imperfect competition, I provide a preliminary result: under imperfect competition, trading is slow because of a static effect and of a dynamic effect. The static effect is classical: at each date dealers care about the impacts of their demand schedules on the contemporaneous asset price, which leads them to realize each period only a fraction of remaining gains from trade. The dynamic effect is that at date 0, dealers care about the impact of their trade on date 1 price. By selling more at date 0, sellers reduce their need to sell at date 1, which raises date 1 price and thus increases the

[1]Forward contracts are traded bilaterally, futures are listed. Swaps are portfolios of forward contracts (*cf. e.g.* Hull 2003, Stulz 2004).

[2]*E.g.* a farmer sells his/her crop forward before it has grown, or a company buys domestic currency forward because it anticipates future cash inflows in foreign currency from sales abroad.

[3]McDonald and Paulson (2015) notice that in 2007, AIG not only bought exposures to corporate bonds credit risk through credit default swaps, but also to the interest rate component through interest rate swaps. They could have purchased corporate bonds directly instead.

[4]In the US, dealers are economic agents who trade in financial markets for their own account as a regular business. The largest are within large banking or asset management groups. Their activity is regulated by the SEC.

profitability of selling one unit at date 1: therefore selling more at date 0 implies an increased opportunity cost of not delaying trade to date 1. A symmetric mechanism operates for buyers. As reviewed shortly, this second effect drives the impact of forward contracts on the pace of trading. By contrast, under perfect competition, none of the two effects play: therefore all gains from trade are realized and dealers equate their inventories at date 0.

To get the main result of the paper, I first introduce the risk that forward contracts allow to share: at date 1, exogenous customers buy or sell a quantity that is unknown at date 0, which makes date 1 price increase or decrease. Under imperfect competition, buyers and sellers have opposite exposure to this risk: sellers at date 0 are still willing to sell at date 1 and they fear that customers will simultaneously sell, which would lower the date 1 price; symmetrically buyers fear that customers buy at date 1. Thus buyers and sellers are exposed in an opposite way to the supply shock, *i.e.* there are gains from trading this risk.

Then I introduce forward contracts maturing at date 1, thus indexed on the supply shock, as a means of trading exposures to the supply shock. Forward contracts by definition pay off the difference between the realized price of the asset, and a pre-agreed forward price. In my model, there are two sources of variation between date 0 and date 1 price: information on dealers' final payoff that arrive at date 1, and the supply shock. I study two contracts separately. The first one, contract a, has a payoff linear in the supply shock, and can be implemented as a forward if the price moves between date 0 and date 1 only because of the supply shock. The second contract, labelled b, has a payoff linear in both the supply shock and date 1 information and is implemented as a forward even with date 1 information about the asset payoff and is approximated by contract a if date 1 information is not too high; the latter condition can be viewed as a shortcut for date 1 being not too far away in calendar time. I refer to both as forwards for simplicity. In equilibrium, sellers of the underlying asset sell forward contracts, so that they gain if the price is lower than expected, and buyers take the other side.

The second important result of this is that introducing forward contracts slows down and decrease risk sharing in the underlying asset. To understand this, I show that uncertainty about the supply shock accelerates trading in the underlying asset, because it decreases the dynamic effect of imperfect competition: unexpected customer buys or sells make date 1 price move in unexpected direction, making the value of managing date 1 price lower. uncertainty about the supply shock in turn mitigates incentives to postpone trade to date 1. As forward contracts hedge against adverse movement of prices, they allow dealers to behave as if there was no uncertainty about the supply shock, which slows down trading. In addition, as postponing one unit of date 0 asset trade translates into less than one unit of date 1 trade, the total quantity traded is decreased with forward contracts. The effect is also present for contract b if the variance of date 1 information is not too high. Thus instead of equating buyers' and sellers' marginal utilities of asset holdings, forward contracts make them differ even more at date 1.

The third important result is that in spite of the cost of decreased underlying risk sharing, dealers are better off with contract a: this is because of the compensating benefit of making date 1 surplus more certain and more controllable.

Finally, I observe that the value dealers derive from trading increases with the square of initial inventory differences, *i.e.* with gains from trade. This gives incentives for dealers to trade bilaterally to create inventory imbalances and capture the trading rent. I run a simple exercise to show that derivatives increase the willingness of dealers to trade over-the-counter, since trading over-the-counter allow to create interdealer gains from trade.

This setting with forward contracts resembles the situation before the 2007-2009 crisis, with large dealers holding very large asset and derivative positions. Dealers held large bond positions hedged by derivatives, credit default swaps in particular which can be viewed as portfolios of forward contracts. After the crisis, several regulations have impeded holding of risk for large dealers through various balance sheet costs on bank-affiliated dealers: the setting without forward contracts is reminiscent of this situation.

Consistently with the empirical findings that after the crisis dealers hold positions for a shorter period of time (Dick-Nielsen and Rossi 2018), dealers in my model are quicker to revert inventory imbalances. However, my model suggests that the risk bearing capacity of dealers does not only serve to bring more liquidity, but is also useful for rent extraction.

Literature review. This paper first contributes to a growing literature on the economics of derivatives. Allaz and Vila (1993) also study how forward contracts emerge under imperfect competition: their main point is that forward emerge even without risk, and forward transactions make producers worse off. With a very different setting I find opposite results. Biais et al. (2016) and Biais et al. (2019) study how derivatives impede risk management, and how margin requirements can mitigate the problem. Biais et al. (2019) explore how moral hazard affects risk sharing through derivatives. Gains from trading risk arise from differences in preferences in a competitive setting, while I make gains from trading derivatives through imperfect competition with identical preferences. Oehmke and Zawadowski (2015, 2016) show that when swaps have exogenously low transaction costs, the natural holder of derivatives are those with short-term horizon, while long term investors hold the underlying. In my model transaction costs are endogenously low, because forwards have a shorter maturity than the underlying.

This paper also connects to the literature on dynamic trading with imperfectly competitive double auctions. Vayanos (1999), Du and Zhu (2017) and Rostek and Weretka (2015) study dynamic trading strategies without forward contracts. Duffie and Zhu (2017) and Antill and Duffie (2018) explore the ability of size discovery mechanisms to overcome the inefficiency. My paper is to my knowledge the first to make forward contract emerge in this context.

The paper is organized as follows. Section 2.2 presents the setting and solves the competitive benchmark. Section 2.3 solves the imperfect competition equilibrium without forward contracts, and derives the result that uncertainty on the supply shock accelerates trading. Section 2.4 highlights the gains from trading the risk on the supply shock. Section 2.5 introduces forward contracts and solves for the equilibrium. Section 2.6 gives dealer welfare with and without derivatives, and compares them. Section 2.7 gives an analysis of spreads and trading volumes in OTC markets. Section 2.8 concludes.

2.2 Setting and competitive benchmark

2.2.1 Setting

There are four dates $t = 0, 1, 2$. There is one risky asset that pays off at $t = 2$ an ex ante unknown amount v per unit. At each date, before any action takes place, a public signal ϵ_t is released: ϵ_1 and ϵ_2 are independent and normally distributed with mean 0 and respective variances σ_1^2 and σ_2^2. Thus $v = v_0 + \epsilon_1 + \epsilon_2$ and we denote v_t the expectation of v conditional on information released at t. There is also a riskless asset (cash) that can be purchased or sold without constraint by a perfectly elastic supplier. We normalize its gross return to 1.

There are two types of traders $i = 1, 2$, which I call *dealers* for concreteness. Dealers of class i maximize the expected utility of their terminal wealth W_i described below. Each class contains $N \geq 2$ dealers[5]. Their utility is negative exponential (CARA), with risk aversion parameter γ for both classes. Dealers of class i all start with initial inventory $I_{i,0}$ of the risky asset at date 0: interdealer gains from trade arise from inventory differences. Dealers are all forward-looking and fully rational: in particular, they perfectly anticipate at date 0 the date 1 equilibrium and adjust their actions accordingly. When considering dealers of class i, I use the notation $-i$ to refer to dealers of the other class.

At date 0, dealers can meet in a centralized market, where they post demand schedules. A walrasian auctioneer computes the equilibrium price p_0^c (competitive case) or p_0^* (imperfectly competitive case) that clears the market. All dealers of class i post the same demand schedule $q_{i,0}(p_0)$ and therefore purchase the same equilibrium quantity $q_{i,0}^* \equiv q_{i,0}^*(p_0^*)$ (replace the star by a c for the competitive market). This holds whether the market is perfectly competitive or not, and I further specify the equilibrium concept when dealers are strategic in section 2.3. Symmetry of demand schedules within a class of dealers is an equilibrium outcome when market are competitive, while when traders are strategic, such equilibria exist and are a natural focus of analysis. The market clearing condition at date 0 is thus $Nq_{1,0}^*(p_0^*) + Nq_{2,0}^*(p_0^*) = 0$, *i.e.*

$$q_{1,0}^*(p_0^*) + q_{2,0}^*(p_0^*) = 0 \tag{2.2.1}$$

[5]This makes the total number of traders in each market greater than 3, a necessary condition to have equilibrium in linear strategies. When there are only two traders, Du and Zhu (2017) show existence of equilibria in *non-linear* strategies.

At date 1, the market re-opens, and an external infinitely risk averse customer has a liquidity shock and is willing to sell Q units of the security (thus when $Q > 0$, the customer is willing to sell and vice versa).[6] Again the market is walrasian, with market clearing price p_1^*, and the assumption about date 1 competitiveness of dealers is naturally consistent with that of date 0. dealers arrive in the date 1 market with inventories $I_{i,1} = I_{i,0} + q_{i,0}$, where $q_{i,0}$ is the quantity traded at date 0. Market clearing at $t = 1$ thus writes $Nq_{1,1}^*(p_1^*) + Nq_{2,1}^*(p_1^*) = Q$, *i.e.*

$$q_{1,1}^*(p_1^*) + q_{2,1}^*(p_1^*) = \frac{Q}{N} \tag{2.2.2}$$

Dealers of both classes do not know the value of Q at date 0, and it is publicly revealed simultaneously with ϵ_1 at date 1 before the market opens. A public noisy signal on Q is released at date 0 before the market opens (thus there is no information asymmetry): all dealers share the common belief at date 0 that Q is normally distributed with mean $\mathbb{E}_0[Q]$ and variance $N^2\sigma_q^2$. In addition, we assume that Q is independent of ϵ_1 and ϵ_2, and that this is common knowledge. Independence of Q and ϵ_t means that Q is a pure private value or liquidity shock: this captures the real life feature that an investor may sometime need cash when other market participants don't, *e.g.* a mutual fund or a life insurer facing indiosyncratic withdrawals, or that the investor has got news on its future cash needs and adjusts his portfolio (maturity, liquidity, ...) accordingly.

With initial inventory $I_{i,0}$, quantities $q_{i,0}$ and $q_{i,1}$ purchased at $t = 0$ and $t = 1$ at respective prices p_0 and p_1, the terminal wealth of class i traders is

$$W_i = I_{i,0}v + q_{i,0}(v - p_0) + q_{i,1}(v - p_1) \tag{2.2.3}$$

Equilibria are solved by backward induction, consistently with dealer full rationality. The appropriate equilibrium concepts are defined in the relevant sections.

2.2.2 Competitive equilibrium

To give the intuitions which and where derivative contracts may be used in equilibrium, I first solve the date 1 competitive equilibrium and compute the associated equilibrium utility which is a function of Q. Then I derive intuitions on the impact of Q and the associated risk that carry over for the imperfect competition case, which is formally very similar in this respect.

I look for competitive equilibria defined as sets of demand schedules $(q_{i,0}^*(p_0), q_{i,1}^*(p_1))$ $(i = 1, 2)$ and equilibrium prices p_0^c, p_1^c such that:

1. All traders are price-takers;

2. For each class of trader i, date 1 demand schedules $q_{i,1}^*(p_1)$ maximize their expecetd utility of terminal wealth W_i given information available at date 1;

[6] The analysis is isomorphic to that of a liquidity shock of size Q/N that hits each trader of a given class i (e.g. a customer who has traded only with traders of class i), although the equilibrium allocations and prices differ in the imperfect competition setting.

3. For each class of trader i, date 0 demand schedules $q^*_{i,0}(p_0)$ maximize their expected utility of terminal wealth W_i given information available at date 0 and anticipated equilibrium outcomes at date 1;

4. The market clearing conditions (3.2.2) and (3.2.3) hold.

Again I make the slight abuse of notation that that symmetry of traders of class i is included in the definition, while it is in fact an equilibrium outcome. I look for equilibria by backward induction.

Date 1 equilibrium

Dealers of class i maximize over $q_{i,1}$ the expected utility of trader i is, since the only uncertainty is on the normally distributed variable ϵ_2,

$$\mathbb{E}_1\left[-e^{-\gamma W_i}\right] = -\exp\left\{-\gamma \widetilde{W}_{i,1}\right\}$$

where

$$\widetilde{W}_{i,1} = I_{i,0}v_1 + q_{i,0}(v_1 - p_0) + q_{i,1}(v_1 - p_1) - \frac{\gamma}{2}\sigma_2^2(I_{i,1} + q_{i,1})^2 \tag{2.2.4}$$

As the utility function is increasing, all happens as if dealers of class i maximized the certainty equivalent $\widetilde{W}_{i,1}$ of their wealth. From the first order condition of this maximization problem one easily derives the optimal competitive demand schedule:

$$q^c_{i,1}(p_1) = \frac{v_1 - p_1}{\gamma\sigma_2^2} - I_{i,1} \tag{2.2.5}$$

Demand increases when the expected terminal payoff v_1 is larger with respect to the purchase price p_1, when dealers' risk aversion γ is low, and when the terminal payoff variance σ_2^2 is low. Plugging optimal demands into the market clearing condition (3.2.3), it is straightforward to derive the equilibrium price

$$p^c_1 = v_1 - \gamma\sigma_2^2\frac{Q^*_c}{2} \tag{2.2.6}$$
$$\text{with} \quad Q^*_c = I_{1,1} + I_{2,1} + \frac{Q}{N}$$

Notice that with date 0 market clearing condition (3.2.2), one has $I_{1,1} + I_{2,1} = I_{1,0} + I_{2,0}$. The equilibrium price is therefore equal to the expected value of the asset minus a risk premium that increases if risk aversions increase, if the uncertainty σ_2^2 over the asset terminal payoff v increases, and if the quantity held by traders after date 1 trade increases. In particular, if customers are net sellers ($Q > 0$), then the equilibrium price decreases and vice versa, which is intuitive.

Plugging equilibrium price (2.2.6) into optimal demand schedule (2.2.5), one gets the equilibrium quantities purchased and post-trade inventories held by traders of class i (denoting the other class of traders by $-i$):

$$I_{i,1} + q^c_{i,1} = \frac{Q^*_c}{2} \tag{2.2.7}$$

$$q^c_{i,1} = \frac{Q^*_c}{2} - I_{i,1} = \frac{I_{-i,1} - I_{i,1}}{2} + \frac{Q}{2N} \tag{2.2.8}$$

Equilibrium post trade inventories show that the total inventory in the market is evenly split across dealers.

Dealer valuation of the *surplus* of date 1 trade

After date 1 trade, from 2.2.4, dealer i certainty equivalent of wealth can be decomposed as

$$\widetilde{W}^c_{i,1} = I_{i,0}v_1 + q_{i,0}(v_1 - p_0) - \frac{\gamma\sigma_2^2}{2}(I_{i,0} + q_{i,0})^2 + S^c_1 \tag{2.2.9}$$

$$\text{with } S^c_1 = q^c_{i,1}(v_1 - p^c_1) - \left(\frac{\gamma\sigma_2^2}{2}(I_{i,1} + q^c_{i,1})^2 - \frac{\gamma\sigma_2^2}{2}(I_{i,1})^2\right) \tag{2.2.10}$$

The first terms are the classical mean-variance value of date 0 inventory position after date 0 trade, while S^c_1 is the net surplus of date 1 transaction. S^c_1 is the sum of two terms: $q^c_{i,1}(v_1 - p^c_1)$ is the expected payoff from the trade, while the difference in bracket is the impact of the change in dealer i's inventory position on her risk holding cost.

Plugging equilibrium price 2.2.6, inventory 2.2.7 and quantity traded 2.2.8 into the expression for S_1 leads to

$$S^c_1 = \frac{\gamma\sigma_2^2}{2}\left(q^*_{1,1}\right)^2 \tag{2.2.11}$$

$$= \frac{\gamma\sigma_2^2}{2}\left(\frac{Q^*_c}{2} - I_{i,0} - q_{i,0}\right)^2 \tag{2.2.12}$$

S^c_1 is proportional to the square of the quantity traded, which also holds under imperfect competition. It can also be influenced by date 0 trading choice $q_{i,0}$. Finally, it is quadratic in Q: dealer i cares both about the price and the quantity traded, and both are impacted by Q.

Plugging 2.2.12 into 2.2.10 and taking the certainty equivalent with respect to both ϵ_1 and Q using lemma 11 in the appendix, one gets the date 0 certainty equivalent of wealth for dealer i:

$$\widetilde{W}_{i,0} = I_{i,0}v_0 + q_{i,0}(v_0 - p_0) - \frac{\gamma}{2}(\sigma_1^2 + \sigma_2^2)(I_{i,0} + q_{i,0})^2$$

$$+ \frac{\gamma\sigma_2^2}{2}\frac{1}{1 + x_c}\left(\mathbb{E}_0[q^c_{i,1}]\right)^2 + \frac{1}{2\gamma}\ln(1 + x_c) \tag{2.2.13}$$

$$\text{with } q^c_{i,1} = \frac{Q^*_c}{2} - I_{i,0} - q_{i,0} \qquad \text{and} \qquad x_c = \bar{\gamma}_c^2\sigma_2^2\sigma_q^2$$

The first term is the expected payoff from the initial inventory $I_{i,0}$, the second term if the expected profit from date 0 trade, the third term is the cost of risk associated with holding asset inventory until maturity. These terms reflect a classical mean-variance trade-off.

The fourth term is new and reflects the opportunity of short-term capital gains for dealer i: it shows that y maximizing the difference between his post-trade inventory $Q^*/2$ and his initial position, dealer i maximizes the surplus from the transaction (and thus the share he gets from it). Therefore, by choosing his demand $q_{i,0}$, a class i dealer faces a potential tradeoff between optimizing the Hold-To-Maturity (HTM) component of his payoff, and the short-term component.

However, the fourth term decreases when the uncertainty σ_q^2 over the supply shock increases: when uncertainty about the liquidity shock Q is large, there is a high probability that Q takes extreme values, which 1) makes date 1 surplus very large irrespective of the direction of the liquidity shock and 2) increase the probability that optimization was in the wrong direction (e.g. large ex ante sells expecting large customer sells, but large customer buys realize). Therefore the marginal payoff of optimizing is lower, and indeed I show in the next subsection that the weight of this term in traders' certainty equivalent of wealth and optimal demand decreases.

The fifth term $1/2\gamma \times \ln(1 + x_c)$ also comes from the quadratic dependence of interim utility on the supply shock, but is independent of trader i's demand. It increases with uncertainty about the supply shock σ_q^2: when it is more likely that the supply shock takes more extreme values, it is more likely that the date 1 surplus is higher, which makes trader i better off. It does not intervene in dealer i's optimization problem however.

Competitive equilibrium: date 0

Optimal demand schedules. The optimal demand $q_{i,0}^c(p_0)$ maximizes the certainty equivalent of wealth (2.2.13). The problem is solved by the unique solution to the following first order condition:[7]

$$v_0 - p_0 = \gamma(\sigma_1^2 + \sigma_2^2)\left(I_{i,0} + q_{i,0}^c(p_0)\right) - \frac{\bar{\gamma}_c\sigma_2^2}{1+x_c}\left(\mathbb{E}_0[Q_c^*] - \frac{\gamma_i}{\bar{\gamma}_c}\left(I_{i,0} + q_{i,0}^c(p_0)\right)\right)$$
$$x_c = \bar{\gamma}_c^2\sigma_2^2\sigma_q^2$$

Rearranging leads to

$$q_{i,0}^c(p_0) = \frac{v_0 - p_0}{\gamma(\sigma_1^2 + \delta_c\sigma_2^2)} - \frac{\sigma_2^2}{\sigma_1^2 + \delta_c\sigma_2^2}\frac{1}{1+x_c}\mathbb{E}_0\left[\frac{Q_c^*}{2}\right] - I_{i,0} \qquad (2.2.14)$$

$$\text{with } \delta_c = \frac{x_c}{1+x_c} \in [0,1)$$

The optimal demand is the sum of a quasi hold-to-maturity demand (first term), analogous to the two periods demand (2.2.5), and the short term profit demand,

[7]It is straightforward to check that the problem is strictly concave, as $1/(1 + x_{2c}) < 1$.

that appears as an arbitrage demand (second term). Both terms are impacted by the uncertainty about the liquidity shock σ_q^2.

Consider the second term first. The term inside the expectation is the equilibrium date 0 inventory that trader i expects to carry from date 1 to date 2. Suppose for simplicity that $\sigma_1^2 = \sigma_2^2$ and that the variance of the supply shock is zero: then the term boils down to minus the date 1 equilibrium inventory. This means that at date 0, trader i is willing to sell the asset inventory he expects to repurchase at date 1, likely at a profit: it thus reflects an opportunity to realize a short-term capital gain.

However when the uncertainty about the supply shock Q increases, the coefficient in front of the expected inventory decreases: if $\sigma_1^2 = \sigma_2^2$, then the coefficient equals $1/(1+2x_c)$, which reflect the fact that the reversal of the date 0 selling is uncertain. Then trader 1 may well end up holding too little or too much inventory to maturity with respect to the optimal quantities $q_{i,0}, q_{i,1}$ when the liquidity shock is known with certainty.

Uncertainty about the reversal leads trader i to put less weight on the arbitrage term, while he becomes more conservative on the Hold-to-Maturity component of his demand. As reversal of date 0 selling is more uncertain, holding the excess or deficit of asset inventory to maturity is more likely: this translates into a variance that becomes closer to $\sigma_1^2 + \sigma_2^2$, the variance of v from date 0.

Uncertainty about Q and trader horizon. When the liquidity shock is known ex ante for sure ($\sigma_q^2 = 0 \Rightarrow x_c = 0$), the factor δ_c is zero: thus the effective holding horizon is $t = 1$, as only $\gamma\sigma_1^2$ appears in the denominator of the first term. Therefore trader i has optimally short horizon whenever he is sure to be able to reverse his position in the short term. Conversely, when a trader is not sure to be able to reverse his position in the short run, he puts some probability of holding it for a longer time and therefore has a longer horizon.

As σ_q^2 increases, the coefficient δ_c in front of σ_2^2 in (2.2.14) increases and becomes closer to 1 as σ_q^2 becomes arbitrarily large. This is as if the trader was increasingly concerned about the terminal payoff at date 3, rather than about the short term return.

Competitive equilibrium. The equilibrium prices and quantities are stated in the following proposition, proven in the appendix.

Proposition 1. *The equilibrium price is*

$$p_0^c = v_0 - \gamma(\sigma_1^2 + \sigma_2^2)\frac{I_{1,0} + I_{2,0}}{2} - \frac{\gamma\sigma_2^2}{1 + x_c}\,\mathbb{E}_0\left[\frac{Q}{2N}\right] \tag{2.2.15}$$

The risk premium is the sum of a hold-to-maturity component (second term), and of an short-term arbitrage component (third term) that is proportional to the expected date 1 liquidity shock: the price is higher when customer purchases are expected ($\mathbb{E}_0[Q] < 0$*) and vice versa. The sensitivity to the date 1 liquidity shock decreases as uncertainty about it increases.*

Equilibrium trade and post-trade inventories are

$$q^c_{i,0} = \frac{I_{-i,0} - I_{i,0}}{2} \tag{2.2.16}$$

$$I_{i,0} + q^c_{i,0} = \frac{I_{1,0} + I_{2,0}}{2} \tag{2.2.17}$$

$$q^c_{i,1} = \frac{Q}{2N} \tag{2.2.18}$$

Risk sharing is Pareto optimal.

Inventories are equalized right after date 0 trade: all interdealer gains from trade are realized at date 0. By market clearing, as date 1 customers are price inelastic, the short term capital gain demand has no impact on the quantities traded and all effect goes in the price.

2.3 Equilibrium with imperfect competition and uncertainty about future customer demand

In this section I derive the equilibrium when dealers are imperfectly competitive in that they manage the price impact of their trades (*cf.* subsection 3.3.1), but without derivative contracts allowed.

Dealers then reach only progressively their final inventory positions. In fact, at date 0 imperfect competition plays both in a static and in a dynamic way: dealers manage the impact of their trade on the contemporaneous price (static effect) and on subsequent price (dynamic effect).

Crucially, I show in subsection 2.3.3 that when uncertainty about Q is higher, dealers converge more quickly to the final allocation, and eventually get closer to the efficient allocation.

2.3.1 Date 1

In this section, I look for Nash equilibria in demand schedules in the date 1 market. As there is no asymmetric information, there is an equilibrium multiplicity problem (Klemperer and Meyer 1989). I use the usual trembling-hand stability criterion to select a unique equilibrium (*cf.* Vayanos 1999).

At date 1, the expected utility to be maximized by trader k in class i is given by (2.2.4). By contrast with competitive markets, traders now take the impact of their demand on the equilibrium price into account: they conjecture the equilibrium residual demand curve that is the sum of all other traders' demand curves. For a given quantity $q_{i,1}$ demanded by trader i, this residual demand curve implies an equilibrium price p_1, and a marginal increase in the quantity demanded by trader i implies a marginal price impact $\partial p_1/\partial q_{k,i,1}$. Differentiating the certainty equivalent of wealth (2.2.4), trader k' first order condition is:

$$v_1 - p_1 - q_{k,i,1}\frac{\partial p_1}{\partial q_{k,i,1}} = \gamma\sigma_2^2(I_{i,1} + q_{k,i,1})$$

Following the literature on market with Nash equilibria in demand schedules,[8] I conjecture linear strategies in equilibrium: there optimal strategies are linear whenever other market participants use linear strategies, but individual strategies are not constrained to be linear. Thus trader k in class i expects to face a linear residual demand curve of conjectured slope $1/\lambda_{k,i,1}$, so that $\partial p_1/\partial q_{i,1} = \lambda_{k,i,1}$. To ease notation I slightly anticipate on the equilibrium result that all traders k within class i follow symmetric strategies, thus I drop the k subscript, so that her optimal demand is:

$$q^*_{k,i,1}(p_1, \lambda_{i,1}) = \frac{v_1 - p_1}{\lambda_{i,1} + \gamma\sigma_2^2} - \frac{\gamma\sigma_2^2}{\lambda_{i,1} + \gamma\sigma_2^2}I_{i,1} \tag{2.3.1}$$

Therefore the residual demand curve faced by trader k in class i, summing optimal demand (3.3.1) over other traders, has slope $(N-1)(\lambda_{i,1} + \gamma\sigma_2^2) + N(\lambda_{-i,1} + \gamma\sigma_2^2)$. Requiring consistency of conjectured equilibrium slope of the residual demand curve and the actual ones:

$$\lambda_{i,1} = \left((N-1)(\lambda_{i,1} + \gamma\sigma_2^2)^{-1} + N(\lambda_{-i,1} + \gamma\sigma_2^2)^{-1}\right)^{-1} \tag{2.3.2}$$

Definition 1. *A date 1 equilibrium with imperfect competition is a set of demand schedules as in* (3.3.1)*, of* $\lambda_{1,1}$ *and* $\lambda_{2,1}$ *that solve* (3.3.2) *and a price* p_1^* *such that the market clearing condition* (3.2.3) *holds.*

Proposition 2 (Vayanos (1999), Malamud and Rostek (2017))**.** *A date 1 equilibrium in linear strategies with imperfect competition exists and is unique. In this equilibrium,* $\lambda_{1,1} = \lambda_{2,1} = \frac{\gamma\sigma_2^2}{2N-2}$ *so that equilibrium demand schedules are:*

$$q^*_{i,1}(p_1) = \frac{2N-2}{2N-1}\left[\frac{v_1 - p_1}{\gamma\sigma_2^2} - I_{i,1}\right] \tag{2.3.3}$$

The equilibrium quantities traded and post trade inventories are

$$q^*_{i,1} = \frac{2N-2}{2N-1}\frac{I_{i,1} - I_{-i,1}}{2} + \frac{Q}{2N} \tag{2.3.4}$$

$$I_{i,1} + q^*_{i,1} = \frac{Q^*}{2} + \frac{1}{2N-1}I_{i,1} \tag{2.3.5}$$

The equilibrium price is

$$p_1^* = v_1 - \bar{\gamma}\sigma_2^2 Q^* \tag{2.3.6}$$

$$\text{with} \begin{cases} Q^* & = \frac{2N-2}{2N-1}\left(I_{1,1} + I_{2,1}\right) + \frac{Q}{N} \\ \bar{\gamma} & = \frac{2N-1}{2N-2}\frac{\gamma}{2} \end{cases}$$

[8]Cf. Kyle (1989), Vayanos (1999), Malamud and Rostek (2017) among many others.

The quantity traded by dealer i is reduced by a factor $(2N-2)/(2N-1)$ with respect to the competitive equilibrium, as shown in equation (3.3.4). This naturally results in an equilibrium quantity traded (3.3.4) reduced by the same factor and in imperfect risk sharing: class i traders retain a fraction $1/(2N-1)$ of their initial inventory $I_{i,1}$ as shown in equation (3.3.5).

Notice that the quantity traded by each class of dealers depends on date 0 equilibrium trade, as $I_{i,1} = I_{i,0} + q_{1,0}$. The equilibrium date 1 trade is fully solved in subsection 2.3.3.

The equilibrium price equals the underlying asset expected payoff conditional on information available at date 1 minus a risk premium that is the risk aversion γ times a markup factor $(2N-1)/(2N-2)$ related to imperfect competition, times the asset variance σ_2^2, while Q^* is the total quantity put on the market by all traders divided by N.

2.3.2 Date 0 equilibrium

I look for a Nash equilibrium in demand schedules as for date 1, with the additional requirement that dealers' strategies are conditional only on their initial inventories, so that all traders within a class i follow symmetric strategies.

At date 0, dealers perform the same intertemporal arbitrage as they do in the competitive case. However, they now take the impact of their trades on two prices into account: the direct impact on the contemporaneous price p_0, analogous to the date 1 equilibrium, and the indirect impact on future price p_1.

The indirect price impact plays as follows. Suppose customers learn at date 0 that customers are expected to sell. Then they decrease their demand in order to be able to sell the asset if possible at a high price, to repurchase at date 1 at a lower price because customers sell. But conditional on other traders' strategies, if a trader succeeds in selling, he will arrive at date 1 with a lower inventory, thus a higher demand, which tends to raise the price and lowers the spread he can make.

I do not model the indirect impact as an impact on a date 1 residual demand curve. Instead, I assume rational expectations on equilibrium price and quantities: dealer k within class i conjectures symmetric equilibrium trades $q^e_{l,-i,0} = q^e_{-i,0}$ for all dealers in the other class $-i$, and $q^e_{l,i,1} = q^e_{i,1}$ for other dealers $(l \neq k)$ in his class i. This leads to conjectured date 1 initial inventory $I^e_{-i,1} = I_{-i,0} + q^e_{-i,0}$, and $I^e_{l,i,1}$ $(l \neq k)$. Dealer k optimizes according to this conjecture. In equilibrium, these conjectures coincide with actual equilibrium quantities:

$$q^e_{i,0} = q^*_{i,0}(p^*_0). \tag{2.3.7}$$

The certainty equivalent of wealth can be written (*cf.* lemma 18 in the appendix):

$$\widehat{W}_{k,i,0} = I_{i,0}v_0 + q_{k,0}(v_0 - p_0) - \frac{\gamma(\sigma_1^2 + \sigma_2^2)}{2}(I_{k,1})^2 + \widehat{S}^*_{k,1}(q_{k,0}) \tag{2.3.8}$$

where

$$\widehat{S}^*_{k,1}(q_{k,0}) = \frac{2N}{2N-2}\frac{1}{1+\alpha x}\frac{\gamma\sigma_2^2}{2}\left(\frac{\bar{\gamma}}{\gamma}\mathbb{E}_0\left[q^*_{i,1}\right]\right)^2$$
$$= \frac{\alpha}{1+\alpha x}\frac{\gamma\sigma_2^2}{2}\left(\frac{\bar{\gamma}}{\gamma}\mathbb{E}_0\left[\frac{Q}{N}\right] + \frac{\bar{I}^e_{-i,1}}{2} + \frac{N-1}{N}\frac{\bar{I}^e_{i,1}}{2} - \left(1-\frac{1}{2N}\right)I_{k,1}\right)^2$$

and $x = \bar{\gamma}^2\sigma_2^2\sigma_q^2$ and $\alpha = \frac{2N(2N-2)}{(2N-2)^2}$ is increasing with N and is strictly between 0 and 1. The first order condition of the maximization of the above criterion, together with a consistency condition of price impacts $\lambda_{i,0}$ analogous to (3.3.2), leads to equilibrium demand schedules that are identical across dealers of class i (*cf.* lemma 19 in appendix):

$$q^*_{i,0}(p_0) = \frac{2N-2}{2N-1}\left[\frac{v_0-p_0}{\gamma(\sigma_1^2+\delta\sigma_2^2)} - I_{i,0}\right.$$
$$\left. -\frac{2N-2}{2N-1}\frac{1}{1+\alpha x}\frac{\gamma\sigma_2^2}{\gamma(\sigma_1^2+\delta\sigma_2^2)}\left(\frac{\bar{\gamma}}{\gamma}\mathbb{E}_0\left[\frac{Q}{N}\right] + \frac{\bar{I}^e_{-i,1}}{2} + \frac{N-1}{N}\frac{\bar{I}^e_{i,1}}{2}\right)\right] \tag{2.3.9}$$

where $\delta = 1 - \frac{N-1}{N}\frac{1}{1+\alpha x} \in [0,1)$. There are a few differences between this imperfectly competitive demand schedule and the competitive one. Similarly to the date 1 equilibrium, all terms in the demand schedule are reduced by a factor by $(2N-2)/(2N-1) < 1$. In addition, consider the first term representing the hold-to-maturity component of demand: it is divided by a variance $\sigma_1^2 + \delta\sigma_2^2$ where $\delta \in [0,1)$ and is analogous to the competitive case. But unlike the competitive case, when the supply shock Q is known for sure ($\sigma_q^2 = 0$), $\delta = 1/N > 0$. It reflects the fact that dealer i has to keep part of his position built at date 0 until maturity, as risk sharing is limited by imperfect competition at date 1. Otherwise δ increases with σ_q^2 and converges to 1 as σ_q^2 becomes arbitrarily large, as in the competitive case: dealer k's effective horizon converges to $t=2$.

Plugging (3.3.9) into the market clearing condition (3.2.2), and imposing consistency of conjectures on others' trades (3.3.7), I derive the following proposition, fully proven in the appendix.

Proposition 3. *At date 0, the equilibrium quantity traded by class 1 traders is*

$$q^*_{1,0} = \frac{1}{1+A(\sigma_q^2)}\frac{I_{2,0}-I_{1,0}}{2} \tag{2.3.10}$$

where $A(\cdot)$ is the demand reduction rate. It is positive and decreases with σ_q^2.

$A(\cdot)$ is the sum of a static demand reduction rate and of a positive dynamic demand reduction rate. The static demand reduction rate is the same as the one obtained in the static date 1 market. The dynamic demand reduction rate is positive and falls to zero as σ_q^2 goes to infinity.

The date 0 equilibrium price is:

$$p_0^* = v_0 - \gamma(\sigma_1^2 + \sigma_2^2)\frac{I_{1,0} + I_{2,0}}{2} - \frac{\gamma\sigma_2^2}{1+\alpha x}\,\mathbb{E}_0\left[\frac{Q}{2N}\right] \tag{2.3.11}$$

As expected with imperfect competition, the equilibrium quantity traded $|q_{1,0}^*|$ is lower than the competitive quantity $q_{1,0}^c = (I_{2,1} - I_{1,1})/2$ since $A(\cdot)$ is positive. The properties of A also imply that for finite σ_q^2, the reduction $1/(1+A)$ in quantity traded is lower than in the static case: this simply reflects the fact that dealer i cares about the impact of his trade on two prices p_0 and p_1.

However, as σ_q^2 increases, a dealer of class i cares less about what happens at date 1, including her impact on date 1 price: she restricts her trade less. When σ_q^2 becomes very large, the demand reduction factor converges to the static case: dealer i trades as if there was no trading opportunity at date 1.

2.3.3 Uncertainty about Q slows down trading

I am now able to fully solve for the equilibrium quantities: combining 3.3.4 and 3.3.10, one gets the date 1 quantity traded by dealers 1 (and symmetrically for dealers 2):

$$q_{1,1}^* = \frac{2N-2}{2N-1} \times \frac{A\left(\sigma_q^2\right)}{1+A\left(\sigma_q^2\right)} \times \frac{I_{2,0} - I_{1,0}}{2} + \frac{Q}{2N} \tag{2.3.12}$$

There is a trade postponement effect: when A increases, date 0 quantity is reduced as shown by (3.3.10), but date 1 quantity increases. However not all quantity is postponed to date 1 when A is higher, as shown by the total quantity traded at dates 0 and 1:

$$q_{1,0}^* + q_{1,1}^* = \frac{1 + \frac{2N-2}{2N-1}A(\sigma_q^2)}{1+A(\sigma_q^2)} \times \frac{I_{2,0} - I_{1,0}}{2} + \frac{Q}{2N} \tag{2.3.13}$$

The total quantity traded increases when $A(\sigma_q^2)$ increases. Given that $A(\sigma_q^2)$ decreases with σ_q^2, one can conclude the following.

Theorem 1. *When the uncertainty σ_q^2 over date 1 customer supply shock* decreases, *dealers postpone and reduce risk sharing.*

The intuition is easy to grasp. Consider sellers: when σ_q^2 decreases, the uncertainty about date 1 price p_1^* decreases, so that it the gains from postponing trade to date 1 to limit price impact increase.

This result is crucial because I show in later section that with derivatives, dealers behave as if σ_q^2 was lower than it actually is.

2.4 Imperfect competition creates gains from trading risk over Q

In this section I show that imperfect competition creates gains from trading the risk on Q because buyers and sellers then have opposite exposure to a risk on Q. There are two effects that play in the same direction. To show these effects I come back to dealer utilities post date 1 trade under perfect and imperfect competition. In both cases, they can be written:

$$W_{i,1}^{eq} = I_{i,0}v_1 + q_{i,0}^{eq}(v_1 - p_0) - \frac{\gamma\sigma_2^2}{2}\left(I_{i,1}^{eq}\right)^2 + S_{i,1}^{eq} \tag{2.4.1}$$

$$\text{with} \quad S_{i,1}^{eq} = q_{i,1}^{eq}(v_1 - p_1^{eq}) - \frac{\gamma\sigma_2^2}{2}\left[\left(I_{i,1}^{eq} + q_{i,1}^{eq}\right)^2 - (I_{i,1}^{eq})^2\right] \tag{2.4.2}$$

where the superscript $eq \in \{c, *\}$ refers to equilibrium variables under perfect and imperfect competition. All effects of Q on dealer i's marginal utilities go through the net surplus $S_{i,1}^{eq}$ from date 1 transaction. $S_{i,1}^{eq}$ is composed of the expected payoff component $q_{i,1}^{eq}(v_1 - p_1^{eq})$, and the impact on risk holding cost $\frac{\gamma\sigma_2^2}{2}\left[\left(I_{i,1} + q_{i,1}^{eq}\right)^2 - (I_{i,1}^{eq})^2\right]$. The two effects relate to each of these components. To compute the equilibrium values of these components, notice that the date 1 equilibrium price is

$$p_1^{eq} = v_1 - \gamma\sigma_2^2\frac{I_{1,0} + I_{2,0}}{2} - C^{eq}\gamma\sigma_2^2\frac{Q}{2N} \tag{2.4.3}$$

where under perfect competition $C^c = 1$, and under imperfect competition $C^* = \frac{2N-1}{2N-2} > 1$. Optimal quantity traded and inventories are

$$q_{i,1}^{eq} = B^{eq}\frac{I_{-i,1}^{eq} - I_{i,1}^{eq}}{2} + \frac{Q}{2N} \tag{2.4.4}$$

$$I_{i,1}^{eq} + q_{i,1}^{eq} = \frac{B^{eq}}{2}I_{-i,1}^{eq} + \left(1 - \frac{B^{eq}}{2}\right)I_{i,1}^{eq} + \frac{Q}{2N} \tag{2.4.5}$$

where $B^c = 1$ under perfect competition: this is because all gains from trade are exhausted at date 0 so that dealers arrive at date 1 with equal inventories. Under imperfect competition, $B^* = \frac{2N-2}{2N-1} \in (0, 1)$.

First effect: price risk. Under imperfect competition, dealers with higher initial inventory at date 0 still have higher inventory and are willing to sell at date 1: thus they dislike when customers sell at the same time as them because it decreases the price at which they sell. Symmetrically dealers with low initial inventory dislike when customers buy at the same time as them. This is not the case under perfect competition, because all interdealer gains from trade are exhausted at date 0 and all dealers arrive with symmetric inventories at date 1.

Formally, this effect relates to the expected payoff component of date 1 surplus: using equilibrium price 2.4.3 and quantity 2.4.4, one sees that

$$q^{eq}_{i,1}(v_1 - p^{eq}_1) = \underbrace{\left(B\frac{I^{eq}_{-i,1} - I^{eq}_{i,1}}{2} + \frac{Q}{2N}\right)}_{q^{eq}_{i,1}} \underbrace{\gamma\sigma_2^2\left(\frac{I^{eq}_{i,1} + I^{eq}_{-i,1}}{2} + C\frac{Q}{2N}\right)}_{v_1 - p^{eq}_1}$$

A given realization of Q has three effects on the expected payoff of date 1 transaction. The first is that the quantity Q impacts expected return at which interdealer gains from trade are realized, which is the term $\gamma\sigma_2^2 Q/2N \times (I^{eq}_{-i,1} - I^{eq}_{i,1})/2$: dealers with high initial inventories are sellers $(I^{eq}_{-i,1} - I^{eq}_{i,1} < 0)$ and make an unexpected profit when customers are buyers $(Q < 0)$, while they make an unexpected loss when customers sell at the same time as them $Q > 0)$. By contrast, dealers with low initial inventory make a loss when customers buy at the same time as them, and make an unexpected profit when customers are sellers. This effect is not present under perfect competition, because gains from trade are exhausted at date 1, so that $I^c_{i,1} = I^c_{-i,1}$.

Second, Q affects the quantity traded given the price, to which the term $Q/N \times \gamma\sigma_2^2(I^{eq}_{i,1} + I^{eq}_{-i,1})/2$ corresponds. Both classes of dealers are exposed in the same way to this risk under perfect and imperfect competition, with the same marginal utility: it is not obvious that there is room for trading this part of the risk.

Third, Q has a second order effect represented by $(Q/2N)^2$, which is always positive: unexpected customer sales occur at an unexpectedly low price, which increases the surplus both classes of dealers earn. Again all dealers are exposed in the same way to this risk, with the same marginal utilities related to this effect.

Overall I conclude that only the price effect leads marginal utilities between buyers and sellers make marginal expected payoff from date 1 transaction differ for each dealers.

Second effect: asymmetric effect of Q on holding costs. The quadratic form of risk holding costs $\frac{\gamma\sigma_2^2}{2}(I^{eq}_{i,1} + q^{eq}_{i,1})^2$, and the fact that all dealers get the same share of customer trades, implies that dealers with large date 1 initial inventory $I^{eq}_{i,1}$ incur a larger cost (relief) than buying dealers when customers sell (buy) than dealers with low date 1 initial inventory. Indeed the marginal holding cost for dealer i is

$$\frac{\partial \frac{\gamma\sigma_2^2}{2}(I^{eq}_{i,1} + q^{eq}_{i,1})^2}{\partial Q} = -\frac{\gamma\sigma_2^2}{4N}\left(\frac{B^{eq}}{2}I^{eq}_{-i,1} + \left(1 - \frac{B^{eq}}{2}\right)I^{eq}_{i,1} + \frac{Q}{2N}\right)$$

As under imperfect competition $B^* < 1$, selling dealers (with $I^{eq}_{i,1} > I^{eq}_{-i,1}$) face a larger marginal cost of customer trades than buying dealers.

The first and the second effect go in the same direction, which leads to the following proposition.

Proposition 4. *Dealers are exposed in an opposite way to the risk on date 1 price that customers' supply shock Q generates. When customers sell $(Q > 0)$, dealers* *starting date 1 with a higher inventory are marginally worse off than dealers with low inventory. The relation is reversed when customers buy. Thus there are gains from trading this risk.*

Thus under imperfect competition where dealers still have unequal inventories after date 0 trade, there are gains from dealers trading risk on Q.

Under perfect competition, dealers have equal inventories after date 0 trade and there are no gains from trading risk on Q.

2.5 Equilibrium with derivatives and imperfect competition

In this section I study the equilibrium when derivatives are allowed. Specifically, I study two equilibria with imperfect competition, each with a different derivative contract. The first contract is linear in Q and allows to isolate the effect of hedging the associated risk; however it can be implemented as a forward contract under more restrictive conditions. The second contract can be implemented as a forward contract in general.

As both contracts have qualitatively similar impact on the dynamic of trading, I expose equilibrium results simultaneously.

2.5.1 The derivative contracts

I thus introduce several contracts that are prima facie abstract, and equilibrium resolution shows that they can be implemented as forward or swap contracts. They are indexed on plausibly not directly observable variables, but in equilibrium they are function of observable prices.

Contract a: Linear payoff in Q. As uncertainty about Q is the root of the effects studied above, I study a contract whose payoff depends only on Q first: it isolates the effect of hedging of Q that intervenes in forward contracts. Beyond the theoretical interest of isolating the pure effect of trading the risk on Q, I also later show that they can be interpreted as forward contracts when $\sigma_1^2 = 0$. With this contract, one unit purchased of such contract pays off at $t = 1$

$$v_a = \alpha\bar{\gamma}\sigma_2^2 \frac{Q}{N},$$

where the normalization by the constant $\alpha\bar{\gamma}\sigma_2^2$ is convenient and without loss of generality. At date 0, v_a is normally distributed, with mean $\mu_a = \alpha\bar{\gamma}\sigma_2^2\,\mathbb{E}_0\left[\frac{Q}{N}\right]$ and and variance $\sigma_a^2 = \alpha^2\bar{\gamma}^2\sigma_2^4\sigma_q^2 = \alpha^2\sigma_2^2 x$. Trader k of class i purchases a_k units of the contract at unit price π_a, which is in zero net supply. Slightly anticipating on the equilibrium result that dealers within class i all post the same demand schedule and denoting p_0^a the equilibrium price of the underlying asset when contract a is traded, the market clearing condition is

$$a_1^*(p_0^a, \pi_a^*) + a_2^*(p_0^a, \pi_a^*) = 0 \tag{2.5.1}$$

Purchasing the contract means receiving a cash compensation when external customers sell more. It therefore acts as a hedge against unexpectedly low price for the reason that customers unexpectedly sold.

Contract b: forward contract. Next I study a contract whose payoff is given by:

$$v_b = -\alpha\epsilon_1 + \alpha\bar{\gamma}\sigma_2^2 \frac{Q}{N}$$

I show later that they can it implemented as a forward contract even when $\sigma_1^2 > 0$: buying the contract is equivalent to selling a forward.

One already sees that they are conditonal on all random variables which realize at date 1. It is straightforward to show that $v_b = \alpha(\mathbb{E}_0[p_1^*] - p_1^*)$. The proof that contract b can be implemented as a forward is in subsection 2.5.3.

The expected payoff and variance of the payoff v_b are $\mu_b = \alpha\bar{\gamma}\sigma_2^2\,\mathbb{E}_0\left[\frac{Q}{N}\right]$ and $\sigma_b^2 = \alpha^2\sigma_1^2 + \alpha^2\bar{\gamma}^2\sigma_2^4\sigma_q^2 = \alpha^2(\sigma_1^2 + \sigma_2^2 x)$. Dealer k of class i purchase a quantity b_k of this contract at price π_b. Denoting p_0^b the equilibrium price of the underlying asset when contract b is traded, and slightly anticipating on the equilibrium result that dealers within class i all post the same demand schedule $b_i^*(p_0^b, \pi_b)$, a similar market clearing condition applies:

$$b_1^*(p_0^b, \pi_b^*) + b_2^*(p_0^b, \pi_b^*) = 0 \tag{2.5.2}$$

2.5.2 What derivatives hedge

Here I show what hedging role derivatives have by exhibiting the marginal valuations of the underlying asset and that of derivatives. In particular I show why both imperfect competition and risk on Q are necessary to generate trading in the underlying asset.

Contract a. Denote $\widehat{W}_{i,0}^a(q_{i,0}, a_i)$ the date 0 certainty equivalent of wealth for trader i.

Lemma 1. *The marginal valuation for the derivative is*

$$\frac{\partial \widehat{W}_{k,i,0}^a}{\partial a_k} = \frac{\mu_a - (1+\alpha x)\pi_a - \gamma\sigma_a^2 a_k}{1+\alpha x} - \lambda_{aq} q_{k,0} - \lambda_{aa} a_k + \nu\gamma\sigma_2^2 I_{k,1} - \frac{2N}{2N-1}\nu\gamma\sigma_2^2\left(\frac{\bar{I}_{-i,1}^e}{2} + \frac{N-1}{N}\frac{\bar{I}_{i,1}^e}{2}\right)$$

where $x = \bar{\gamma}^2\sigma_2^2\sigma_q^2$ *and* $\nu = \frac{2N-2}{2N-1}\frac{\alpha x}{1+\alpha x} \in [0, 1)$.

The first term reflects the classical mean-variance trade-off for buying the derivative, irrespective of what it hedges. It is discounted by a factor $1+\alpha x$ because dealer i is sensitive to the square of Q. The second and third terms in λ_{aq} and λ_{aa} reflect dealer i's price impact management, which results in imperfect competition.

The fourth term in $I_{k,1}$ is the hedging term: it is positive, so that a higher long position in the underlying inventory $I_{k,1}$ is associated with a higher marginal value of the contract.

Lemma 2. *The marginal valuation for the underlying asset, conditional on purchasing the derivative in quantity a_i, is*

$$\frac{\partial \widehat{W}^a_{k,i,0}}{\partial q_{k,0}} = v_0 - p_0 - \gamma(\sigma_1^2 + \delta\sigma_2^2)(I_{i,0} + q_{k,0}) - \lambda_{qq} q_{i,0} - \lambda_{qa} a_k$$
$$- \frac{2N-1}{2N}\frac{\mu_a}{1+\alpha x} - \frac{2N-2}{2N-1}\frac{\gamma\sigma_2^2}{1+\alpha x_2}\left(\frac{\bar{I}^e_{-i,1}}{2} + \frac{N-1}{N}\frac{\bar{I}^e_{i,1}}{2}\right) + \nu\gamma\sigma_2^2 a_k$$

where $\delta = 1 - \frac{N-1}{N}\frac{1}{1+\alpha x}$ and $\nu = \frac{2N-2}{2N-1}\frac{\alpha x}{1+\alpha x} \in [0,1)$.

The first line reflects the classical marginal value of holding the asset until maturity, and dealer i's price impact management (last two terms). A marginal unit of the underlying asset brings expected profit $v_0 - p_0$, but dealer i also incurs a marginal cost from holding asset risk $\gamma(\sigma_1^2 + \sigma_2^2)(I_{i,0} + q_{i,0})$ for two periods. It also has an impact on both the underlying asset price p_0, and on the derivative price π_a. On the second line, the first term reflects the marginal value on date 1 surplus. The last term shows that purchasing the derivative (selling a forward) also increases the marginal value of holding the underlying asset whenever $\sigma_q^2 > 0$: this is consistent with the derivative allowing to keep inventory for a longer period of time.

Contract *b*. For contract b, marginal valuations are similar to that of contract a, except for an additional dependence on σ_1^2 since the payoff of contract b is proportional to ϵ_1.

Lemma 3. *For contract b, the marginal valuation fo the underlying asset and contract b are*

$$\frac{\partial \widehat{W}^b_{k,i,0}}{\partial q_{k,0}} = v_0 - p_0 - \lambda^b_{qq} q_{k,0} - \lambda^b_{qb} b_k - \gamma(\sigma_1^2 + \delta\sigma_2^2) I_{k,1} + \alpha\gamma\left(\sigma_1^2 + \frac{2N-1}{2N}\nu_b\sigma_2^2\right) b_k$$
$$- \frac{\gamma\sigma_2^2}{1+\alpha x}\mathbb{E}_0\left[\frac{Q}{2N}\right] - \frac{2N-2}{2N-1}\frac{\gamma\sigma_2^2}{1+\alpha x}\left(\frac{\bar{I}^e_{-i,1}}{2} + \frac{N-1}{N}\frac{\bar{I}^e_{i,1}}{2}\right) \tag{2.5.3}$$

with $\delta = 1 - \frac{N-1}{N}\frac{1}{1+\alpha x}$ and $\nu_b = \frac{\alpha x}{1+\alpha x}$, and

$$\frac{\partial \widehat{W}^b_{k,i,0}}{\partial b_k} = \frac{\mu_b}{1+\alpha x} - \pi_b - \alpha\gamma\left(\alpha\sigma_1^2 + \nu_b\sigma_2^2\right) b_k - \lambda^b_{bq} q_{k,0} - \lambda^b_{bb} b_k$$
$$+ \alpha\gamma\left(\sigma_1^2 + \frac{2N-1}{2N}\nu_b\sigma_2^2\right) I_{k,1} - \alpha\nu_b\gamma\sigma_2^2\left(\frac{\bar{I}^e_{-i,1}}{2} + \frac{N-1}{N}\frac{\bar{I}^e_{i,1}}{2}\right)$$

2.5.3 Implementation as forward contracts

The following proposition now shows when contracts a and that contract b can be implemented as forward contracts.

Proposition 5. *In the equilibrium with contract a, the equilibrium price for contract a is:*

$$\pi_a^* = \frac{\mu_a}{1 + \alpha x}$$

In the equilibrium with contract b, the equilibrium price for contract b is:

$$\pi_b^* = \frac{\mu_b}{1 + \alpha x} + \alpha\gamma\sigma_1^2 \frac{I_{1,0} + I_{2,0}}{2}$$

The underlying asset price p_0^{*a} in the equilibrium with contract a, and p_0^{*b} in the equilibrium with contract b, are both equal to the price with imperfect competition and without derivatives:

$$p_0^a = p_0^b = p_0^* \tag{2.5.4}$$

The payoff of buying contract b is that of a selling *a forward:*

$$v_b - \pi_b = \alpha \left(p_0^f - p_1^* \right)$$

where the forward price is

$$p_0^f = v_0 - \gamma(\sigma_1^2 + \sigma_2^2) \frac{I_{1,0} + I_{2,0}}{2} - \frac{\bar{\gamma}\sigma_2^2}{1 + \alpha x} \mathbb{E}_0 \left[\frac{Q}{N} \right]$$

When there no information about the terminal payoff arrives at date 1 ($\sigma_1^2 = 0$), contract a can be similarly implemented as a forward, with the forward price equal to p_0^f.

The proposition is proven in appendix. The price of contract a is simply the expected value of the contract, discounted at rate αx proportional to σ_q^2. There is no derivative inventory risk premium because the contract is in zero net supply. For contract b, the price is similar but a premium attached to underlying inventory risk appears: it comes from the fact that contract b's payoff is a function of ϵ_1.

2.5.4 Equilibrium trades

Theorem 1 shows that without derivatives, when the uncertainty about customer date 1 trade σ_q^2 decreases, interdealer trading is slowed down and risk sharing is decreased. As contracts a and b both hedge against this risk, I expect that dealers would behave as if the uncertainty σ_q^2 was low. This is what the following propositions confirm.

Proposition 6 (Contract a)**.** *In the equilibrium with contract a, the date 0 and date 1 equilibrium quantities are for dealer i*

$$q^a_{i,0} = \frac{1}{1 + A(0)} \frac{I_{-i,0} - I_{i,0}}{2} \tag{2.5.5}$$

$$q^a_{i,1} = \frac{2N-2}{2N-1} \times \frac{A(0)}{1+A(0)} \times \frac{I_{-i,0} - I_{i,0}}{2} + \frac{Q}{2N} \tag{2.5.6}$$

$$a^e_{i,1} = -\frac{1}{2} q^{ID,a}_{i,1} \tag{2.5.7}$$

where $A(0)$ is the date 0 demand reduction rate $A(\sigma^2_q)$ from proposition 12 when $\sigma^2_q = 0$, and $q^{ID,a}_{i,1} = q^a_{i,1} - \frac{Q}{2N}$ the quantity traded that stems from difference in dealer inventories just before date 1 trade.

In equilibrium, dealers with high inventory, who are the underlying asset sellers, purchase the derivative. This was expected from the computation of marginal valuations in subsection 2.5.2.

For contract b, the equilibrium quantities traded of the underlying asset are formally the same, except that the demand reduction rate $A_b(\sigma^2_q)$ differs. A_b is compared to $A(\sigma^2_q)$ in subsection 2.6.1.

Proposition 7 (Contract b)**.** *In the equilibrium with contract b, the date 0 and date 1 equilibrium quantities are*

$$q^b_{i,0} = \frac{1}{1 + A_b(\sigma^2_q)} \frac{I_{-i,0} - I_{i,0}}{2} \tag{2.5.8}$$

$$q^b_{i,1} = q^{ID,b}_{i,1} + \frac{Q}{2N} \tag{2.5.9}$$

$$b^e_i = -\frac{1+\kappa^b_2}{2} q^{ID,b}_{i,1} \tag{2.5.10}$$

where

$$q^{ID,b}_{i,1} = \frac{2N-2}{2N-1} \times \frac{A_b(\sigma^2_q)}{1+A_b(\sigma^2_q)} \times \frac{I_{-i,0} - I_{i,0}}{2}$$

and where $A_b(\sigma^2_q)$ is the date 0 demand reduction rate with contract b, and κ^b_2 is a positive (negative) constant whenever σ^2_1/σ^2_2 is above (below) a threshold.

The quantity of forward contract traded is also proportional to the interdealer quantity $q^{ID,b}_{i,1}$ and goes in the same direction (sellers of the underlying buy the contract, i.e. sell the forward), but the coefficient of proportionality differs.

The amount of derivative trading, both contract a and contract b, is proportional to the quantity of trades related to interdealer gains from trade: it is zero in the perfect competition limit ($N \to \infty$) as the date 0 demand reduction rate $A(0)$ becomes zero. This leads to the following.

Corollary 1. *Under perfect competition, contracts a and b are not traded.*

2.6 Dealer welfare effects of adding contract a

2.6.1 The cost of contract a and b: interdealer risk sharing is slowed down and decreased

Contract a allows dealers to hedge the price risk they face when they are willing to postpone from date 0 to date 1: this implies that they are more willing to postpone trade, which is intuitive and follows theorem 1. The following proposition confirms this for contract b.

Proposition 8. *One has for σ_1^2/σ_2^2 not too high, or for σ_1^2/σ_2^2 higher and σ_q^2 not too high,*

$$A_b(\sigma_q^2) > A(\sigma_q^2)$$

so that

$$\begin{aligned} q_{i,0}^b, q_{i,0}^a &< q_{i,0}^* \\ q_{i,1}^b, q_{i,1}^a &> q_{i,1}^* \end{aligned}$$

Contracts a and b thus slow down interdealer trading and decrease interdealer risk sharing.

The first part of this proposition is lemma 10 in the appendix, and inequalities in quantity traded immediately follow with inspection of equilibrium quantities.

2.6.2 Benefit of contract a: hedging of date 1 surplus

In section 2.5.2 I showed that under imperfect competition, as there remains interdealer gains from underlying asset trade, dealers face opposite exposure to the risk on Q. This comes from a price effect, and from an effect on the holding cost that goes in the same direction. Contract a is designed to realize these gains from trading the risk on Q.

This results in a higher value of date 1 surplus from transaction with customers. To show this, proposition 9 first decomposes dealer i's equilibrium utility into a classical mean-variance component, and date 0 and date 1 surpluses from transactions.

Proposition 9. *Under imperfect competition with no derivatives, the equilibrium certainty equivalent of wealth for dealers of class i is*

$$\widehat{W}_i^* = I_{i,0}v_0 - \frac{\gamma(\sigma_1^2 + \sigma_2^2)}{2}(I_{1,0})^2 + \widehat{S}_{i,0}^* + \widehat{S}_{i,1}^* \tag{2.6.1}$$

where the $\widehat{S}_{i,t}^a$ are the certainty equivalents of the surpluses from date t transactions that dealer i get:

$$\begin{aligned} \widehat{S}_{i,0}^* &= \frac{\gamma}{2}\left(\sigma_1^2 + \sigma_2^2\right)(1 + 2A_0)\left(q_{1,0}^*\right)^2 + \frac{\gamma\sigma_2^2}{1+\alpha x} q_{i,0}^* \mathbb{E}_0\left[\frac{Q}{2N}\right] \\ \widehat{S}_{i,1}^* &= (1 + 2A_1)\frac{\gamma\sigma_2^2}{2}\frac{\left(\mathbb{E}_0\left[q_{i,1}^*\right]\right)^2}{1+\alpha x} \end{aligned}$$

where $A_0 \equiv A\left(\sigma_q^2\right)$ *and* $A_1 \equiv \frac{1}{2N-2}$.

The proof of this proposition is in the appendix. The first two terms are the classical mean-variance preferences associated with dealer i's initial inventory. The third and fourth term come from the gains from date 0 and date 1 trade. $\widehat{S}^*_{i,0}$ is the sum of one term proportional to the square of the quantity traded by dealer i, and to one term proportional to the expectation of future customer demand. The latter is a transfer between dealer i and dealer $-i$, as the sum of these terms over all dealers equals zero by market clearing.

The certainty equivalent from date 1 surplus $\widehat{S}^*_{1,1}$ discounts the surplus by a factor $(1+\alpha\bar{\gamma}^2\sigma_2^2\sigma_q^2)^{-1}$, reflecting the risk associated with Q: the higher σ_q^2, the more the surplus is discounted and $\widetilde{S}^*_{1,1}$ becomes arbitrarily close to zero as σ_q^2 becomes arbitrarily large.

It is also interesting to compare the utilities of buyers and sellers as a function of the expectation of date 1 customer trades. Suppose that customers are expected to sell ($\mathbb{E}_0[Q] > 0$). Then at date 0, the date 0 equilibrium price is lower (*cf.* equation (3.3.11)), which raises the utility of buyers of the asset ($q^*_{i,0} > 0$): this shows up in $S^*_{i,0}$ as a term proportional to $\mathbb{E}_0[Q/N] \times q^*_{i,0}$. The same effect plays at date 1: date 1 price is expected to be low, and date 0 purchasers continue to purchase at date 1. This shows up as $\mathbb{E}_0[q^*_{i,1}] = \mathbb{E}_0\left[\frac{Q}{2N}\right] + \frac{2N-2}{2N-1}\frac{A}{1+A}\frac{I_{-i,0}-I_{i,0}}{2}$ is higher for dealers with low initial inventory.

With derivative contract a, dealer i's utility takes a similar form, but the expression of date 0 and date 1 surpluses is modified. Proposition 10 shows in particular that with contract a, the certainty equivalent of date 1 surplus does not shrink to zero anymore as σ_q^2 grows.

Proposition 10. *Under imperfect competition with derivatives, the equilibrium utility of dealers of class 1 is*

$$\widehat{W}^a_{i,0} = I_{i,0}v_0 - \frac{\gamma(\sigma_1^2+\sigma_2^2)}{2}(I_{i,0})^2 + \widehat{S}^a_{i,0} + \widehat{S}^a_{i,1} \tag{2.6.2}$$

where the $\widehat{S}^a_{i,t}$ *are the certainty equivalents of the surpluses from date* t *transactions that dealer* i *get:*

$$\widehat{S}^a_{i,0} = \frac{\gamma}{2}\left(\sigma_1^2+\sigma_2^2\right)\left(1+2A_0^a\right)\left(q^a_{i,0}\right)^2 + \frac{\alpha\bar{\gamma}\sigma_2^2}{1+\alpha x}q^a_{i,0}\,\mathbb{E}_0\left[\frac{Q}{2N}\right]$$

$$\widehat{S}^a_{i,1} = \gamma\sigma_2^2\left\{\frac{N\eta}{2N-1}\left(1-\frac{2N-2}{2N-1}\frac{\eta}{4}\right)\left(q^{ID,a}_{i,1}\right)^2 + \frac{1+2A_1}{1+\alpha x}\left(\left(1-\frac{2N-2}{2N-1}\frac{\eta}{2}\right)q^{ID,a}_{i,1} + \mathbb{E}_0\left[\frac{Q}{2N}\right]\right)^2\right\} + \frac{N\eta}{2N-2}\frac{\gamma\sigma_2^2}{1+\alpha x}q^{ID,a}_{i,1}\,\mathbb{E}_0\left[\frac{Q}{2N}\right]$$

with $A_0^a = A(0)$, $A_1 = \frac{1}{2N-2}$ *and* $q^{ID,a}_{1,1} = \frac{2N-2}{2N-1}\frac{A(0)}{1+A(0)}\frac{I_{2,0}-I_{1,0}}{2}$ *is the quantity corresponding to pure interdealer gains from trade at date 1, and* $\eta = \frac{\sigma_1^2+\sigma_2^2}{N\sigma_1^2+\sigma_2^2}$.

In particular, $\widehat{S}^a_{i,1}$ *does* not *converge to zero as* σ_q^2 *becomes infinite.*

This proposition is proven in the appendix. Date 1 surplus is now the sum of a positive term that does not vary with uncertainty σ_q^2 over Q, and a term that shrinks to zero as σ_q^2 grows larger: thus date 1 surplus remains bounded below by a strictly positive value, reflecting the hedging provided by contract a.

In the following subsection I show that the value of date 1 surplus is higher with derivatives is higher for all values of σ_q^2.

2.6.3 Contract a increases dealers' welfare

Here I show that for for all values of σ_q^2, and all values of parameters, introducing derivatives raises dealers' utility.

Theorem 2. *Suppose* $\mathbb{E}_0[Q] = 0$. *Then all dealers' utility is raised with the introduction of contract a:*

$$\widehat{W}_{i,0}^{a} \ > \ \widehat{W}_{i,0}^{*}, \qquad i = 1, 2$$

This is because the cumulative value of trading the asset is higher with derivatives:

$$\widehat{S}_{i,0}^{a} + \widehat{S}_{i,1}^{a} \ > \ \widehat{S}_{i,0}^{*} + \widehat{S}_{i,1}^{*}$$

The proof is in the appendix, with the second statement being immediate from dealer utilities in propositions 9 and 10. The theorem thus says that the derivative cost of posponing and reducing profitable trade is more than offset by the benefit of sharing risk over Q.

2.7 Spreads and trading volume in over-the-counter market with and without derivatives

In this section I investigate how forward contracts impact spreads in an over-the-counter market. Utilities derived in previous sections are increasing in trading values, therefore in inventory differences between dealers: this creates a value for decentralized trading where one class of dealers purchases a block to benefit from the associated gain from trade.

In this section I run a simple exercise: I introduce unmodelled customers at date 0 willing to trade a quantity Q_0 that is inelastic, and I allow them to deal with a single class of dealers in a separate market, making it resemble an OTC market. In particular this assumes that the interdealer market is shut down to customers, which is often the case empirically. For simplicity I assume $\mathbb{E}_0[Q] = 0$ in this section.

Dealer valuations of OTC trades increase with derivatives. Dealer date 0 utilities shown in propositions 9 and 10 can generally be written,

$$\widehat{W}_{i,0} = I_{i,0}v_0 - \frac{\gamma(\sigma_1^2 + \sigma_2^2)}{2}(I_{i,0})^2 + \underbrace{\gamma\theta\left(\frac{I_{2,0} - I_{1,0}}{2}\right)^2}_{\widehat{S}_{i,0} + \widehat{S}_{i,1}}$$

The first two terms reflect dealers' valuation of their endowment if they held it to maturity: it is a simple mean-variance criterion. The second term is the trading value, which is the sum of surpluses from future transactions. The parameter θ is derived from the proofs of propositions 9 and 10, depend on the setting:

$$\theta^c = \gamma \frac{\sigma_1^2 + \sigma_2^2}{2}$$
$$\theta^* = \theta^c - \gamma \left(\frac{A(\sigma_q^2)}{1 + A(\sigma_q^2)} \right)^2 \left(\frac{\sigma_1^2 + \sigma_2^2}{2} - \frac{\alpha}{2} \frac{\sigma_2^2}{1 + \alpha x} \right)$$
$$\theta^a = \theta^c - \frac{1}{2} \left(\frac{A(0)}{1 + A(0)} \right)^2 \left\{ \sigma_1^2 + (1 - \alpha) \sigma_2^2 + \alpha \sigma_2^2 \frac{\alpha x}{1 + \alpha x} \left(1 - \frac{2N - 2}{2N - 1} \frac{\eta}{2} \right)^2 \right\}$$

Crucially, the trading value depends on the difference between dealer 1 and dealer 2 inventories. A way of making inventory difference appear for dealers is to trade with external customers bilaterally, *i.e.* in an over-the-counter (OTC) market.

I assume that before trading OTC, dealers start with symmetric inventory I. Dealer 1's utilities derived from purchasing $q_{1,OTC}$ at price p_{OTC} are then

$$\widehat{W}_{i,0} = I v_0 + q_{1,OTC}(v_0 - p_{OTC}) - \frac{\gamma(\sigma_1^2 + \sigma_2^2)}{2} (I + q_{1,OTC})^2 + \gamma \theta \left(q_{1,OTC} \right)^2$$

Notice that theorem 2 has shown that $\theta^a > \theta^*$.

Corollary 2. *With contract a, dealers are more willing to pay for OTC trades than without derivatives.*

A natural consequence of this is that dealers' willingness to pay for the asset in the OTC market increases with derivatives. Assume σ_1^2/σ_2^2 is sufficiently large so that dealer utilities remain concave in $q_{1,OTC}$. Suppose that customers choose to trade an equal quantity with dealers of class 1, whose number N is greater than 3 to ensure existence of an equilibrium in linear strategies. Class 1 dealers post demand schedules by taking care about their price impact, and I denote p_{OTC} the market clearing price. Dealer 1's utilities derived from purchasing $q_{1,OTC}$ at price p_{OTC} are

$$\widehat{W}_{i,0} = I v_0 + q_{1,OTC}(v_0 - p_{OTC}) - \frac{\gamma(\sigma_1^2 + \sigma_2^2)}{2} (I + q_{1,OTC})^2 + \gamma \theta \left(q_{1,OTC} \right)^2$$

Their marginal valuation for the asset is therefore, again assuming linear prices:

$$\begin{aligned} \frac{\partial \widehat{W}_{1,0}}{\partial q_{1,OTC}} &= v_0 - p_{OTC} - q_{i,OTC} \lambda_{OTC} - \gamma \left(\sigma_1^2 + \sigma_2^2 \right) I - \gamma \left(\sigma_1^2 + \sigma_2^2 - 2\theta \right) q_{1,OTC} \\ &= v_0 - p_{OTC} - q_{i,OTC} \lambda_{OTC} - \gamma \left(\sigma_1^2 + \sigma_2^2 \right) I - 2\gamma \left(\theta^c - \theta \right) q_{1,OTC} \end{aligned}$$

In equilibrium, following similar analysis to previous section, one has $\lambda = (N - 2)^{-1}$. Equating the marginal valuation to zero and rearranging leads to the following demand function:

$$q_{1,OTC}(p_{OTC}) = \frac{N-2}{N-1} \frac{v_0 - p_{OTC}}{\gamma(\sigma_1^2 + \sigma_2^2 - 2\theta)} - \frac{N-2}{N-1} \frac{\sigma_1^2 + \sigma_2^2}{\gamma(\sigma_1^2 + \sigma_2^2 - 2\theta)} I$$

Dealers' demand increases, and their price elasticity decreases, with the trading valuation parameter θ.

Spread with inelastic customers. The market clearing condition is $Nq_{i,OTC}(p^*_{OTC}) = Q_0$, leading to the equilibrium price

$$p^*_{OTC} = v_0 - \gamma(\sigma_1^2 + \sigma_2^2)I - 2\frac{N-1}{N-2}\gamma(\theta^c - \theta)\frac{Q_0}{N}$$

Each dealer of class 1 gets a quantity Q_0/N after OTC trade. The date 0 price in the interdealer market is, following (3.3.11):

$$p^*_0 = p^a_0 = v_0 - \gamma(\sigma_1^2 + \sigma_2^2)\left(I + \frac{Q_0}{2N}\right)$$

Therefore the spread with respect to the interdealer price that dealers charge in the OTC market is

$$S = p^*_{OTC} - p^*_0 = -\gamma\left(\frac{N}{2(N-2)}\left(\sigma_1^2 + \sigma_2^2\right) + 2\frac{N-1}{N-2}\theta\right)\frac{Q_0}{2N}$$

Trading volume with elastic customers. It is easy to see that a higher willingness to pay for the underlying also creates more trading volume when customers are elastic. Assume that the opposite of the equilibrium demand schedule of customers is

$$Q_0(p_{OTC}) = Q_0 + \lambda p_{OTC}$$

with $\lambda > 0$. The market clearing condition is $Nq_{1,OTC}(p'_{OTC}) = Q_0 + \lambda p'_{OTC}$ and it is easy to see that the equilibrium quantity increases in trading valuation parameter θ.

2.8 Conclusion

In this paper, I provide a model in which imperfect competition in a dynamic trading context generates gains from trading risk over a supply shock in the future: risk averse dealers trade an asset slowly, and in the course of trading they fear that external customers unexpectedly buy or sell. Sellers fear that customers sell at the same time as them, lowering the average selling price; buyers fear the opposite: buyers and sellers have an opposite exposure to this supply shock. With perfect competition, dealers gains from trade are realized immediately, so that they are exposed to the supply shock in the same way: there are no gains from trading the supply shock risk.

I allow market participants, labelled dealers, to trade contracts that are linear in the supply shock and can be implemented as forward contracts. In equilibrium, sellers sell the foward contracts to buyers, and trade less initially: they keep their inventory for a longer period in spite of gains from trade. This is a cost that is more than compensated for dealers by the following benefit: derivatives both reduce the risk from future transaction income, and allos them to better to control their dynamic trading strategies.

In equilibrium, dealer utilities are the sum of a risk-adjusted return on their initial inventory, and of a trading value which is the sum of the surpluses made from trading. The trading value increases with interdealer gains from trade, *i.e.* inventory imbalances. Thus dealers have incentives to trade in decentralized markets to create imbalances and capture the trading value. Forward contracts increase the trading value: therefore it increases the willingness to pay for the asset in decentralized markets: this leads to tighter spreads and/or larger trading volume.

This paper thus contributes to the understanding of what financial institutions do with derivatives.

2.9 Appendix: Proofs

2.9.1 Proof of proposition 1

Plugging optimal demand (2.2.14) into the market clearing condition (3.2.2), one gets

$$\left(\frac{1}{\gamma_1}+\frac{1}{\gamma_2}\right)\frac{v_0-p_0^*}{\sigma_1^2+\delta_c\sigma_2^2}-\frac{\sigma_2^2}{(\sigma_1^2+\delta_c\sigma_2^2)(1+x_{2c})}\left(\frac{\gamma_1}{\gamma_1+\gamma_2}+\frac{\gamma_2}{\gamma_1+\gamma_2}\right)\mathbb{E}_0[Q_c^*]=0$$

Rearranging one gets the equilibrium price formula (2.2.15). Plugging the equilibrium price formula into the optimal demand schedule (2.2.14), one gets

$$\begin{aligned}
I_{i,0}+q_{i,0}^*(p_0^*) &= \frac{\bar{\gamma}_c(\sigma_1^2+\sigma_2^2)(I_{1,0}+I_{2,0})+\frac{\bar{\gamma}_c\sigma_2^2}{1+x_{2,c}}\mathbb{E}_0\left[\frac{Q}{N}\right]}{\gamma_i(\sigma_1^2+\delta_c\sigma_2^2)}-\frac{\sigma_2^2}{\sigma_1^2+\delta_c\sigma_2^2}\frac{1}{1+x_{2,c}}\mathbb{E}_0\left[\frac{\gamma_{-i}}{\gamma_1+\gamma_2}Q_c^*\right]\\
&=\frac{\gamma_{-i}}{\gamma_1+\gamma_2}\frac{1}{\sigma_1^2+\delta_c\sigma_2^2}\left(\sigma_1^2+\sigma_2^2-\frac{\sigma_2^2}{1+x_{2,c}}\right)(I_{1,0}+I_{2,0})\\
&=\frac{\gamma_{-i}}{\gamma_1+\gamma_2}(I_{1,0}+I_{2,0})
\end{aligned}$$

Optimal risk sharing comes from the competitiveness of the market (first welfare theorem).

2.9.2 Proof of proposition 12: date 0 imperfect competition equilibrium without derivatives

Demand schedules

I first general results when risk aversions differ across groups: $\gamma_1 \neq \gamma_2$. From proposition 11, the post-trade certainty equivalent of wealth at date 1 is given by the following lemma, proven in the appendix.

Lemma 4. *The interim expected utility for a trader k in class i is* $-\exp\left\{-\gamma_i\widehat{W}_{k,i,1}\right\}$, *where $\widehat{W}_{i,1}$ is the interim certainty equivalent of wealth given by:*

$$\widehat{W}_{k,i,1}=I_{k,0}v_1+q_{k,0}(v_1-p_0)-\frac{\gamma\sigma_2^2}{2}(I_{k,1})^2+\alpha\frac{\gamma\sigma_2^2}{2}\left(\frac{\bar{\gamma}}{\gamma}Q^*-I_{k,1}\right)^2 \tag{2.9.1}$$

$$=I_{k,0}v_1+q_{k,0}(v_1-p_0)-\frac{\gamma\sigma_2^2}{2}(I_{k,1})^2+\frac{2N}{2N-2}\frac{\gamma\sigma_2^2}{2}\left(q_{k,1}^*\right)^2 \tag{2.9.2}$$

where $\alpha=\frac{2N(2N-2)}{(2N-1)^2}=1-\frac{1}{(2N-1)^2}$.

Proof. Plugging equilibrium price (3.3.6) and quantities (3.3.4) into the date 1 certainty equivalent of wealth (2.2.4), one gets

$$\widehat{W}_{k,1}=I_{k,0}v_1+q_{k,0}(v_1-p_0)-\frac{\gamma\sigma_2^2}{2}(I_{k,1})^2+q_{k,1}^*(v_1-p_1^*)-\frac{\gamma\sigma_2^2}{2}\left(\left(I_{k,1}+q_{k,1}^*\right)^2-(I_{k,1})^2\right)$$

Recognizing $\frac{v_1-p_1^*}{\gamma\sigma_2^2} = \frac{2N-1}{2N-2}q_{k,1}^* + I_{k,1}$ and rearranging one get

$$\begin{aligned}
\widehat{W}_{k,i,1} &= I_{k,0}v_1 + q_{k,0}(v_1 - p_0) - \frac{\gamma\sigma_2^2}{2}(I_{k,1})^2 \\
&\quad + q_{k,1}^*\gamma\sigma_2^2\left(\frac{2N-1}{2N-2}q_{k,1}^* + I_{k,1}\right) - \frac{\gamma\sigma_2^2}{2}\left(2I_{k,1} + q_{k,1}^*\right)q_{k,1}^* \\
&= I_{k,0}v_1 + q_{k,0}(v_1 - p_0) - \frac{\gamma\sigma_2^2}{2}(I_{k,1})^2 + \left(\frac{2N-1}{2N-2} - \frac{1}{2}\right)\gamma\sigma_2^2\left(q_{k,1}^*\right)^2
\end{aligned}$$

which leads to the desired formulas. ■

It is then possible to compute the certainty equivalent of wealth at date 0.

Lemma 5. *The date 0 certainty equivalent of wealth for trader k in class i is:*

$$\widehat{W}_{k,i,0} = I_{i,0}v_0 + q_{k,0}(v_0 - p_0) - \frac{\gamma(\sigma_1^2 + \sigma_2^2)}{2}(I_{k,1})^2 + \frac{1}{2}\frac{2N}{2N-2}\frac{\gamma\sigma_2^2}{1+\alpha x}\left(\mathbb{E}_0\left[q_{k,1}^*\right]\right)^2 \tag{2.9.3}$$

$$\begin{aligned}
&= I_{k,0}v_0 + q_{k,0}(v_0 - p_0) - \frac{\gamma(\sigma_1^2 + \sigma_2^2)}{2}(I_{k,1})^2 \\
&\quad + \frac{\alpha}{2}\frac{\gamma\sigma_2^2}{1+\alpha x}\left(\frac{\bar{\gamma}}{\gamma}\mathbb{E}_0\left[\frac{Q}{N}\right] + \frac{\bar{I}_{-i,1}^e}{2} + \frac{N-1}{N}\frac{\bar{I}_{i,1}^e}{2} - \frac{2N-1}{2N}I_{k,1}\right)^2
\end{aligned} \tag{2.9.4}$$

where $x = \bar{\gamma}^2\sigma_2^2\sigma_q^2$ and $\bar{\gamma} = \frac{1}{2}\frac{2N-1}{2N-2}$. $\bar{I}_{-i,1}^e$ and $\bar{I}_{i,1}^e$ are the rational expectations of average dealer inventories after date 1 trade.

Proof. Start from interim expected utility (3.7.2). Take the certainty equivalent with respect to ϵ_1 first, which gives

$$\widehat{W}_{k,0}|Q = I_{i,0}v_1 + q_{k,0}(v_1 - p_0) - \frac{\gamma(\sigma_1^2 + \sigma_2^2)}{2}(I_{k,1})^2 + \alpha\frac{\gamma\sigma_2^3}{2}\left(\frac{\bar{\gamma}}{\gamma}Q^* - I_{k,1}\right)^2$$

Then take the certainty equivalent with respect to Q following lemma 11, which gives the desired formula.

For the rational expectations: Q^* is an outcome of date 0 trade, as it depends on dealers' average inventories in each class. Dealer k's trade has an impact on date his class' average inventory

$$\begin{aligned}
\bar{I}_{i,1}^e &\equiv \frac{1}{N}\sum_{l=1,l\neq k}^{N} I_{l,i,1}^e + \frac{I_{k,1}}{N} \\
&= \frac{N-1}{N}\bar{I}_{i,1}^e + \frac{I_{k,1}}{N}
\end{aligned}$$

where the second line follows from rational expectation of a symmetric equilibrium. ■

Lemma 6. *In equilibrium, all dealers within class i submit the same optimal demand schedules as follows:*

$$q^*_{i,0}(p_0) = \frac{2N-2}{2N-1}\left[\frac{v_0 - p_0}{\gamma(\sigma_1^2 + \delta\sigma_2^2)} - I_{k,0} \right.$$
$$\left. -\frac{2N-2}{2N-1}\frac{1}{1+\alpha x}\frac{\sigma_2^2}{\sigma_1^2+\delta\sigma_2^2}\left(\frac{\bar{\gamma}}{\gamma}\mathbb{E}_0\left[\frac{Q}{N}\right] + \frac{\bar{I}^e_{-i,1}}{2} + \frac{N-1}{N}\frac{\bar{I}^e_{i,1}}{2}\right)\right] \quad (2.9.5)$$

It depends on trader k's expectation on other traders' equilibrium trades.

Proof. Differentiate the certainty equivalent of wealth (3.7.4) with respect to $q_{i,0}$, taking into account its price impact that is conjectured to be constant (and denoted $\lambda_{i,0}$). Equating to zero to get the first-order condition:

$$\begin{aligned} v_0 - p_0 &= q_{k,0}(\lambda_{k,0} + \gamma(\sigma_1^2+\sigma_2^2))q_{k,0} + \gamma(\sigma_1^2+\sigma_2^2)I_{k,0} \\ &\quad + \frac{2N-2}{2N-1}\frac{\gamma\sigma_2^2}{1+\alpha x}\left(\frac{\bar{\gamma}}{\gamma}\mathbb{E}_0\left[\frac{Q}{N}\right] + \frac{\bar{I}^e_{-i,1}}{2} + \frac{N-1}{N}\frac{\bar{I}^e_{i,1}}{2} - \frac{2N-1}{2N}I_{k,1}\right) \\ &= (\lambda_{k,0} + \gamma(\sigma_1^2+\delta\sigma_2^2))q_{i,0} + \gamma(\sigma_1^2+\delta\sigma_2^2)I_{i,0} \\ &\quad + \frac{2N-2}{2N-1}\frac{\gamma\sigma_2^2}{1+\alpha_i x_2}\left(\frac{\bar{\gamma}}{\gamma}\mathbb{E}_0\left[\frac{Q}{N}\right] + \frac{\bar{I}^e_{-i,1}}{2} + \frac{N-1}{N}\frac{\bar{I}^e_{i,1}}{2}\right) \end{aligned}$$

where

$$\delta = 1 - \frac{N-1}{N}\frac{1}{1+\alpha x} \in [0,1].$$

It is thus straightforward to check that the second derivative of $\widehat{W}_{i,0}$ is negative, so that the problem is strictly concave. Using proposition 1 of Malamud and Rostek (2017)

$$\lambda_{k,0} = \frac{\gamma(\sigma_1^2+\delta\sigma_2^2)}{2N-2}$$

Plugging equilibrium price impacts $\lambda_{k,0}$ in the first order condition and rearranging, one gets the desired formula. ■

Equilibrium price and quantity

The date 0 market clearing condition can be written:

$$\frac{v_0 - p^*_0}{\gamma(\sigma_1^2+\delta\sigma_2^2)} = \frac{I_{1,0}+I_{2,0}}{2} + \frac{2N-2}{2N-1}\frac{1}{1+\alpha x}\frac{\gamma\sigma_2^2}{\gamma(\sigma_1^2+\delta\sigma_2^2)}\left(\frac{\bar{\gamma}}{\gamma}\mathbb{E}_0\left[\frac{Q}{N}\right] + \frac{2N-1}{2N}\frac{\bar{I}^e_{1,1}+\bar{I}^e_{2,1}}{2}\right)$$

By market clearing at date 0, $\bar{I}^e_{1,1} + \bar{I}^e_{2,1} = I_{1,0} + I_{2,0}$; in addition, recalling $\bar{\gamma} = \frac{1}{2}\frac{2N-1}{2N-2}$, one has

$$v_0 - p^*_0 = \gamma(\sigma_1^2+\delta\sigma_2^2)\frac{I_{1,0}+I_{2,0}}{2} + \frac{2N-2}{2N}\frac{\gamma\sigma_2^2}{1+\alpha x}\frac{I_{1,0}+I_{2,0}}{4} + \frac{\gamma\sigma_2^2}{1+\alpha x}\mathbb{E}_0\left[\frac{Q}{2N}\right]$$

Recalling the definition of δ, the equilibrium price is therefore:

$$p_0^* = v_0 - \gamma(\sigma_1^2+\sigma_2^2)\frac{I_{1,0}+I_{2,0}}{2} - \frac{\gamma\sigma_2^2}{1+\alpha x}\mathbb{E}_0\left[\frac{Q}{2N}\right] \tag{2.9.6}$$

Plugging 3.7.6 into the equilibrium demand schedule for class 1 traders:

$$\begin{aligned}
q_{1,0}^* = \frac{2N-2}{2N-1}\Bigg[\frac{I_{2,0}-I_{1,0}}{2} &+ \frac{2N-2}{2N-1}\frac{1}{1+\alpha x}\frac{\sigma_2^2}{\sigma_1^2+\delta\sigma_2^2}\left(\frac{\bar\gamma}{\gamma}\mathbb{E}_0\left[\frac{Q}{N}\right]+\frac{2N-1}{2N}\frac{\bar I_{1,1}^e+\bar I_{2,1}^e}{2}\right) \\
&- \frac{2N-2}{2N-1}\frac{1}{1+\alpha x}\frac{\sigma_2^2}{\sigma_1^2+\delta\sigma_2^2}\left(\frac{\bar\gamma}{\gamma}\mathbb{E}_0\left[\frac{Q}{N}\right]+\frac{\bar I_{2,1}^e}{2}+\frac{N-1}{N}\frac{\bar I_{1,1}^e}{2}\right)\Bigg] \\
= \frac{2N-2}{2N-1}\Bigg[\frac{I_{2,0}-I_{1,0}}{2} &+ \frac{2N-2}{2N-1}\frac{1}{1+\alpha x}\frac{\sigma_2^2}{\sigma_1^2+\delta\sigma_2^2}\left(\frac{\bar I_{1,1}^e-\bar I_{2,1}^e}{4N}\right)\Bigg] \\
= \frac{2N-2}{2N-1}\Bigg[\frac{I_{2,0}-I_{1,0}}{2} &+ \frac{1}{2N-1}\frac{N-1}{N}\frac{1}{1+\alpha x}\frac{\sigma_2^2}{\sigma_1^2+\delta\sigma_2^2}\left(\frac{I_{1,0}-I_{2,0}}{2}+q_{i,0}^*\right)\Bigg]
\end{aligned}$$

where the third line used the equilibrium condition $q_{2,0}^e = q_{2,0}^*$ and market clearing 3.2.2. Thus

$$\left(\frac{2N-1}{2N-2} - \frac{N-1}{N(2N-1)}\frac{1}{1+\alpha x}\frac{\sigma_2^2}{\sigma_1^2+\delta\sigma_2^2}\right)q_{i,0}^* = \left(1 - \frac{N-1}{N(2N-1)}\frac{1}{1+\alpha x}\frac{\sigma_2^2}{\sigma_1^2+\delta\sigma_2^2}\right)\frac{I_{2,0}-I_{1,0}}{2}$$

Notice that with $\delta = 1 - \frac{N-1}{N}\frac{1}{1+\alpha x}$,

$$1 - \frac{N-1}{N(2N-1)}\frac{1}{1+\alpha x}\frac{\sigma_2^2}{\sigma_1^2+\delta\sigma_2^2} = \frac{\sigma_1^2+\left(1-\frac{2N-2}{2N-1}\frac{1}{1+\alpha x}\right)\sigma_2^2}{\sigma_1^2+\left(1-\frac{2N-2}{2N}\frac{1}{1+\alpha x}\right)\sigma_2^2} > 0$$

This and rearranging leads to the desired equilibrium quantity:

$$q_{i,0}^* = \frac{1}{1+A(\sigma_q^2)}\frac{I_{2,0}-I_{1,0}}{2} \tag{2.9.7}$$

where

$$A(\sigma_q^2) = \frac{1}{2N-2}\frac{\sigma_1^2+\left(1-\frac{N-1}{N}\frac{1}{1+\alpha x}\right)\sigma_2^2}{\sigma_1^2+\left(1-\frac{2N-2}{2N-1}\frac{1}{1+\alpha x}\right)\sigma_2^2}$$

where the dependence in σ_q^2 in the right-hand side goes through δ and x, to get formula 3.3.10. The properties of $A(\sigma_q^2)$ are derived in lemma 7 in appendix 2.9.3.

It is also possible to write

$$A(\sigma_q^2) = \underbrace{\frac{1}{2N-2}}_{A^{static}} + \underbrace{\frac{1}{1+\alpha x}\frac{1}{(2N-2)(2N-1)}\frac{\sigma_2^2}{\sigma_1^2+\left(1-\frac{\alpha}{2}\frac{1}{1+\alpha x}\right)\sigma_2^2}}_{A^{dynamic}}$$

The static demand reduction rate deserves its name because $1/(1+A^{static}) = \frac{2N-2}{2N-1}$, which is the same reduction factor as in the date 1 market which is a static game. It is straightforward to show that $A^{dynamic}$ converges to zero as σ_q^2, thus x, tends to infinity.

The date 1 quantity is computed straightforwardly from 3.7.7 and 3.3.4.

2.9.3 Properties of the demand reduction rate $A(\sigma_q^2)$

Define for $z = \frac{1}{1+\alpha\bar{\gamma}^2\sigma_2^2\sigma_q^2} \in [0,1]$ the ratio

$$\widetilde{A}(z) = \frac{1}{2N-2}\,\frac{\sigma_1^2 + \left(1 - \frac{N-1}{N}z\right)\sigma_2^2}{\sigma_1^2 + \left(1 - \frac{2N}{2N-1}\frac{N-1}{N}z\right)\sigma_2^2}$$

so that $A(\sigma_q^2) = \widetilde{A}(z)$.

Lemma 7. *Then whatever the finite parameters $N \geq 2$, $\sigma_1^2 \geq 0$ and $\sigma_2^2 > 0$:*

1. *The function $\widetilde{A} : z \mapsto \widetilde{A}(z)$ is strictly increasing so that $A(\sigma_q^2)$ strictly decreases in σ_q^2.*

2. $$1 < (2N-2)A(\sigma_q^2) < \frac{3}{2}$$
   The lower bound 1 the limit of $(2N-2)A(\sigma_q^2)$ when σ_q^2 becomes infinite. The upper bound $3/2$ is attained only in the perfect competition limit ($N \to \infty$) when both $\sigma_1^2 = 0$ and $\sigma_q^2 = 0$.

3. *Therefore*
   $$0 < A(\sigma_q^2) < \frac{3}{4}$$
   $$\frac{1}{2} < \frac{2N-2}{2N-\frac{1}{2}} < \frac{1}{1+A(\sigma_q^2)} < \frac{2N-2}{2N-1}$$
   $$\frac{1}{2N-1} < \frac{A(\sigma_q^2)}{1+A(\sigma_q^2)} \leq \frac{3}{2}\frac{1}{2N-\frac{1}{2}} \leq \frac{3}{7}$$

Proof. For 1., compute the derivative

$$\widetilde{A}'(z) = \frac{\sigma_2^2}{2N(2N-2)}\,\frac{\sigma_1^2+\sigma_2^2}{\left(\sigma_1^2 + \left(1 - \frac{2N-2}{2N-1}\frac{N-1}{N}z\right)\sigma_2^2\right)^2} > 0$$

For 2., the first inequality is easily derived from $z \geq 0$; the case $z = 0$ corresponds to $\sigma_q^2 \to \infty$. For the second inequality, given that $\widetilde{A}(\cdot)$ is increasing,

$$(2N-2)\widetilde{A}(z) \leq (2N-2)\widetilde{A}(1) = \frac{\frac{\sigma_1^2}{\sigma_2^2} + 1 - \frac{N-1}{N}}{\frac{\sigma_1^2}{\sigma_2^2} + 1 - \frac{2N}{2N-1}\frac{N-1}{N}} \leq \frac{1 - \frac{N-1}{N}}{1 - \frac{2N}{2N-1}\frac{N-1}{N}} = 2 - \frac{1}{N}$$

where the last inequality follows from the fact that the ratio $\widetilde{A}(1)$ is decreasing in the ratio σ_1^2/σ_2^2. Given that $N \geq 2$, one finally gets the desired inequality

$$(2N-2)A(z) \leq \frac{3}{2}.$$

For 3. and 4., applying the mappings $x \mapsto 1/(1+x)$ and $x \mapsto x/(1+x)$ to inequalities derived in 2. (all members in these inequalities are greater than -1 so the first mapping reverses ordering, the second preserves it), one gets the desired inequalities. The last inequality is found by applying $N = 2$. ■

2.9.4 Resolution of date 0 equilibrium with contract a

Proof of lemmas 1 and 2

To compute the marginal valuations, I first compute the date 0 utility as a function of $q_{i,0}$ and a_i.

Lemma 8. *The certainty equivalent of wealth of trader k in class i is at date 0, given conjectures on other traders' equilibrium average post trade inventories* $\bar{I}^e_{,1} = I_{i,0}+q^e_{i,0}$ $\bar{I}^e_{-i,1} = I_{-i,0} + q^e_{-i,0}$*:*

$$\widehat{W}^a_{k,i,0}(q_{i,0}, a_i) = I_{i,0}v_0 + q_{k,0}(v_0 - p_0) - a_k \left(\pi_a + \alpha\gamma\sigma_2^2 \frac{\bar{I}^e_{i,1} + \bar{I}^e_{-i,1}}{2} \right)$$
$$- \frac{\gamma}{2} \left(\left(\sigma_1^2 + (1-\alpha)\sigma_2^2\right)(I_{k,1})^2 + \alpha\sigma_2^2(I_{k,1} - a_k)^2 \right)$$
$$+ \frac{1}{2}\frac{\alpha\gamma\sigma_2^2}{1+\alpha x} \left(\frac{\bar{\gamma}}{\gamma}\mathbb{E}_0[Q^*] - I_{k,1} + a_k \right)^2 + cst \qquad (2.9.8)$$

Proof. From the proof of lemma 18, one has with contract a's payoff purchased in quantity a_k:

$$\widehat{W}_{k,i,1} = I_{k,0}v_1 + q_{k,0}(v_1 - p_0) - \frac{\gamma\sigma_2^2}{2}(I_{k,1})^2 + \frac{N}{2N-2}\gamma\sigma_2^2\left(q^*_{k,1}\right)^2 + a_k \left(\alpha\bar{\gamma}\sigma_2^2 \frac{Q}{N} - \pi_a \right)$$
$$= I_{k,0}v_1 + q_{k,0}(v_1 - p_0) - \frac{\gamma\sigma_2^2}{2}(I_{k,1})^2 + a_k \left(\alpha\bar{\gamma}\sigma_2^2 Q^* - \alpha\gamma\sigma_2^2 \frac{\bar{I}^e_{i,1} + \bar{I}^e_{-i,1}}{2} - \pi_a \right)$$
$$+ \frac{\alpha}{2}\gamma\sigma_2^2 \left(\left(\frac{\bar{\gamma}}{\gamma}Q^*\right)^2 - 2\frac{\bar{\gamma}}{\gamma}Q^* I_{k,1} + (I_{k,1})^2 \right)$$
$$= I_{k,0}v_1 + q_{k,0}(v_1 - p_0) - \frac{\gamma\sigma_2^2}{2}\left((1-\alpha)(I_{k,1})^2 + \alpha(I_{k,1} - a_k)^2\right) - a_k \left(\alpha\gamma\sigma_2^2 \frac{\bar{I}^e_{i,1} + \bar{I}^e_{-i,1}}{2} + \pi_a \right)$$
$$+ \frac{\alpha}{2}\gamma\sigma_2^2 \left(\left(\frac{\bar{\gamma}}{\gamma}Q^*\right)^2 - 2\frac{\bar{\gamma}}{\gamma}Q^*(I_{k,1} - a_k) + (I_{k,1} - a_k)^2 \right)$$
$$= I_{k,0}v_1 + q_{k,0}(v_1 - p_0) - \frac{\gamma\sigma_2^2}{2}\left((1-\alpha)(I_{k,1})^2 + \alpha(I_{k,1} - a_k)^2\right)$$
$$- a_k \left(\alpha\gamma\sigma_2^2 \frac{\bar{I}^e_{i,1} + \bar{I}^e_{-i,1}}{2} + \pi_a \right) + \frac{\alpha}{2}\gamma\sigma_2^2 \left(\frac{\bar{\gamma}}{\gamma}Q^* - I_{k,1} + a_k \right)^2$$

Then take the certainty equivalent of this with respect to ϵ_1 and apply lemma 11 to get 2.9.8. ■

I conjecture linear demand schedules, thus linear dependence of prices in $q_{i,0}$ and a_i. With two assets, a trader cares about cross-asset price impacts $\partial p_0/\partial a_i$ and $\partial \pi/\partial q_{i,0}$. All derivatives of price with respect to quantities are thus constant.

By symmetry of risk aversions, the price impacts are the same for all traders (*cf.* Proposition 1 in Malamud and Rostek 2017) I denote the matrix of price impacts

$$\Lambda_a = \begin{pmatrix} \lambda_{qq} & \lambda_{qa} \\ \lambda_{aq} & \lambda_{aa} \end{pmatrix}$$

Differentiating 2.9.8 with respect to $q_{k,0}$, remembering that $\frac{\partial \bar{I}^e_{i,1}}{\partial q_{k,1}} = \frac{1}{N}$:

$$\begin{aligned}
\frac{\partial \widehat{W}_{k,i,0}}{\partial q_{k,0}} &= v_0 - p_0 - \lambda_{qq} q_{k,0} - \lambda_{qa} a_k - \gamma(\sigma_1^2 + \sigma_2^2)(I_{i,0} + q_{k,0}) + \alpha\gamma\sigma_2^2 a_k - \frac{\alpha\gamma\sigma_2^2}{2N} a_k \\
&\quad - \frac{2N-1}{2N} \frac{\alpha\gamma\sigma_2^2}{1+\alpha x} \left(\frac{\bar{\gamma}}{\gamma} \mathbb{E}_0 \left[\frac{Q}{N} \right] + \frac{\bar{I}^e_{-i,1}}{2} + \frac{N-1}{N} \frac{\bar{I}^e_{i,1}}{2} - \frac{2N-1}{2N} I_{k,1} + a_k \right) \\
&= v_0 - p_0 - \lambda_{qq} q_{i,0} - \lambda_{qa} a_i - \gamma(\sigma_1^2 + \delta\sigma_2^2)(I_{i,0} + q_{k,0}) + \frac{2N-2}{2N-1} \nu\gamma\sigma_2^2 a_k \\
&\quad - \frac{2N-2}{2N-1} \frac{\gamma\sigma_2^2}{1+\alpha x} \left(\frac{\bar{\gamma}}{\gamma} \mathbb{E}_0 \left[\frac{Q}{N} \right] + \frac{\bar{I}^e_{-i,1}}{2} + \frac{N-1}{N} \frac{\bar{I}^e_{i,1}}{2} \right)
\end{aligned}$$

$$\begin{aligned}
\frac{\partial \widehat{W}_{k,i,0}}{\partial q_{k,0}} &= v_0 - p_0 - \frac{2N-1}{2N} \frac{\mu_a}{1+\alpha x} - \lambda_{qq} q_{i,0} - \lambda_{qa} a_k - \gamma(\sigma_1^2 + \delta\sigma_2^2)(I_{i,0} + q_{k,0}) \\
&\quad + \nu\gamma\sigma_2^2 a_k - \frac{2N-2}{2N-1} \frac{\gamma\sigma_2^2}{1+\alpha x_2} \left(\frac{\bar{I}^e_{-i,1}}{2} + \frac{N-1}{N} \frac{\bar{I}^e_{i,1}}{2} \right)
\end{aligned}$$

where $\delta = 1 - \frac{N-1}{N} \frac{1}{1+\alpha x}$ and $\nu = \frac{2N-2}{2N-1} \frac{\alpha x}{1+\alpha x} \in [0, 1)$. This proves lemma 2.

Then differentiating 2.9.8 with respect to a_k:

$$\begin{aligned}
\frac{\partial \widehat{W}_{k,i,0}}{\partial a_k} &= -\pi_a - \lambda_{aq} q_{k,0} - \lambda_{aa} a_k + \alpha\gamma\sigma_2^2 (I_{k,1} - a_k) - \alpha\gamma\sigma_2^2 \frac{\bar{I}^e_{i,1} + \bar{I}^e_{-i,1}}{2} \\
&\quad + \frac{\alpha\gamma\sigma_2^2}{1+\alpha x} \left(\frac{\bar{\gamma}}{\gamma} \mathbb{E}_0 \left[\frac{Q}{N} \right] + \frac{\bar{I}^e_{-i,1}}{2} + \frac{N-1}{N} \frac{\bar{I}^e_{i,1}}{2} - \frac{2N-1}{2N} I_{k,1} + a_k \right) \\
&= -\pi_a - \lambda_{aq} q_{k,0} - \lambda_{aa} a_k + \frac{2N-1}{2N} \alpha \left(1 - \frac{1}{1+\alpha x} \right) \gamma\sigma_2^2 I_{k,1} - \alpha \left(1 - \frac{1}{1+\alpha x} \right) \gamma\sigma_2^2 a_k \\
&\quad + \frac{\alpha\bar{\gamma}\sigma_2^2}{1+\alpha x} \mathbb{E}_0 \left[\frac{Q}{N} \right] - \alpha\gamma\sigma_2^2 \left(1 - \frac{1}{1+\alpha x} \right) \left(\frac{\bar{I}^e_{-i,1}}{2} + \frac{N-1}{N} \frac{\bar{I}^e_{i,1}}{2} \right) \\
&= \frac{\mu_a}{1+\alpha x} - \pi_a - \lambda_{aq} q_{k,0} - \lambda_{aa} a_k + \nu\gamma\sigma_2^2 I_{k,1} - \frac{2N}{2N-1} \nu\gamma\sigma_2^2 a_k \\
&\quad - \frac{2N}{2N-1} \nu\gamma\sigma_2^2 \left(\frac{\bar{I}^e_{-i,1}}{2} + \frac{N-1}{N} \frac{\bar{I}^e_{i,1}}{2} \right)
\end{aligned}$$

$$\frac{\partial \widehat{W}_{k,i,0}}{\partial a_k} = \frac{\mu_a - \gamma\sigma_a^2 a_k}{1+\alpha x} - \pi_a - \lambda_{aq} q_{k,0} - \lambda_{aa} a_k + \nu\gamma\sigma_2^2 I_{k,1} - \frac{2N}{2N-1} \nu\gamma\sigma_2^2 \left(\frac{\bar{I}^e_{-i,1}}{2} + \frac{N-1}{N} \frac{\bar{I}^e_{i,1}}{2} \right)$$

The last line has used $\frac{2N}{2N-1}\nu = \frac{2N(2N-2)}{(2N-1)^2} \frac{\alpha x}{1+\alpha x} = \frac{\alpha^2 x}{1+\alpha x}$. This proves lemma 1.

Proof of proposition 5 (contract a)

The first order conditions equate the derivatives of lemma 1 and 2 to zero. With matrix notations, they can be written

$$M_a \begin{pmatrix} v_0 \\ \frac{\mu_a}{1+\alpha x} \end{pmatrix} - \begin{pmatrix} p_0 \\ \pi_a \end{pmatrix} = (\Lambda_a + \gamma\Sigma_a) \begin{pmatrix} q_{k,0} \\ a_k \end{pmatrix} + \gamma(\Sigma_a + K_a) \begin{pmatrix} I_{i,0} \\ \frac{\bar{I}^e_{-i,1}}{2} + \frac{N-1}{N}\frac{\bar{I}^e_{i,1}}{2} \end{pmatrix} \tag{2.9.9}$$

where

$$M_a = \begin{pmatrix} 1 & -\left(1-\frac{1}{2N}\right) \\ 0 & 1 \end{pmatrix},$$

$$\Sigma_a = \begin{pmatrix} \sigma_1^2 + \delta\sigma_2^2 & -\left(1-\frac{1}{2N}\right)\frac{\sigma_a^2}{1+\alpha x} \\ -\left(1-\frac{1}{2N}\right)\frac{\sigma_a^2}{1+\alpha x} & \frac{\sigma_a^2}{1+\alpha x} \end{pmatrix},$$

$$K_a = \begin{pmatrix} 0 & \frac{2N-2}{2N-1}\sigma_2^2 \\ 0 & 0 \end{pmatrix} \quad \Rightarrow \quad \Sigma_a + K_a = \begin{pmatrix} \sigma_1^2 + \delta\sigma_2^2 & \frac{2N-2}{2N-1}\frac{\sigma_2^2}{1+\alpha x} \\ -\left(1-\frac{1}{2N}\right)\frac{\sigma_a^2}{1+\alpha x} & \frac{\sigma_a^2}{1+\alpha x} \end{pmatrix}$$

Σ_a is symmetric, now check that it is positive definite: it has a first diagonal coefficient that is positive, and its determinant is positive:

$$\begin{aligned} |\Sigma_a| &= \nu\sigma_2^2\left(\frac{2N}{2N-1}\sigma_1^2 + \left(\frac{2N}{2N-1}\delta - \nu\right)\sigma_2^2\right) \\ &= \frac{2N}{2N-1}\nu\sigma_2^2\left(\sigma_1^2 + \left(1 - \frac{N-1}{N}\frac{1}{1+\alpha x} - \frac{2N-1}{2N}\frac{2N-2}{2N-1}\frac{\alpha x}{1+\alpha x}\right)\sigma_2^2\right) \\ &= \frac{2N}{2N-1}\nu\sigma_2^2\left(\sigma_1^2 + \left(1 - \frac{N-1}{N}\right)\sigma_2^2\right) \end{aligned}$$

Therefore

$$|\Sigma_a| = \frac{\sigma_a^2}{1+\alpha x}\left(\sigma_1^2 + \frac{\sigma_2^2}{N}\right) > 0 \tag{2.9.10}$$

Σ_a is *not* the covariance matrix of the underlying asset and of contract a (their covariance should be zero): wealth is not normally distributed because of Q. As $\nu = 0$ whenever $\sigma_q^2 = 0$, the matrix is not invertible whenever the supply shock is known for sure.

The equilibrium price impact matrix λ is given by the equation (assuming invertibility, which is verified in equilibrium):

$$\Lambda_a = \left((2N-1)(\Lambda_a + \gamma\Sigma)^{-1}\right)^{-1}$$

so that

$$\Lambda_a = \frac{1}{2N-2}\gamma\Sigma_a \tag{2.9.11}$$

The first order conditions become

$$M_a \begin{pmatrix} v_0 \\ \frac{\mu_a}{1+\alpha x} \end{pmatrix} - \begin{pmatrix} p_0 \\ \pi_a \end{pmatrix} = \frac{2N-1}{2N-2}\gamma \Sigma_a \begin{pmatrix} q^a_{k,0}(p_0,\pi_a) \\ a^*_k(p_0,\pi_a) \end{pmatrix} + \gamma(\Sigma_a + K_a) \begin{pmatrix} I_{i,0} \\ \frac{\bar{I}^e_{-i,1}}{2} + \frac{N-1}{N}\frac{\bar{I}^e_{i,1}}{2} \end{pmatrix}$$

This gives the equilibrium demand schedules:

$$\begin{pmatrix} q^*_{i,0}(p_0,\pi) \\ a^*_i(p_0,\pi) \end{pmatrix} = \frac{2N-2}{2N-1}\gamma^{-1}\Sigma_a^{-1}\left(M_a \begin{pmatrix} v_0 \\ \frac{\mu_a}{1+\alpha x} \end{pmatrix} - \begin{pmatrix} p_0 \\ \pi_a \end{pmatrix} \right) - \frac{2N-2}{2N-1}(Id_2 + \Sigma_a^{-1}K_a) \begin{pmatrix} I_{i,0} \\ \frac{\bar{I}^e_{-i,1}}{2} + \frac{N-1}{N}\frac{\bar{I}^e_{i,1}}{2} \end{pmatrix} \tag{2.9.12}$$

where Id_2 is the two-dimensional identity matrix. From (2.9.10) one has

$$\Sigma_a^{-1} = \begin{pmatrix} \left(\sigma_1^2 + \sigma_2^2/N\right)^{-1} & (1-1/2N)\left(\sigma_1^2 + \sigma_2^2/N\right)^{-1} \\ (1-1/2N)\left(\sigma_1^2 + \sigma_2^2/N\right)^{-1} & \left(\frac{\sigma_a^2}{1+\alpha x}\right)^{-1} + (1-1/2N)^2\left(\sigma_1^2 + \sigma_2^2/N\right)^{-1} \end{pmatrix}$$

Now I compute

$$Id_2 + \Sigma_a^{-1}K_a = \begin{pmatrix} 1 & \kappa_{a1} \\ 0 & 1+\kappa_{a2} \end{pmatrix} \qquad \text{with} \qquad \begin{cases} \kappa_{a1} &= \frac{2N-2}{2N-1}\frac{\sigma_2^2}{\sigma_1^2+\sigma_2^2/N} \\ \kappa_{a2} &= \frac{N-1}{N}\frac{\sigma_2^2}{\sigma_1^2+\sigma_2^2/N} \end{cases}$$

Applying market clearing conditions (3.2.2) and (2.5.1), one gets the equilibrium risk premia:

$$M_a \begin{pmatrix} v_0 \\ \frac{\mu_a}{1+\alpha x} \end{pmatrix} - \begin{pmatrix} p^a_0 \\ \pi^*_a \end{pmatrix} = \gamma(\Sigma_a + K_a) \begin{pmatrix} \frac{I_{1,0}+I_{2,0}}{2} \\ \frac{2N-1}{2N}\frac{\bar{I}^e_{1,1}+\bar{I}^e_{2,1}}{2} \end{pmatrix} \tag{2.9.13}$$

The equilibrium asset prices are given by the following, given that $I^e_{i,1} + I^e_{-i,1} = I_{1,0} + I_{2,0}$

$$\begin{aligned} M_a \begin{pmatrix} v_0 \\ \frac{\mu_a}{1+\alpha x} \end{pmatrix} - \begin{pmatrix} p^a_0 \\ \pi^*_a \end{pmatrix} &= \gamma(\Sigma_a + K_a) \begin{pmatrix} 1 \\ 1 - \frac{1}{2N} \end{pmatrix} \frac{I_{1,0}+I_{2,0}}{2} \\ &= \begin{pmatrix} \sigma_1^2 + \left(\delta + \frac{2N-2}{2N}\frac{1}{1+\alpha x}\right)\sigma_2^2 \\ 0 \end{pmatrix} \frac{I_{1,0}+I_{2,0}}{2} \end{aligned}$$

Recognizing $\delta + \frac{2N-1}{2N}\frac{2N-2}{2N-1}\frac{1}{1+\alpha x} = 1$ and $\alpha\bar{\gamma} = \frac{2N}{2N-1}\gamma/2$, this leads to:

$$p^a_0 = v_0 - \gamma\left(\sigma_1^2 + \sigma_2^2\right)\frac{I_{1,0}+I_{2,0}}{2} - \frac{\gamma\sigma_2^2}{1+\alpha x_2}\mathbb{E}_0\left[\frac{Q}{2N}\right] \tag{2.9.14}$$

$$\pi^*_a = \frac{2N}{2N-1}\frac{\gamma\sigma_2^2}{1+\alpha x}\mathbb{E}_0\left[\frac{Q}{2N}\right] \tag{2.9.15}$$

Proof of proposition 6

Plugging equilibrium risk premia 2.9.13 into class 1 traders' equilibrium demands schedule 2.9.12:

$$\begin{pmatrix} q^a_{1,0}(p^a_0,\pi^*_a) \\ a^*_1(p^a_0,\pi^*_a) \end{pmatrix} = \frac{2N-2}{2N-1}(Id_2+\Sigma_a^{-1}K_a)\begin{pmatrix} \frac{I_{1,0}+I_{2,0}}{2} \\ \frac{2N-1}{2N}\frac{\bar{I}^e_{1,1}+I^e_{2,1}}{2} \end{pmatrix}$$
$$- \frac{2N-2}{2N-1}(Id_2+\Sigma_a^{-1}K_a)\begin{pmatrix} I_{1,0} \\ \frac{\bar{I}^e_{2,1}}{2}+\frac{N-1}{N}\frac{\bar{I}^e_{1,1}}{2} \end{pmatrix}$$
$$= \frac{2N-2}{2N-1}(Id_2+\Sigma_a^{-1}K_a)\begin{pmatrix} \frac{I_{2,0}-I_{1,0}}{2} \\ \frac{1}{2N}\frac{\bar{I}^e_{1,1}-\bar{I}^e_{2,1}}{2} \end{pmatrix}$$

Therefore the equilibrium quantities verify the following equations:

$$q^a_{1,0}(p^a_0,\pi^*_a) = \frac{2N-2}{2N-1}\frac{I_{2,0}-I_{1,0}}{2} + \frac{\kappa_{a1}}{2N}\frac{2N-2}{2N-1}\frac{I_{1,0}+q^e_{1,0}-I_{2,0}-q^e_{2,0}}{2}$$
$$a^*_1(p^a_0,\pi^*_a) = \frac{1+\kappa_{a2}}{2N}\frac{2N-2}{2N-1}\frac{I_{1,0}+q^e_{1,0}-I_{2,0}-q^e_{2,0}}{2}$$

imposing consistency conditions $q^a_{i,0} = q^e_{i,0}$ and date 0 market clearing $q^a_{2,0} = -q^a_{1,0}$:

$$q^a_{1,0} = \left(1-\frac{\kappa_{a1}}{2N}\right)\frac{2N-2}{2N-1}\frac{I_{2,0}-I_{1,0}}{2} + \frac{\kappa_{a1}}{2N}\frac{2N-2}{2N-1}q^a_{1,0}$$
$$a^*_1 = \frac{1+\kappa_{a2}}{2N}\frac{2N-2}{2N-1}\left(\frac{I_{1,0}-I_{2,0}}{2}+q^a_{1,0}\right)$$

This leads to

$$q^a_{1,0} = \frac{1-\kappa_{a1}/2N}{1-\frac{\kappa_{a1}}{2N}\frac{2N-2}{2N-1}}\frac{2N-2}{2N-1}\frac{I_{2,0}-I_{1,0}}{2}$$
$$= \frac{1-\kappa_{a1}/2N}{1-\kappa_{a1}/2N+\frac{1}{2N-2}}\frac{I_{2,0}-I_{1,0}}{2}$$
$$= \frac{1}{1+\frac{1}{2N-2}\left(1-\frac{\kappa_{a1}}{2N}\right)^{-1}}\frac{I_{2,0}-I_{1,0}}{2}$$

Notice that

$$1-\frac{\kappa_{a1}}{2N} = \frac{\sigma_1^2+\sigma_2^2/N-\frac{2N-2}{2N-1}\frac{1}{2N}\sigma_2^2}{\sigma_1^2+\sigma_2^2/N} = \frac{\sigma_1^2+\frac{\sigma_2^2}{2N-1}}{\sigma_1^2+\sigma_2^2/N} < 1$$
$$= \frac{1}{(2N-2)A(0)}$$

where $A(0)$ is the demand reduction rate when there is no uncertainty on Q. Plugging the expression for $1 - \kappa_a/2N$ into equilibrium quantity and small rearranging leads to:

$$q^a_{1,0} = \frac{1}{1 + A(0)} \frac{I_{2,0} - I_{1,0}}{2} \tag{2.9.16}$$

The equilibrium quantity of contract a is

$$a^*_1 = -\frac{1 + \kappa_{a2}}{2N} \frac{2N - 2}{2N - 1} \left(1 - \frac{1}{1 + A(0)}\right) \frac{I_{2,0} - I_{1,0}}{2}$$

$$a^*_1 = -\frac{1 + \kappa_{a2}}{2N} \frac{2N - 2}{2N - 1} \frac{A(0)}{1 + A(0)} \frac{I_{2,0} - I_{1,0}}{2} \tag{2.9.17}$$

$$\text{with} \qquad \kappa_{a2} = \frac{N - 1}{N} \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2/N}$$

Another way to write the result is

$$a^*_1 = -\frac{\eta}{2} q^{ID,a}_{1,1} \tag{2.9.18}$$

$$\eta = \frac{\sigma_1^2 + \sigma_2^2}{N\sigma_1^2 + \sigma_2^2}$$

Finally, the date 1 quantity traded is obtained straightforwardly from 3.3.4 and $q^*_{1,0}$.

2.9.5 Resolution of date 0 equilibrium with forward contracts

For computational convenience, I study the contract with payoff

$$v_b = -\alpha\epsilon_1 + \alpha\bar{\gamma}\sigma_2^2 \frac{Q}{N}$$

Proof of lemma 3

I first compute the date 0 certainty equivalent of wealth as a function of $q_{i,1}$ and b_i.

Lemma 9. *The date 0 certainty equivalent of wealth is*

$$\widehat{W}^b_{k,0} = I_{i,0}v_0 + q_{i,0}(v_0 - p_0) - b_k\pi_b - \frac{\gamma}{2}\left(\sigma_1^2(I_{i,1} - \alpha b_k)^2 + \alpha\sigma_2^2(I_{k,1} - b_k)^2 + (1 - \alpha)(I_{k,1})^2\right)$$
$$+ \frac{1}{2}\frac{\alpha\gamma\sigma_2^2}{1 + \alpha x}\left(\frac{\bar{\gamma}}{\gamma}\mathbb{E}_0\left[\frac{Q}{N}\right] + \frac{\bar{I}^e_{-i,1}}{2} + \frac{N - 1}{N}\frac{\bar{I}_{i,1}}{2} - \frac{2N - 1}{2N}I_{k,1} + b_k\right)^2 - \alpha\gamma\sigma_2^2\frac{\bar{I}^e_{-i,1} + \bar{I}_{i,1}}{2} \tag{2.9.19}$$

Proof. After date 1 trade, the certainty equivalent of wealth of trader i is, from 3.3.10 and 3.3.6 and the derivative payoff

$$\begin{aligned}
\widehat{W}^b_{k,i,1} &= I_{i,0}v_0 + q_{k,0}(v_0 - p_0) - b_k\pi_b + (I_{i,0} + q_{k,0} - \alpha b_k)\epsilon_1 - \frac{\gamma\sigma_2^2}{2}(I_{i,0} + q_{k,0})^2 \\
&\quad + \frac{\alpha\gamma\sigma_2^2}{2}\left(\frac{\bar{\gamma}}{\gamma}\mathbb{E}_0\left[Q^*\right] - I_{k,1}\right)^2 + b_k\alpha\bar{\gamma}\sigma_2^2\frac{Q}{N} \\
&= I_{i,0}v_0 + q_{k,0}(v_0 - p_0) - b_k\pi_b + (I_{i,0} + q_{k,0} - \alpha b_k)\epsilon_1 - \frac{\gamma\sigma_2^2}{2}(I_{i,0} + q_{k,0})^2 \\
&\quad + \frac{\alpha\gamma\sigma_2^2}{2}\left(\left(\frac{\bar{\gamma}}{\gamma}Q^*\right)^2 - 2\frac{\bar{\gamma}}{\gamma}Q^*(I_{i,1} - b_i) + (I_{k,1} - b_k)^2\right) + \frac{\alpha\gamma\sigma_2^2}{2}(I_{k,1})^2 \\
&\quad - \frac{\alpha\gamma\sigma_2^2}{2}(I_{k,1} - b_k)^2 - \alpha\gamma\sigma_2^2\frac{\bar{I}^e_{-i,1} + \bar{I}^e_{i,1}}{2}b_k \\
&= I_{i,0}v_0 + q_{i,0}(v_0 - p_0) - b_k\pi_b + (I_{k,1} - \alpha b_k)\epsilon_1 - \frac{\gamma}{2}\left((1-\alpha)\sigma_2^2(I_{k,1})^2 + \alpha\sigma_2^2(I_{k,1} - b_k)^2\right) \\
&\quad + \frac{\alpha\gamma\sigma_2^2}{2}\left(\frac{\bar{\gamma}}{\gamma}Q^* - (I_{k,1} - b_k)\right)^2 - \alpha\gamma\sigma_2^2\frac{\bar{I}^e_{-i,1} + \bar{I}^e_{i,1}}{2}b_k
\end{aligned}$$

Taking the certainty equivalent of wealth with respect to ϵ_1 and Q leads to

$$\begin{aligned}
\widehat{W}^b_{k,0} &= I_{k,0}v_0 + q_{k,0}(v_0 - p_0) - b_k\pi_b - \frac{\gamma}{2}\left(\sigma_1^2(I_{k,1} - \alpha b_k)^2 + \alpha\sigma_2^2(I_{k,1} - b_k)^2 + (1-\alpha)(I_{k,1})^2\right) \\
&\quad + \frac{1}{2}\frac{\alpha\gamma\sigma_2^2}{1+\alpha x}\left(\frac{\bar{\gamma}}{\gamma}\mathbb{E}_0\left[Q^*\right] - I_{k,1} + b_k\right)^2 - \alpha\gamma\sigma_2^2\frac{\bar{I}^e_{-i,1} + \bar{I}^e_{i,1}}{2}b_k
\end{aligned}$$

Taking the expression in 3.3.6 for Q^* leads to the desired formula. ∎

Then differentiating (2.9.19) with respect to $q_{i,0}$:

$$\begin{aligned}
\frac{\partial\widehat{W}^b_{k,i,0}}{\partial q_{i,0}} &= v_0 - p_0 - \lambda^b_{qq}q_{k,0} - \lambda^b_{qb}b_k - \gamma\left(\sigma_1^2(I_{k,1} - \alpha b_k) + \alpha\sigma_2^2(I_{k,1} - b_k) + (1-\alpha)\sigma_2^2 I_{k,1}\right) \\
&\quad - \frac{2N-1}{2N}\frac{\alpha\gamma\sigma_2^2}{1+\alpha x}\left(\frac{\bar{\gamma}}{\gamma}\mathbb{E}_0\left[\frac{Q}{N}\right] + \frac{\bar{I}^e_{-i,1}}{2} + \frac{N-1}{N}\frac{\bar{I}^e_{i,1}}{2} - \frac{2N-1}{2N}I_{k,1} + b_k\right) \\
&\quad - \frac{1}{2N}\frac{\alpha\gamma\sigma_2^2}{2}b_k \\
&= v_0 - p_0 - \lambda^b_{qq}q_{k,0} - \lambda^b_{qb}b_k - \gamma\left(\sigma_1^2 + \sigma_2^2\right)I_{k,1} + \left(\alpha\sigma_1^2 + \alpha\frac{2N-1}{2N}\right)b_k \\
&\quad - \frac{2N-1}{2N}\frac{\alpha\gamma\sigma_2^2}{1+\alpha x}\left(\frac{\bar{\gamma}}{\gamma}\mathbb{E}_0\left[\frac{Q}{N}\right] + \frac{\bar{I}^e_{-i,1}}{2} + \frac{N-1}{N}\frac{\bar{I}^e_{i,1}}{2} - \frac{2N-1}{2N}I_{k,1} + b_k\right)
\end{aligned}$$

Rearranging leads to

$$\frac{\partial \widehat{W}^b_{k,i,0}}{\partial q_{k,0}} = v_0 - p_0 - \lambda^b_{qq} q_{k,0} - \lambda^b_{qb} b_k - \gamma(\sigma_1^2 + \delta\sigma_2^2) I_{k,1} + \alpha\gamma \left(\sigma_1^2 + \frac{2N-1}{2N}\nu_b\sigma_2^2\right) b_k$$
$$- \frac{\gamma\sigma_2^2}{1+\alpha x}\mathbb{E}_0\left[\frac{Q}{2N}\right] - \frac{2N-2}{2N-1}\frac{\gamma\sigma_2^2}{1+\alpha x}\left(\frac{\bar{I}^e_{-i,1}}{2} + \frac{N-1}{N}\frac{\bar{I}^e_{i,1}}{2}\right) \tag{2.9.20}$$

with

$$\delta = 1 - \frac{N-1}{N}\frac{1}{1+\alpha x} \qquad \text{and} \qquad \nu_b = \frac{\alpha x}{1+\alpha x}$$

Then differentiating 2.9.19 with respect to b_k:

$$\frac{\partial \widehat{W}^b_{k,i,0}}{\partial b_k} = -\pi_b - \lambda^b_{bq} q_{k,0} - \lambda^b_{bb} b_k + \gamma\left(\alpha\sigma_1^2(I_{k,1} - \alpha b_k) + \alpha\sigma_2^2(I_{k,1} - b_k)\right)$$
$$+ \frac{\alpha\gamma\sigma_2^2}{1+\alpha x}\left(\frac{\bar{\gamma}}{\gamma}\mathbb{E}_0\left[\frac{Q}{N}\right] + \frac{\bar{I}^e_{-i,1}}{2} + \frac{N-1}{N}\frac{\bar{I}^e_{i,1}}{2} - \frac{2N-1}{N} I_{k,1} + b_k\right)$$
$$- \alpha\gamma\sigma_2^2 \frac{\bar{I}^e_{-i,1} + \bar{I}^e_{i,1}}{2}$$
$$= -\pi_b - \lambda^b_{bq} q_{k,0} - \lambda^b_{bb} b_k + \alpha\gamma\left(\sigma_1^2 + \sigma_2^2\right) I_{k,1} - \alpha\gamma\left(\alpha\sigma_1^2 + \sigma_2^2\right) b_k$$
$$+ \frac{\alpha\gamma\sigma_2^2}{1+\alpha x}\left(\frac{\bar{\gamma}}{\gamma}\mathbb{E}_0\left[\frac{Q}{N}\right] + \frac{\bar{I}^e_{-i,1}}{2} + \frac{N-1}{N}\frac{\bar{I}^e_{i,1}}{2} - \frac{2N-1}{N} I_{k,1} + b_k\right)$$
$$- \alpha\gamma\sigma_2^2\left(\frac{\bar{I}^e_{-i,1}}{2} + \frac{N-1}{N}\frac{\bar{I}^e_{i,1}}{2} + \frac{I_{k,1}}{2N}\right)$$

Rearranging leads to

$$\frac{\partial \widehat{W}^b_{k,i,0}}{\partial b_k} = -\pi_b - \lambda^b_{bq} q_{k,0} - \lambda^b_{bb} b_k + \alpha\gamma\left(\sigma_1^2 + \frac{2N-1}{2N}\nu_b\sigma_2^2\right) I_{k,1} - \alpha\gamma\left(\alpha\sigma_1^2 + \nu_b\sigma_2^2\right) b_k$$
$$+ \frac{\gamma\sigma_2^2}{1+\alpha x}\mathbb{E}_0\left[\frac{Q}{2N}\right] - \alpha\nu_b\gamma\sigma_2^2\left(\frac{\bar{I}^e_{-i,1}}{2} + \frac{N-1}{N}\frac{\bar{I}^e_{i,1}}{2}\right) \tag{2.9.21}$$

which proves lemma 3.

Proof of proposition 5 (contract b)

Optimal demand schedules. From 2.9.20 and 2.9.21, the first order conditions can be expressed in matrix form as follows:

$$M_b \begin{pmatrix} v_0 \\ \frac{\mu_b}{1+\alpha x} \end{pmatrix} - \begin{pmatrix} p_0 \\ \pi_b \end{pmatrix} = (\Lambda_b + \gamma\Sigma_b)\begin{pmatrix} q_{k,0} \\ b_k \end{pmatrix} + \gamma(\Sigma_b + K_b)\begin{pmatrix} I_{i,0} \\ \frac{\bar{I}^e_{-i,1}}{2} + \frac{N-1}{N}\frac{\bar{I}^e_{i,1}}{2} \end{pmatrix} \tag{2.9.22}$$

where

$$M_b = \begin{pmatrix} 1 & -\left(1-\frac{1}{2N}\right) \\ 0 & 1 \end{pmatrix},$$

$$\Sigma_b = \begin{pmatrix} \sigma_1^2 + \delta\sigma_2^2 & -\alpha\left(\sigma_1^2 + \frac{2N-1}{2N}\nu_b\sigma_2^2\right) \\ -\alpha\left(\sigma_1^2 + \frac{2N-1}{2N}\nu_b\sigma_2^2\right) & \alpha\left(\alpha\sigma_1^2 + \nu_b\sigma_2^2\right) \end{pmatrix},$$

$$K_b = \begin{pmatrix} 0 & \alpha\sigma_1^2 + \frac{2N-2}{2N-1}\sigma_2^2 \\ 0 & -\alpha^2\sigma_1^2 \end{pmatrix} \qquad \Rightarrow \qquad \Sigma_b + K_b = \begin{pmatrix} \sigma_1^2 + \delta\sigma_2^2 & \frac{2N-2}{2N-1}\frac{\sigma_2^2}{1+\alpha x} \\ -\alpha\left(\sigma_1^2 + \frac{2N-1}{2N}\nu_b\sigma_2^2\right) & \alpha\nu_b\sigma_2^2 \end{pmatrix}$$

This gives the equilibrium demand schedules, which are identical across all dealers of the same class i:

$$\begin{pmatrix} q_{i,0}^b(p_0,\pi_b) \\ b_i^*(p_0,\pi_b) \end{pmatrix} = \frac{2N-2}{2N-1}\gamma^{-1}(\Sigma_b)^{-1}\left(M_b\begin{pmatrix} v_0 \\ \frac{\mu_b}{1+\alpha x} \end{pmatrix} - \begin{pmatrix} p_0 \\ \pi_b \end{pmatrix}\right)$$
$$- \frac{2N-2}{2N-1}\left(Id_2 + (\Sigma_b)^{-1}K_b\right)\begin{pmatrix} I_{i,0} \\ \frac{\bar{I}^e_{-i,1}}{2} + \frac{N-1}{N}\frac{\bar{I}^e_{i,1}}{2} \end{pmatrix} \qquad (2.9.23)$$

where Id_2 is the two-dimensional identity matrix. Now I compute $(\Sigma_b)^{-1}$. The determinant of Σ_b is

$$|\Sigma_b| = \alpha\left\{(\sigma_1^2+\delta\sigma_2^2)(\alpha\sigma_1^2+\nu_b\sigma_2^2) - \alpha\left(\sigma_1^2 + \frac{2N-1}{2N}\nu_b\sigma_2^2\right)^2\right\}$$
$$= \alpha\sigma_2^2\left\{\left[\alpha\delta + \nu_b - 2\alpha\left(1-\frac{1}{2N}\right)\nu_b\right]\sigma_1^2 + \left[\delta - \alpha\left(1-\frac{1}{2N}\right)^2\nu_b\right]\nu_b\sigma_2^2\right\}$$

Recognizing $\delta = 1 - \frac{N-1}{N}\frac{1}{1+\alpha x}$, one gets

$$|\Sigma_b| = \alpha\sigma_2^2\left\{\left[\nu_b + \alpha\left(1 - \frac{1}{1+\alpha x} + \frac{1}{N}\frac{1}{1+\alpha x} - \left(2-\frac{1}{N}\right)\frac{\alpha x}{1+\alpha x}\right)\right]\sigma_1^2 + \frac{\nu_b}{N}\sigma_2^2\right\}$$
$$= \alpha\sigma_2^2\left\{\left[(1-\alpha)\nu_b + \frac{1}{N}\right]\sigma_1^2 + \frac{\nu_b}{N}\sigma_2^2\right\}$$

which is strictly positive. Therefore as the first diagonal term of Σ^b is also positive, Σ^b is positive definite and dealer i's problem is strictly concave. Thus

$$(\Sigma_b)^{-1} = |\Sigma^b|^{-1}\begin{pmatrix} \alpha(\alpha\sigma_1^2+\nu_b\sigma_2^2) & \alpha\left(\sigma_1^2 + \left(1-\frac{1}{2N}\right)\nu_b\sigma_2^2\right) \\ \alpha\left(\sigma_1^2 + \left(1-\frac{1}{2N}\right)\nu_b\sigma_2^2\right) & \sigma_1^2+\delta\sigma_2^2 \end{pmatrix}$$
$$= \frac{1}{\sigma_2^2}\begin{pmatrix} \frac{\alpha\sigma_1^2+\nu_b\sigma_2^2}{\left((1-\alpha)\nu_b+\frac{1}{N}\right)\sigma_1^2+\frac{\nu_b}{N}\sigma_2^2} & \frac{\sigma_1^2+\left(1-\frac{1}{2N}\right)\nu_b\sigma_2^2}{\left((1-\alpha)\nu_b+\frac{1}{N}\right)\sigma_1^2+\frac{\nu_b}{N}\sigma_2^2} \\ \frac{\sigma_1^2+\left(1-\frac{1}{2N}\right)\nu_b\sigma_2^2}{\left((1-\alpha)\nu_b+\frac{1}{N}\right)\sigma_1^2+\frac{\nu_b}{N}\sigma_2^2} & \frac{1}{\alpha}\frac{\sigma_1^2+\delta\sigma_2^2}{\left((1-\alpha)\nu_b+\frac{1}{N}\right)\sigma_1^2+\frac{\nu_b}{N}\sigma_2^2} \end{pmatrix}$$

Equilibrium prices. Applying market clearing conditions (3.2.2) and (2.5.1), one gets the equilibrium risk premia:

$$M_b \begin{pmatrix} v_0 \\ \frac{\mu_b}{1+\alpha x} \end{pmatrix} - \begin{pmatrix} p_0^b \\ \pi_b^* \end{pmatrix} = \gamma(\Sigma_b + K_b) \begin{pmatrix} \frac{I_{1,0}+I_{2,0}}{2} \\ \frac{2N-1}{N} \frac{\bar{I}_{1,1}^e + \bar{I}_{2,1}^e}{2} \end{pmatrix} \tag{2.9.24}$$

Explicit computation

Proof of proposition 7

Plugging equilibrium risk premia 2.9.24 into class 1 traders' equilibrium demand schedules 2.9.23:

$$\begin{aligned} \begin{pmatrix} q_{1,0}^b(p_0^b, \pi_b^*) \\ b_1^*(p_0^*, \pi_b^*) \end{pmatrix} &= \frac{2N-2}{2N-1}(Id_2 + \Sigma_b^{-1}K_b) \begin{pmatrix} \frac{I_{1,0}+I_{2,0}}{2} \\ \frac{2N-1}{2N} \frac{\bar{I}_{1,1}^e + \bar{I}_{2,1}^e}{2} \end{pmatrix} \\ &\quad - \frac{2N-2}{2N-1}(Id_2 + \Sigma_b^{-1}K_b) \begin{pmatrix} I_{1,0} \\ \frac{\bar{I}_{2,1}^e}{2} + \frac{N-1}{N}\frac{\bar{I}_{1,1}^e}{2} \end{pmatrix} \\ &= \frac{2N-2}{2N-1}(Id_2 + \Sigma_b^{-1}K_b) \begin{pmatrix} \frac{I_{2,0}-I_{1,0}}{2} \\ \frac{\bar{I}_{1,1}^e - \bar{I}_{2,1}^e}{4N} \end{pmatrix} \end{aligned}$$

With

$$Id_2 + \Sigma_b^{-1}K_b = \begin{pmatrix} 1 & \kappa_1^b \\ 0 & 1+\kappa_2^b \end{pmatrix}$$

one gets, using the equilibrium conditions $q_{i,0}^e = q_{i,0}^*(p_0^*, \pi_b^*)$ and market clearing condition 3.2.2:

$$\begin{pmatrix} q_{1,0}^b(p_1^b, \pi_b^*) \\ b_1^*(p_1^b, \pi_b^*) \end{pmatrix} = \frac{2N-2}{2N-1} \begin{pmatrix} \frac{I_{2,0}-I_{1,0}}{2} + \frac{\kappa_1^b}{2N}\frac{I_{1,0}-I_{2,0}+2q_{1,0}^*}{4} \\ \frac{1+\kappa_2^b}{2N}\frac{I_{1,0}-I_{2,0}+2q_{1,0}^b}{2} \end{pmatrix}$$

This implies

$$q_{1,0}^b\left(1 - \frac{\kappa_1^b}{2N}\frac{2N-2}{2N-1}\right) = \left(1 - \frac{\kappa_1^b}{2N}\right)\frac{2N-2}{2N-1}\frac{I_{2,0}-I_{1,0}}{2}$$

and after rearranging:

$$q_{1,0}^b = \frac{1}{1+A_b(\sigma_q^2)}\frac{I_{2,0}-I_{1,0}}{2} \tag{2.9.25}$$

$$\text{with} \qquad A_b(\sigma_q^2) = \frac{1}{2N-2}\left(1 - \frac{\kappa_1^b}{2N}\right)^{-1}$$

Plugging this into the expression for b_1^* leads to

$$b_1^* = -\frac{2N-2}{2N-1} \times \frac{1+\kappa_2^b}{2N} \times \frac{A_b(\sigma_q^2)}{1+A_b(\sigma_q^2)} \times \frac{I_{2,0}-I_{1,0}}{2} \tag{2.9.26}$$

Computation of $q^b_{1,0}$. One has

$$\Sigma_b^{-1} K_b = \begin{pmatrix} 0 & \kappa_1^b \\ 0 & \kappa_2^b \end{pmatrix}$$

where

$$\begin{aligned}
\kappa_1^b &= \frac{\left(\alpha\sigma_1^2 + \nu_b\sigma_2^2\right)\left(\alpha\sigma_1^2 + \frac{2N-2}{2N-1}\sigma_2^2\right) - \alpha^2\sigma_1^2\left(\sigma_1^2 + \left(1-\frac{1}{2N}\right)\nu_b\sigma_2^2\right)}{\sigma_2^2\left[\left((1-\alpha)\nu_b + \frac{1}{N}\right)\sigma_1^2 + \frac{\nu_b}{N}\sigma_2^2\right]} \\
&= \frac{\left(\alpha\left(1-\alpha\frac{2N-1}{2N}\right)\nu_b + \alpha\frac{2N-2}{2N-1}\right)\sigma_1^2 + \frac{2N-2}{2N-1}\nu_b\sigma_2^2}{\left((1-\alpha\nu_b) + \frac{1}{N}\right)\sigma_1^2 + \frac{\nu_b}{N}\sigma_2^2} \\
&= \frac{\alpha\left(\frac{2N-2}{2N-1} + \left(1-\frac{2N-2}{2N-1}\right)\nu_b\right)\sigma_1^2 + \frac{2N-2}{2N-1}\nu_b\sigma_2^2}{\left((1-\alpha)\nu_b + \frac{1}{N}\right)\sigma_1^2 + \frac{\nu_b}{N}\sigma_2^2} \\
&= \frac{\alpha\frac{2N-2+\nu_b}{2N-1}\sigma_1^2 + \frac{2N-2}{2N-1}\nu_b\sigma_2^2}{\left((1-\alpha)\nu_b + \frac{1}{N}\right)\sigma_1^2 + \frac{\nu_b}{N}\sigma_2^2}
\end{aligned}$$

The second line has used $\alpha\frac{2N-1}{2N} = \frac{2N-2}{2N-1}$. In particular

$$\begin{aligned}
1 - \frac{\kappa_1^b}{2N} &= \frac{\left((1-\alpha)\nu_b + \frac{1}{N} - \frac{\alpha}{2N}\frac{2N-2+\nu_b}{2N-1}\right)\sigma_1^2 + \left(\frac{\nu_b}{N} - \frac{1}{2N}\frac{2N-2}{2N-1}\nu_b\right)\sigma_2^2}{\left((1-\alpha)\nu_b + \frac{1}{N}\right)\sigma_1^2 + \frac{\nu_b}{N}\sigma_2^2} \\
&= \frac{\left(\left(1-\alpha\left(1+\frac{1}{2N(2N-1)}\right)\right)\nu_b + \frac{1}{N}\left(1-\frac{\alpha}{2}\frac{2N-2}{2N-1}\right)\right)\sigma_1^2 + \frac{1}{N}\left(1-\frac{1}{2}\frac{2N-2}{2N-1}\right)\nu_b\sigma_2^2}{\left((1-\alpha)\nu_b + \frac{1}{N}\right)\sigma_1^2 + \frac{\nu_b}{N}\sigma_2^2} \\
&= \frac{\left(\left(1-\alpha\frac{2N-1}{2N}\right)\nu_b + \frac{1}{N}\left(1-\frac{\alpha}{2}\frac{2N-2}{2N-1}\right)\right)\sigma_1^2 + \frac{1}{2N-1}\nu_b\sigma_2^2}{\left((1-\alpha)\nu_b + \frac{1}{N}\right)\sigma_1^2 + \frac{\nu_b}{N}\sigma_2^2} \\
&= \frac{\left(\frac{1}{2N-1}\nu_b + \frac{1}{N}\left(1-\frac{\alpha}{2}\frac{2N-2}{2N-1}\right)\right)\sigma_1^2 + \frac{1}{2N-1}\nu_b\sigma_2^2}{\left((1-\alpha)\nu_b + \frac{1}{N}\right)\sigma_1^2 + \frac{\nu_b}{N}\sigma_2^2}
\end{aligned}$$

so that

$$A_b(\sigma_q^2) = \frac{1}{2N-2} \times \frac{\left((1-\alpha)\nu_b + \frac{1}{N}\right)\sigma_1^2 + \frac{\nu_b}{N}\sigma_2^2}{\left(\frac{1}{2N-1}\nu_b + \frac{1}{N}\left(1-\frac{\alpha}{2}\frac{2N-2}{2N-1}\right)\right)\sigma_1^2 + \frac{1}{2N-1}\nu_b\sigma_2^2} \tag{2.9.27}$$

Together with 2.9.25, this gives equilibrium date 0 quantity traded $q^b_{i,0}$.

Computation of b_i^* The second coefficient κ_2^b of $\Sigma_b^{-1}K_b$ is, recognizing $\alpha\left(1-\frac{1}{2N}\right)=\frac{2N-2}{2N-1}$,

$$\begin{aligned}
\kappa_2^b &= \frac{\left(\sigma_1^2+\left(1-\frac{1}{2N}\right)\sigma_2^2\right)\left(\alpha\sigma_1^2+\frac{2N-2}{2N-1}\sigma_2^2\right)-\alpha^2\sigma_1^2\alpha^{-1}(\sigma_1^2+\delta\sigma_2^2)}{\sigma_2^2\left[\left((1-\alpha)\nu_b+\frac{1}{N}\right)\sigma_1^2+\frac{\nu_b}{N}\sigma_2^2\right]}\\
&= \frac{\left(\frac{2N-2}{2N-1}+\alpha\left(1-\frac{1}{2N}\right)-\alpha\delta\right)\sigma_1^2+\left(1-\frac{1}{2N}\right)\frac{2N-2}{2N-1}\sigma_2^2}{\left((1-\alpha)\nu_b+\frac{1}{N}\right)\sigma_1^2+\frac{\nu_b}{N}\sigma_2^2}\\
&= \frac{\alpha\left(2\left(1-\frac{1}{2N}\right)-1+\frac{N-1}{N}\frac{1}{1+\alpha x}\right)\sigma_1^2+\frac{N-1}{N}\sigma_2^2}{\left((1-\alpha)\nu_b+\frac{1}{N}\right)\sigma_1^2+\frac{\nu_b}{N}\sigma_2^2}\\
&= \frac{N-1}{N}\frac{\alpha\left(1+\frac{1}{1+\alpha x}\right)\sigma_1^2+\sigma_2^2}{\left((1-\alpha)\nu_b+\frac{1}{N}\right)\sigma_1^2+\frac{\nu_b}{N}\sigma_2^2}
\end{aligned}$$

And it is easy to see that $\kappa_2^b>0$. In addition, κ_2^b decreases as σ_q^2 increases, since the numerator depends on σ_q^2 through $\frac{1}{1+\alpha x}$ which decreases with σ_q^2, and the denominator through $\nu_b=\frac{\alpha x}{1+\alpha x}$ which increases as σ_q^2 increases.

Computation of $q_{1,1}^*$. From 3.3.10 and $q_{1,0}^*$, one has

$$q_{1,1}^*=\frac{2N-2}{2N-1}\times\frac{A_b(\sigma_q^2)}{1+A_b(\sigma_q^2)}\times\frac{I_{2,0}-I_{1,0}}{2}$$

Proof of proposition 8

The first part of the proposition is the following lemma, partially proven at this stage.

Lemma 10. *There exists $y_0>0$ such that for $\sigma_1^2/\sigma_2^2<y_0$,*

$$A_b(\sigma_q^2)>A(\sigma_q^2)$$

For $\sigma_1^2/\sigma_2^2>y_0$, there exists $z_0(y)$ such that

$$\begin{aligned}
\frac{1}{1+\alpha x}>z_0 &\quad\Rightarrow\quad A_b(\sigma_q^2)>A(\sigma_q^2)\\
\frac{1}{1+\alpha x}<z_0 &\quad\Rightarrow\quad A_b(\sigma_q^2)<A(\sigma_q^2)
\end{aligned}$$

Proof. The goal is to assess when $A_b(\sigma_q^2)-A(\sigma_q^2)>0$. One has

$$\begin{aligned}
(2N-2)\left(A_b(\sigma_q^2)-A(\sigma_q^2)\right) &= \frac{\left((1-\alpha)\nu_b+\frac{1}{N}\right)\sigma_1^2+\frac{\nu_b}{N}\sigma_2^2}{\left(\frac{\nu_b}{2N-1}+\frac{1}{N}\left(1-\frac{\alpha}{2}\frac{2N-2}{2N-1}\right)\right)\sigma_1^2+\frac{\nu_b}{2N-1}\sigma_2^2}\\
&\quad-\frac{\sigma_1^2+\left(1-\frac{N-1}{N}(1-\nu_b)\right)\sigma_2^2}{\sigma_1^2+\left(1-\frac{2N-2}{2N-1}(1-\nu_b)\right)\sigma_2^2}
\end{aligned}$$

I first look at when $A_b(\sigma_q^2) > A(0)$, as the second inequality comes from lemma 7. Reducing the difference to the same denominator, the sign of $A_b(\sigma_q^2) - A(0)$ is that of the numerator

$$\begin{aligned}
\phi = &\left[(1-\alpha)\nu_b + \frac{1}{N} - \frac{\nu_b}{2N-1} - \frac{1}{N}\left(1 - \frac{\alpha}{2}\frac{2N-2}{2N-1}\right)\right]\sigma_1^4 \\
&+ \left[\frac{\nu_b}{N}\left(1 - \frac{2N-2}{2N-1}(1-\nu_b)\right) - \left(1 - \frac{N-1}{N}(1-\nu_b)\right)\frac{\nu_b}{2N-1}\right]\sigma_2^4 \\
&+ \left[\frac{1}{2N-1}\left(\frac{\nu_b}{2N-1} + \frac{1}{N}\right)\left(1 - \frac{2N-2}{2N-1}(1-\nu_b)\right) + \frac{\nu_b}{N} - \frac{\nu_b}{2N-1}\right. \\
&\qquad \left. - \left(1 - \frac{N-1}{N}(1-\nu_b)\right)\left(\frac{\nu_b}{2N-1} + \frac{1}{N}\left(1 - \frac{\alpha}{2}\frac{2N-2}{2N-1}\right)\right)\right]\sigma_1^2\sigma_2^2 \\
= &\frac{2N-2}{(2N-1)^2}\left[\frac{2N-2}{2N-1} - \nu_b\right]\sigma_1^4 + \frac{2N-2}{2N(2N-1)}\nu_b\sigma_2^4 \\
&+ \left[\frac{1}{N}\left(\frac{1}{2N-1}\left(1 + \left(\frac{2N-2}{2N-1}\right)^2\right) - \frac{1}{N}\right) + \left(\frac{1}{(2N-1)^2} + \frac{1}{N} - \frac{2}{2N-1}\right)\nu_b\right. \\
&\qquad \left. + \frac{2N-2}{2N-1}\left(\frac{1}{2N-1} - \frac{1}{2N}\right)\nu_b^2\right]\sigma_1^2\sigma_2^2
\end{aligned}$$

The second line has used $\alpha = 1 - \frac{1}{(2N-1)^2} = \frac{2N(2N-2)}{(2N-1)^2}$. Plotting the expression above for $y = \frac{\sigma_1^2}{\sigma_2^2}$ and $\nu_b = \frac{\alpha x}{1+\alpha x}$ for various values of $N \geq 2$ gives the conjecture. ∎

The remaining part applies lemma 10 to the equilibrium quantities of propositions 12, 6 and 7, with elementary comparisons.

2.9.6 Welfare analysis

Proof of proposition 9

Proof. The certainty equivalent of wealth at date 0 is for trader 1

$$\widehat{W}_{1,0} = \underbrace{I_{1,0}v_0 + q_{1,0}^*(v_0 - p_0^*) - \frac{\gamma}{2}(\sigma_1^2 + \sigma_2^2)(I_{1,1}^*)^2}_{\widehat{W}_{1,0}^{HTM}} + \underbrace{\frac{2N}{2N-2}\frac{\gamma\sigma_2^2}{1+\alpha x}\left(\mathbb{E}_0\left[q_{1,1}^*\right]\right)^2}_{\widehat{S}_{1,1}^*}$$

Compute the HTM value. One has

$$\begin{aligned}
\widehat{W}_{1,0}^{HTM} &= I_{1,0}v_0 - \frac{\gamma(\sigma_1^2+\sigma_2^2)}{2}(I_{1,0})^2 + q_{1,0}^*(v_0 - p_0^*) - \frac{\gamma}{2}(\sigma_1^2+\sigma_2^2)\left[(I_{1,1}^*)^2 - (I_{1,0})^2\right] \\
&= I_{1,0}v_0 - \frac{\gamma(\sigma_1^2+\sigma_2^2)}{2}(I_{1,0})^2 + q_{1,0}^*\left(v_0 - p_0^* - \frac{\gamma}{2}(\sigma_1^2+\sigma_2^2)(2I_{1,0} + q_{1,0}^*)\right) \\
&= I_{1,0}v_0 - \frac{\gamma(\sigma_1^2+\sigma_2^2)}{2}(I_{1,0})^2 + q_{1,0}^*\left(\gamma(\sigma_1^2+\sigma_2^2)\frac{I_{2,0} - I_{1,0} - q_{1,0}^*}{2} + \frac{\gamma\sigma_2^2}{1+\alpha x}\mathbb{E}_0\left[\frac{Q}{2N}\right]\right)
\end{aligned}$$

Denoting $A \equiv A(\sigma_q^2)$ to ease notation and plugging equilibrium quantity 3.7.7:

$$\begin{aligned}
\widehat{W}_{1,0}^{HTM} &= I_{1,0}v_0 - \frac{\gamma(\sigma_1^2+\sigma_2^2)}{2}(I_{1,0})^2 \\
&\quad + \frac{1}{1+A}\frac{I_{2,0}-I_{1,0}}{2}\left(\gamma(\sigma_1^2+\sigma_2^2)\left(1-\frac{1}{2}\frac{1}{1+A}\right)\frac{I_{2,0}-I_{1,0}}{2} + \frac{\gamma\sigma_2^2}{1+\alpha x}\,\mathbb{E}_0\left[\frac{Q}{2N}\right]\right) \\
&= I_{1,0}v_0 - \frac{\gamma(\sigma_1^2+\sigma_2^2)}{2}(I_{1,0})^2 \\
&\quad + \frac{1}{1+A}\frac{I_{2,0}-I_{1,0}}{2}\left(\frac{\gamma(\sigma_1^2+\sigma_2^2)}{2}\frac{1+2A}{1+A}\frac{I_{2,0}-I_{1,0}}{2} + \frac{\gamma\sigma_2^2}{1+\alpha x}\,\mathbb{E}_0\left[\frac{Q}{2N}\right]\right)
\end{aligned}$$

This can also be re-written:

$$\begin{aligned}
\widehat{W}_{1,0}^{HTM} &= I_{1,0}v_0 - \frac{\gamma(\sigma_1^2+\sigma_2^2)}{2}(I_{i,0})^2 \\
&\quad + \underbrace{\frac{\gamma(\sigma_1^2+\sigma_2^2)}{2}(1+2A)\left(q_{1,0}^*\right)^2 + \frac{\gamma\sigma_2^2}{1+\alpha x}q_{1,0}^*\,\mathbb{E}_0\left[\frac{Q}{2N}\right]}_{S_{1,0}^*}
\end{aligned}$$

■

Proof of proposition 10

Proof. The date 0 certainty equivalent of wealth is, omitting the constant $1/2\gamma\ln(1+\alpha x)$:

$$\begin{aligned}
\widehat{W}_{1,0}^{a} &= I_{1,0}v_0 + q_{i,0}^a(v_0-p_0^*) - a_1^*\left(\pi_a^* + \alpha\gamma\sigma_2^2\frac{I_{1,0}+I_{2,0}}{2}\right) \\
&\quad - \frac{\gamma}{2}\left(\left(\sigma_1^2+(1-\alpha)\sigma_2^2\right)(I_{i,1}^a)^2 + \alpha\sigma_2^2(I_{1,1}^a-a_1^*)^2\right) \\
&\quad + \frac{N}{2N-2}\frac{\gamma\sigma_2^2}{1+\alpha x}\left(\mathbb{E}_0\left[q_{1,1}^a\right] + \frac{2N-2}{2N-1}a_i^*\right)^2 \\
&= I_{1,0}v_0 + q_{i,0}^a(v_0-p_0^*) - a_1^*\left(\pi_a^* + \alpha\gamma\sigma_2^2\frac{I_{1,0}+I_{2,0}}{2}\right) \\
&\quad - \frac{\gamma}{2}\left(\left(\sigma_1^2+\sigma_2^2\right)(I_{i,1}^a)^2 + \alpha\sigma_2^2 a_1^*(a_1^*-2I_{1,1}^a)\right) \\
&\quad + \frac{N}{2N-2}\frac{\gamma\sigma_2^2}{1+\alpha x}\left(\mathbb{E}_0\left[q_{1,1}^a\right] + \frac{2N-2}{2N-1}a_i^*\right)^2
\end{aligned}$$

Now denote $A \equiv A(0)$ to ease notation. First notice that $a_1^* = -\frac{\eta}{2} q_{1,1}^{ID,a}$ and write $q_{1,1}^a = q_{1,1}^{ID,a} + \frac{Q}{2N}$. Also plug the date 0 equilibrium prices to get

$$\begin{aligned}
\widehat{W}_{1,0}^a &= I_{1,0}v_0 - \frac{\gamma(\sigma_1^2+\sigma_2^2)}{2}(I_{1,0})^2 + q_{i,0}^a(v_0 - p_0^*) - \frac{\gamma(\sigma_1^2+\sigma_2^2)}{2}\left((I_{1,0}+q_{i,0}^a)^2 - (I_{1,0})^2\right) \\
&\quad - a_1^*\left(\pi_a^* + \alpha\gamma\sigma_2^2 \frac{I_{1,0}+I_{2,0}}{2} + \alpha\gamma\sigma_2^2\left(\frac{a_1^*}{2} - I_{1,0} - q_{1,0}^a\right)\right) \\
&\quad + \frac{N}{2N-2}\frac{\gamma\sigma_2^2}{1+\alpha x}\left(\left(1 - \frac{2N-2}{2N-1}\frac{\eta}{2}\right) q_{1,1}^{ID,a} + \mathbb{E}_0\left[\frac{Q}{2N}\right]\right)^2 \\
&= I_{1,0}v_0 + q_{i,0}^a\left(\gamma(\sigma_1^2+\sigma_2^2)\frac{I_{2,0}+I_{1,0}}{2} + \frac{\gamma\sigma_2^2}{1+\alpha x}\mathbb{E}_0\left[\frac{Q}{2N}\right] + \frac{\gamma(\sigma_1^2+\sigma_2^2)}{2}\left(2I_{1,0} + q_{i,0}^a)\right)\right) \\
&\quad + \frac{\eta}{2} q_{1,1}^{ID,a}\left(\frac{2N}{2N-2}\frac{\gamma\sigma_2^2}{1+\alpha x}\mathbb{E}_0\left[\frac{Q}{2N}\right] + \alpha\gamma\sigma_2^2\left(\frac{I_{2,0}-I_{1,0}}{2} - q_{1,0}^a - \frac{\eta}{4}q_{1,1}^{ID,a}\right)\right) \\
&\quad + \frac{N}{2N-2}\frac{\gamma\sigma_2^2}{1+\alpha x}\left(\left(1 - \frac{2N-2}{2N-1}\frac{\eta}{2}\right) q_{1,1}^{ID,a} + \mathbb{E}_0\left[\frac{Q}{2N}\right]\right)^2 \\
&= I_{1,0}v_0 + q_{i,0}^a\gamma(\sigma_1^2+\sigma_2^2)\left(1 - \frac{1}{2(1+A)}\right)\frac{I_{2,0}-I_{1,0}}{2} + q_{1,0}^a\frac{\gamma\sigma_2^2}{1+\alpha x}\mathbb{E}_0\left[\frac{Q}{2N}\right] \\
&\quad + \frac{\eta}{2} q_{1,1}^{ID,a}\left(\frac{2N}{2N-2}\frac{\gamma\sigma_2^2}{1+\alpha x}\mathbb{E}_0\left[\frac{Q}{2N}\right] + \frac{2N}{2N-1}\gamma\sigma_2^2\left(1 - \frac{2N-2}{2N-1}\frac{\eta}{4}\right) q_{1,1}^{ID,a}\right) \\
&\quad + \frac{N}{2N-2}\frac{\gamma\sigma_2^2}{1+\alpha x}\left(\left(1 - \frac{2N-2}{2N-1}\frac{\eta}{2}\right) q_{1,1}^{ID,a} + \mathbb{E}_0\left[\frac{Q}{2N}\right]\right)^2 \\
&= I_{1,0}v_0 + \frac{\gamma(\sigma_1^2+\sigma_2^2)}{2}\frac{1+2A}{(1+A)^2}\frac{I_{2,0}-I_{1,0}}{2} + \left(q_{1,0}^a + \frac{N\eta}{2N-2}q_{1,1}^{ID,a}\right)\frac{\gamma\sigma_2^2}{1+\alpha x}\mathbb{E}_0\left[\frac{Q}{2N}\right] \\
&\quad + \frac{N\eta}{2N-1}\gamma\sigma_2^2\left(1 - \frac{2N-2}{2N-1}\frac{\eta}{4}\right)\left(q_{1,1}^{ID,a}\right)^2 \\
&\quad + \frac{N}{2N-2}\frac{\gamma\sigma_2^2}{1+\alpha x}\left(\left(1 - \frac{2N-2}{2N-1}\frac{\eta}{2}\right) q_{1,1}^{ID,a} + \mathbb{E}_0\left[\frac{Q}{2N}\right]\right)^2
\end{aligned}$$

■

Proof of theorem 2

Denote $\Delta\widehat{W}_{i,0}^a \equiv \widehat{W}_{i,0}^a - \widehat{W}_{i,0}^*$ and assume $\mathbb{E}_0[Q] = 0$. One has, observing that $\frac{1+2A}{(1+A)^2} = 1 - \left(\frac{A}{1+A}\right)^2$ and denoting for convenience $A \equiv A(\sigma_q^2)$,

$$\begin{aligned}
\Delta\widehat{W}_{1,0}^a &= \frac{\gamma\sigma_2^2}{2}\left\{-\left(\frac{A(0)}{1+A(0)}\right)^2\left[y + 1 - \alpha z - \alpha(1-z)\frac{2N-2}{2N-1}\eta\left(1 - \frac{2N-2}{2N-1}\frac{\eta}{4}\right)\right]\right. \\
&\qquad \left. + \left(\frac{A}{1+A}\right)^2[y + 1 - \alpha z]\right\}\left(\frac{I_{2,0}-I_{1,0}}{2}\right)^2 \qquad (2.9.28) \\
&= \frac{\gamma\sigma_2^2}{2}\left(\frac{A(0)}{1+A(0)}\right)^2\left(\frac{I_{2,0}-I_{1,0}}{2}\right)^2\psi(y,z,N) \qquad (2.9.29)
\end{aligned}$$

where $y = \sigma_1^2/\sigma_2^2$ and $z = \frac{1}{1+\alpha x} \in (0,1]$ and, given

$$\frac{A(\sigma_q^2)}{1+A(\sigma_q^2)} = \frac{1}{1+(2N-2)\frac{y+1-\frac{\alpha}{2}z}{y+1-\frac{\alpha}{4}z}}.$$

it is possible to write ψ as

$$\psi(y,z,N) = \left[(\phi(y,z,N))^2 - 1\right](y+1-\alpha z) + \alpha(1-z)\eta\left(1 - \frac{2N-2}{2N-1}\frac{\eta}{4}\right)$$

$$\text{with} \quad \begin{cases} \phi(y,z,N) & = \frac{1+(2N-2)R(y,1)}{1+(2N-2)R(y,z)} \\ R(y,z) & = \frac{y+1-\frac{\alpha}{2}z}{y+1-\frac{\alpha}{4}z} \\ \eta & = \frac{1}{N}\frac{y+1}{y+1/N} \end{cases}$$

Now I prove that $\psi(y,z,N) > 0$ for all $y \geq 0$, $z \in (0,1]$ and $N \geq 2$. First

$$\phi^2 - 1 = (\phi+1)(\phi-1)$$

and rewrite

$$\begin{aligned}
\phi(y,z,N) &= \frac{y+1-\frac{N-1}{N}z}{y+\frac{1}{N}} \times \frac{y+1-\frac{N-1}{N}+(2N-2)\left(y+1-\left(1+\frac{1}{2N-1}\right)\frac{N-1}{N}\right)}{y+1-\frac{N-1}{N}z+(2N-2)\left(y+1-\left(1+\frac{1}{2N-1}\right)\frac{N-1}{N}\right)z} \\
&= \frac{y+1-\frac{N-1}{N}z}{y+\frac{1}{N}} \times \frac{y+\frac{1}{N}-\left(\frac{2N-2}{2N-1}\right)^2\frac{1}{2N}}{y+1-\frac{N-1}{N}z-\left(\frac{2N-2}{2N-1}\right)^2\frac{1}{2N}z}
\end{aligned}$$

so that

$$\begin{aligned}
\phi - 1 &= \frac{\left(y+1-\frac{N-1}{N}z\right)\left(y+\frac{1}{N}-\left(\frac{2N-2}{2N-1}\right)^2\frac{1}{2N}\right)-\left(y+\frac{1}{N}\right)\left(y+1-\frac{N-1}{N}z-\left(\frac{2N-2}{2N-1}\right)^2\frac{1}{2N}z\right)}{\left(y+\frac{1}{N}\right)\left(y+1-\frac{N-1}{N}z-\left(\frac{2N-2}{2N-1}\right)^2\frac{1}{2N}z\right)} \\
&= \frac{-(y+1)\left(\frac{2N-2}{2N-1}\right)^2\frac{1}{2N}-\left(y+\frac{1}{N}\right)\frac{N-1}{N}z+\frac{N-1}{N}\left(\frac{2N-2}{2N-1}\right)^2\frac{1}{2N}z+\left(y+\frac{1}{N}\right)\frac{N-1}{N}\left(1+\frac{2N-2}{(2N-1)^2}\right)z}{\left(y+\frac{1}{N}\right)\left(y+1-\frac{N-1}{N}z-\left(\frac{2N-2}{2N-1}\right)^2\frac{1}{2N}z\right)} \\
&= \frac{-(y+1)\left(\frac{2N-2}{2N-1}\right)^2\frac{1}{2N}+\frac{N-1}{N}\left(\frac{2N-2}{2N-1}\right)^2\frac{1}{2N}z+\left(y+\frac{1}{N}\right)\frac{2N-2}{2N}\frac{2N-2}{(2N-1)^2}z}{\left(y+\frac{1}{N}\right)\left(y+1-\frac{N-1}{N}z-\left(\frac{2N-2}{2N-1}\right)^2\frac{1}{2N}z\right)} \\
&= \left(\frac{2N-2}{2N-1}\right)^2\frac{1}{2N}\frac{-(y+1)+\frac{N-1}{N}z+\left(y+\frac{1}{N}\right)z}{\left(y+\frac{1}{N}\right)\left(y+1-\frac{N-1}{N}z-\left(\frac{2N-2}{2N-1}\right)^2\frac{1}{2N}z\right)} \\
&= \frac{1}{2}\left(\frac{2N-2}{2N-1}\right)^2\frac{1}{N}\frac{y+1}{y+\frac{1}{N}}\frac{1-z}{\left(y+\frac{1}{N}\right)\left(y+1-\frac{N-1}{N}z-\left(\frac{2N-2}{2N-1}\right)^2\frac{1}{2N}z\right)}
\end{aligned}$$

which yields

$$(\phi^2-1)(y+1-\alpha z) = \frac{1+\phi}{2}\left(\frac{2N-2}{2N-1}\right)^2 \frac{1}{N}\frac{y+1}{y+\frac{1}{N}} \frac{(1-z)(y+1-\alpha z)}{\left(y+\frac{1}{N}\right)\left(y+1-\frac{N-1}{N}z-\left(\frac{2N-2}{2N-1}\right)^2\frac{1}{2N}z\right)}$$

Thus

$$\begin{aligned}\psi = & \left(\frac{2N-2}{2N-1}\right)^2 \frac{1-z}{N}\frac{y+1}{y+\frac{1}{N}}\left\{\frac{2N}{2N-1}\left(1-\frac{2N-2}{2N-1}\frac{1}{4N}\frac{y+1}{y+\frac{1}{N}}\right)\right. \\ & \left. -\frac{1+\phi}{2}\frac{y+1-\alpha z}{y+1-\frac{N-1}{N}\left(1-\left(\frac{2N-2}{2N-1}\right)^2\right)z}\right\} \\ = & \left(\frac{2N-2}{2N-1}\right)^2 \frac{1-z}{N}\frac{y+1}{y+\frac{1}{N}}\left\{1+\frac{1}{2N-1}-\frac{1}{2}\frac{2N-2}{(2N-1)^2}\frac{y+1}{y+\frac{1}{N}}\right. \\ & \left. -\frac{1+\phi}{2}\frac{y+1-\alpha z}{y+1-\frac{N-1}{N}\left(1-\left(\frac{2N-2}{2N-1}\right)^2\right)z}\right\}\end{aligned}$$

If $z=1$, *i.e.* $\sigma_q^2 = 0$, then $\psi = 0$. If $z<1$, the sign of ψ is that of the expression inside the brackets. First, $\frac{y+1}{y+1/N}$ has its maximum at $y=0$ and then equals N, so that

$$\frac{2N-2}{(2N-1)^2}\frac{y+1}{y+1/N} \leq \frac{\alpha}{2}.$$

For the term on the second line, first observe that $\phi < 1$, which comes from the easily proven fact that $R(y,z)$ decreases with z, so that $(1+\phi)/2) < 1$. Then, since

$$\alpha = \frac{2N(2N-2)}{(2N-1)^2} < \frac{2N-2}{2N}\left(1-\left(\frac{2N-2}{2N-1}\right)^2\right)$$

one has

$$\begin{aligned}\frac{y+1-\alpha z}{y+1-\frac{N-1}{N}\left(1-\left(\frac{2N-2}{2N-1}\right)^2\right)z} & \leq \frac{1-\alpha z}{1-\frac{N-1}{N}\left(1-\left(\frac{2N-2}{2N-1}\right)^2\right)z} \\ & \leq \frac{1-\alpha}{1-\frac{N-1}{N}\left(1-\left(\frac{2N-2}{2N-1}\right)^2\right)} \\ & = \frac{1}{(2N-1)^2-\frac{N-1}{N}\left((2N-1)^2-(2N-2)^2\right)} \\ & = \frac{1}{\frac{1}{N}(2N-1)^2+\frac{N-1}{N}(2N-2)^2} \\ & \leq \frac{1}{(2N-2)^2} \leq \frac{1}{2}\end{aligned}$$

Therefore

$$\psi \geq \left(\frac{2N-2}{2N-1}\right)^2 \frac{1-z}{N} \frac{y+1}{y+\frac{1}{N}} \left\{1 + \frac{1}{2N-1} - \frac{\alpha}{4} - \frac{1}{2}\right\}$$

Since $\alpha < 1$, $1 - \alpha/4 - 1/2 > 0$ and $\psi(y, z, N) > 0$. QED.

2.9.7 A technical lemma

Let $X \sim \mathcal{N}(\mu, \Sigma)$ a normal vector of dimension p ($|\Sigma| > 0$), and A a symmetric matrix. Then one seeks to compute $\mathbb{E}[\exp(-\gamma X'AX)]$ where A is a symmetric matrix.

Lemma 11. *Suppose $I + 2\gamma A\Sigma$ is positive definite, then*

$$\mathbb{E}[\exp(-\gamma X'AX)] = \frac{1}{\sqrt{|I + 2\gamma A\Sigma|}} \exp\left\{-\gamma\mu'(I + 2\gamma A\Sigma)^{-1}A\mu\right\}$$

Proof.

$$\mathbb{E}[\exp(-\gamma X'AX)] = \int_{\mathbb{R}^p} \frac{1}{\sqrt{2\pi|\Sigma|}} \exp\left\{-\gamma x'Ax - \frac{1}{2}(x-\mu)'\Sigma^{-1}(x-\mu)\right\} dx$$

where $dx \equiv dx_1 dx_2 ... dx_p$. One first computes

$$\begin{aligned} Q(x) &= -\gamma x'Ax - \frac{1}{2}(x-\mu)'\Sigma^{-1}(x-\mu) \\ &= -\frac{1}{2}(x-\mu)'(\Sigma^{-1} + 2\gamma A)(x-\mu) - 2\gamma\mu'A(x-\mu) - \gamma\mu'A\mu \end{aligned}$$

Suppose that $(\Sigma^{-1} + 2\gamma A)$ is the inverse of a covariance matrix, then the formula will give almost the moment generating function of a normal variable with covariance matrix $[(I + 2\gamma A\Sigma)\Sigma^{-1}]^{-1} = \Sigma(I + 2\gamma A\Sigma)^{-1}$.

$$\begin{aligned} \mathbb{E}\left[e^{-\gamma X'AX}\right] &= \frac{e^{-\gamma\mu'A\mu}}{\sqrt{|I + 2\gamma A\Sigma|}} \int_{\mathbb{R}^p} \frac{1}{\sqrt{2\pi|\Sigma||I + 2\gamma A\Sigma|^{-1}}} e^{-2\gamma\mu'A(x-\mu)} e^{-\frac{1}{2}(x-\mu)'[\Sigma(I+2\gamma A\Sigma)^{-1}]^{-1}(x-\mu)} dx \\ &= \frac{1}{\sqrt{|I + 2\gamma A\Sigma|}} \exp\left\{\gamma\mu'A\Sigma(I + 2\gamma A\Sigma)^{-1}A\mu - \gamma\mu'A\mu\right\} \\ &= \frac{1}{\sqrt{|I + 2\gamma A\Sigma|}} \exp\left\{\gamma\mu'\left(2\gamma A\Sigma - (I + 2\gamma A\Sigma)\right)(I + 2\gamma A\Sigma)^{-1}A\mu\right\} \\ &= \frac{1}{\sqrt{|I + 2\gamma A\Sigma|}} \exp\left\{-\gamma\mu'(I + 2\gamma A\Sigma)^{-1}A\mu\right\} \end{aligned}$$

∎

Chapter 3

Dynamic Trading and Endogenous Market Fragmentation

Abstract

I study the opportunity for dealers, *i.e.* intermediaries in financial markets to open restricted markets parallel to a centralized, all-to-all market. In a dynamic trading model with imperfect competition, dealers have the opportunity to open a parallel market so that a restricted subset of them trades with customers. Dealers in the parallel market choose to have all customer trades in the parallel market, which makes both customers and dealers not trading in the parallel market worse off. Before dealers learn whether they will have an opportunity to trade with customers in the parallel market, they choose to open the parallel market, as long as the surplus from future transactions are sufficiently high compared with the cost of holding the asset until future transaction, highlighting the role of dynamic trading rent.

3.1 Introduction

In many asset markets, traders do not meet all in a single venue, but trade occurs within and across small groups forming a network. Examples include bond or swap markets where investors deal with a single market maker at a time. There have been large debate on whether such structure is good from a social point of view, and on how these network form. For instance, after the 2007-2009 crisis, the Dodd-Frank Act imposed the possibility for investors to trade swaps in centralized markets. But there is evidence that some investors have continued to trade over-the-counter instead (Collin-Dufresne et al. 2018).

In this paper I study a dynamic trading model in which some agents, called dealers, optimally choose to open a parallel market to trade with customers, because it allows a subset of them to create inventory imbalances with other dealers, and earn trading rents. Even if dealers not trading in the parallel market are worse off with parallel trading, they choose to open the market if ex ante they do not know if they will be able to trade in the parallel market. When the parallel market is open, customers trade at a worse price than in the centralized market, because there is less competition in the market with less dealers. Therefore market fragmentation in this case is Pareto dominated.

I study a model in which risk averse traders (dealers), can trade a risky asset in a centralized market at two dates 0 and 1. Before the date 0 market opens, dealers choose whether they open a parallel market, before they know whether they will actually be in the subset that participates in the parallel market. Dealers are assigned to one group or the other with equal probability. This timing captures the idea that the market structure is given before customers express potential trading needs to their dealers. Then customers disclose the quantity they want to trade, and dealers participating in the parallel market decide what fraction of the order they want customers to trade with them, and what fraction in the centralized market. After parallel trading, all dealers, and the customer if required to trade in the parallel market, meet in the date 0 centralized market; all dealers again meet at date 1 in the centralized market. Before date 1 trading, some public news about the terminal payoff of the asset arrives.

I also assume that dealers care about the price impact of their trades, *i.e.* there is imperfect competition: as in Kyle (1989) or Vayanos (1999), I look for Nash equilibria in demand schedules. Under perfect competition, it is easy to show that dealers and customers would be indifferent between opening and not opening the parallel market, provided nothing happens between the two markets.

I first compute the trading rent associated with progressive liquidation of inventory imbalances. To do this I solve the equilibrium at dates 0 and 1 in the centralized market. Because of imperfect competition, dealers realize gains from trade progressively, and they do not realize all gains from trade at the end of date 1. The utility of a dealer can be nicely decomposed into a static, or hold to maturity component, that gives the utility of the dealer if there was no interim trading opportunity, and into a trading rent component that corresponds to the surplus of all date 0 and date

1 transactions.

To assess the opportunity for dealers of each class, those who trade in the parallel market, and those who do not, I solve the equilibrium in the parallel market. All dealers start with the same zero initial inventory. The zero initial position is without loss of generality. In a first step, I assume that dealers take as given the quantity available in the parallel market and the quantity posted by customers in the centralized market. Dealers charge a higher spread to the customers in the parallel market than in the centralized market, because there are less dealers, thus less competition, in the parallel market.

In a second step, I compute each class of dealers' preference over the fraction of customers' quantity to be traded in the parallel market. Dealers trading in the parallel market prefer that all customer trading happens in the parallel market. This is in particular because in the parallel market, competition is softened due to the lower number of dealers. By contrast, dealers not participating in the parallel market prefer that all customer trading occurs in the centralized market, because they get higher rents in the centralized market.

Ex ante, when dealers decide their market structure before they know if their customers will have trading needs, dealers thus face a risk: if they vote for the opening of the centralized market, they face the risk of not being able to decide to direct customer trades to the centralized market and thus to get lower utility; but if they are in the group trading in the parallel market, they get a higher rent. I show that the expected payoff from opening the parallel market is positive. However, dealers are risk averse, so must weigh the expected utility from opening the parallel market versus the expected utility of ex ante directing all customer trades to the centralized market.

I show that when the variance of public news arriving at date 1 is low relative to the total variance of the terminal payoff, dealers choose to open the parallel market. When the variance of date 1 news is relatively high, they choose to direct all customer trades to the centralized market. This suggests that dealers' ability to trade the position they have built in the parallel market in a dynamic way is crucial to the decision to open the market: when variance of date 1 public news, the cost of holding the asset until date 1 is high relative to the date 1 surplus and the date 0 market looks more like a static market. A higher variance of date 1 news may result from a more remote date 1 trading date: therefore, in this model market fragmentation results from the ability of dealers to dynamically trade with trading rounds in a limited period of time.

Opening the parallel market is Pareto-dominated in this setting, as customers trade at a worse price than in the centralized market.

As a robustness check, I study the case where trade in the parallel market and in the date 0 centralized market occur simultaneously rather than sequentially, using the methods of Malamud and Rostek (2017). I show that this results in higher apparent competition in the parallel market, while the equilibrium is unchanged in the centralized market. Consistently with the unchanged centralized market equilibrium, dealers not trading in the parallel market still prefer that customers trade in the parallel market. Dealers in trading in the parallel market still prefer that all customer trading occurs in the parallel market. Before dealers learn whether they will trade in the parallel market, the expected payoff of opening the parallel market is still positive, and for risk aversion not too high, it is still preferable for dealers to open the parallel market.

Literature review. Relevant papers include Duffie et al. (2005), Atkeson et al. (2015) Dugast et al. (2019). Vayanos (1999), Duffie and Zhu (2017), Antill and Duffie (2018), Rostek and Weretka (2015). This paper differs from Malamud and Rostek (2017) in that I endogenize the network structure, while they assess the welfare implication of given network structures.

3.2 Setting

There are three dates $t = 0, 1, 2$. There is one risky asset that pays off at $t = 2$ an ex ante unknown amount v per unit. At each date, before any action takes place, a public signal ϵ_t is released: ϵ_1 and ϵ_2 are independent and normally distributed with mean 0 and respective variances σ_1^2 and σ_2^2. Thus $v = v_0 + \epsilon_1 + \epsilon_2$ and we denote v_t the expectation of v conditional on information released at t. There is also a riskless asset (cash) that can be purchased or sold without constraint by a perfectly elastic supplier. We normalize its gross return to 1.

There are two types of traders $i = 1, 2$, which I call *dealers* for concreteness; there are also customers to be described shortly. Dealers of class i maximize the expected utility of their terminal wealth W_i described below. Each class contains $N \geq 2$ dealers[1]. Their utility is negative exponential (CARA), with risk aversion parameter $\gamma > 0$ for both classes. Dealers of class i all start with zero initial inventory of the risky asset. Dealers are all forward-looking and fully rational: in particular, they perfectly anticipate at date 0 the date 1 equilibrium and adjust their actions accordingly. When considering dealers of class i, I use the notation $-i$ to refer to dealers of the other class.

Dealers can meet at $t = 0$ and $t = 1$ in a centralized market. At date $t = -1$, they vote on the following alternative:

- either they admit customers to a trade with one class i of dealers, customers truthfully disclose the size Q_0 of their order and dealers decide what fraction w of Q_0 they take in the parallel market, and what fraction $1 - w$ they leave for the centralized market

- or they do not open the parallel market.

The assignment to one class of dealers, rather than with a strategically chosen number of dealers, is for simplicity.

[1]This makes the total number of traders in each market greater than 3, a necessary condition to have equilibrium in linear strategies. When there are only two traders, Du and Zhu (2017) show existence of equilibria in *non-linear* strategies.

Importantly, when voting, dealers do not know if they will be in the group trading in the parallel market, or not. Therefore, as dealers are ex ante symmetric, if they are not ex ante indifferent between customers trading in the centralized market and them trading in the parallel market, then unanimity obtains so that the voting rule does not matter.

The random assignment of customer trades to a class of dealers may capture the fact that each class of dealers has its own pool of customers, and customers' trading need in each pool occur randomly; it therefore does not necessarily imply that customers randomly chose the dealers they trade with.

At date 0, if customers trade in the parallel market, they trade a quantity Q_d with the N dealers they are assigned. These dealers post demand schedules, and a walrasian auctioneer computes a market clearing price p_d^*. I look for equilibria in which dealers with the same characteristics post the same demand schedules, which rules out mixed strategies: as dealers differ only by initial inventory, depending on whether they trade or not in the parallel market, all dealers of class i post the same demand schedule $q_{i,0}^*(p_d)$ and therefore purchase the same equilibrium quantity $q_{i,d}^* \equiv q_{i,d}^*(p_d^*)$. The market clearing condition in the parallel market is therefore, assuming without loss of generality that dealers of class 1 are assigned the trade:

$$q_{1,d}^*(p_0^*) = \frac{Q_d}{N} \tag{3.2.1}$$

Then all $2N$ dealers meet in a centralized market with inventory $I_{i,0} = 0$ if they have traded with customers, and $I_{i,0} = q_{i,d}^*$. They post demand schedules in a similar way to the parallel market. A walrasian auctioneer computes the equilibrium price p_0^* that clears the market. I also allow for customers to post an arbitrary quantity Q_c in the centralized market. If dealers have denied access of the centralized market to customers, then $Q_c = 0$. The market clearing condition at date 0 is thus $Nq_{1,0}^*(p_0^*) + Nq_{2,0}^*(p_0^*) = Q_c$, *i.e.*

$$q_{1,0}^*(p_0^*) + q_{2,0}^*(p_0^*) = \frac{Q_c}{N} \tag{3.2.2}$$

Viewing $Q_c + Q_d = Q_0$ as a single customer order, I forbid customers to trade in one direction in the centralized market, and in the other direction in the parallel market. Therefore I denote $Q_d = wQ_0$ and $Q_c = (1-w)Q_0$, with the constraints that $w \in [0,1]$.

At date 1, the centralized market re-opens, and an external infinitely risk averse customer has a liquidity shock and is willing to sell Q units of the security (thus when $Q > 0$, the customer is willing to sell and vice versa).[2] Again the market is walrasian, with market clearing price p_1^*, and the assumption about date 1 competitiveness of dealers is naturally consistent with that of date 0. dealers arrive in the date 1 market

[2]The analysis is isomorphic to that of a liquidity shock of size Q/N that hits each trader of a given class i (e.g. a customer who has traded only with traders of class i), although the equilibrium allocations and prices differ in the imperfect competition setting.

with inventories $I_{i,1} = I_{i,0} + q_{i,0}$, where $q_{i,0}$ is the quantity traded at date 0. Market clearing at $t = 1$ thus writes $Nq^*_{1,1}(p^*_1) + Nq^*_{2,1}(p^*_1) = Q$, *i.e.*

$$q^*_{1,1}(p^*_1) + q^*_{2,1}(p^*_1) = \frac{Q}{N} \tag{3.2.3}$$

Dealers of both classes do not know the value of Q at date 0, and it is publicly revealed simultaneously with ϵ_1 at date 1 before the market opens. A public noisy signal on Q is released at date 0 before the market opens (thus there is no information asymmetry): all dealers share the common belief at date 0 that Q is normally distributed with mean $\mathbb{E}_0[Q] = 0$ and variance $N^2\sigma_q^2$. In addition, we assume that Q is independent of ϵ_1 and ϵ_2, and that this is common knowledge. Independence of Q and ϵ_t means that Q is a pure private value or liquidity shock: this captures the real life feature that an investor may sometime need cash when other market participants don't, *e.g.* a mutual fund or a life insurer facing indiosyncratic withdrawals, or that the investor has got news on its future cash needs and adjusts his portfolio (maturity, liquidity, ...) accordingly.

With initial inventory $I_{i,0}$, quantities $q_{i,0}$ and $q_{i,1}$ purchased at $t = 0$ and $t = 1$ at respective prices p_0 and p_1, the terminal wealth of class i traders is

$$W_i = I_{i,0}v + q_{i,0}(v - p_0) + q_{i,1}(v - p_1) \tag{3.2.4}$$

Dealers are imperfectly competitive in that they manage the price impacts of their trades. I look for Nash equilibria in demand schedules in the date 1 market. As there is no asymmetric information, there is an equilibrium multiplicity problem (Klemperer and Meyer 1989). I use the usual trembling-hand stability criterion to select a unique equilibrium (*cf.* Vayanos 1999).

I solve for equilibrium by backward induction, consistently with dealer full rationality. In particular, at date 0 dealers have rational expectations on how date 0 equilibrium impacts date 1 equilibrium.

3.3 Equilibrium in the centralized market

3.3.1 Date 1

In this section, I look for Nash equilibria in demand schedules in the date 1 market. As there is no asymmetric information, there is an equilibrium multiplicity problem (Klemperer and Meyer 1989). I use the usual trembling-hand stability criterion to select a unique equilibrium (*cf.* Vayanos 1999).

At date 1, the expected utility to be maximized by trader k in class i is given by (2.2.4). By contrast with competitive markets, traders now take the impact of their demand on the equilibrium price into account: they conjecture the equilibrium residual demand curve that is the sum of all other traders' demand curves. For a given quantity $q_{i,1}$ demanded by trader i, this residual demand curve implies an equilibrium price p_1, and a marginal increase in the quantity demanded by trader i

implies a marginal price impact $\partial p_1/\partial q_{k,i,1}$. Differentiating the certainty equivalent of wealth (2.2.4), trader k' first order condition is:

$$v_1 - p_1 - q_{k,i,1}\frac{\partial p_1}{\partial q_{k,i,1}} = \gamma\sigma_2^2(I_{i,1} + q_{k,i,1})$$

Following the literature on market with Nash equilibria in demand schedules,[3] I conjecture linear strategies in equilibrium: there optimal strategies are linear whenever other market participants use linear strategies, but individual strategies are not constrained to be linear. Thus trader k in class i expects to face a linear residual demand curve of conjectured slope $1/\lambda_{k,i,1}$, so that $\partial p_1/\partial q_{i,1} = \lambda_{k,i,1}$. To ease notation I slightly anticipate on the equilibrium result that all traders k within class i follow symmetric strategies, thus I drop the k subscript, so that her optimal demand is:

$$q^*_{k,i,1}(p_1, \lambda_{i,1}) = \frac{v_1 - p_1}{\lambda_{i,1} + \gamma\sigma_2^2} - \frac{\gamma\sigma_2^2}{\lambda_{i,1} + \gamma\sigma_2^2} I_{i,1} \tag{3.3.1}$$

Therefore the residual demand curve faced by trader k in class i, summing optimal demand (3.3.1) over other traders, has slope $(N-1)(\lambda_{i,1} + \gamma\sigma_2^2) + N(\lambda_{-i,1} + \gamma\sigma_2^2)$. Requiring consistency of conjectured equilibrium slope of the residual demand curve and the actual ones:

$$\lambda_{i,1} = \left((N-1)(\lambda_{i,1} + \gamma\sigma_2^2)^{-1} + N(\lambda_{-i,1} + \gamma\sigma_2^2)^{-1}\right)^{-1} \tag{3.3.2}$$

Definition 2. *A date 1 equilibrium with imperfect competition is a set of demand schedules as in (3.3.1), of $\lambda_{1,1}$ and $\lambda_{2,1}$ that solve (3.3.2) and a price p_1^* such that the market clearing condition (3.2.3) holds.*

Proposition 11 (Vayanos (1999), Malamud and Rostek (2017))**.** *A date 1 equilibrium with imperfect competition exists and is unique. In this equilibrium, $\lambda_{1,1} = \lambda_{2,1} = \frac{\gamma\sigma_2^2}{2N-2}$ so that equilibrium demand schedules are:*

$$q^*_{i,1}(p_1) = \frac{2N-2}{2N-1}\left[\frac{v_1 - p_1}{\gamma\sigma_2^2} - I_{i,1}\right] \tag{3.3.3}$$

The equilibrium quantities traded and post trade inventories are

$$q^*_{i,1} = \frac{2N-2}{2N-1}\frac{I_{i,1} - I_{-i,1}}{2} + \frac{Q}{2N} \tag{3.3.4}$$

$$I_{i,1} + q^*_{i,1} = \frac{Q^*}{2} + \frac{1}{2N-1} I_{i,1} \tag{3.3.5}$$

The equilibrium price is

$$p_1^* = v_1 - \bar{\gamma}\sigma_2^2 Q^* \tag{3.3.6}$$

$$\text{with} \begin{cases} Q^* & = \frac{2N-2}{2N-1}(I_{1,1} + I_{2,1}) + \frac{Q}{N} \\ \bar{\gamma} & = \frac{2N-1}{2N-2}\frac{\gamma}{2} \end{cases}$$

[3]Cf. Kyle (1989), Vayanos (1999), Malamud and Rostek (2017) among many others.

3.3.2 Date 0

For date 0, I assume rational expectations on equilibrium price and quantities: dealer k within class i conjectures symmetric equilibrium trades $q^e_{l,-i,0} = q^e_{-i,0}$ for all dealers in the other class $-i$, and $q^e_{l,i,1} = q^e_{i,1}$ for other dealers ($l \neq k$) in his class i. This leads to conjectured date 1 initial inventory $I^e_{-i,1} = I_{-i,0} + q^e_{-i,0}$, and $I^e_{l,i,1}$ ($l \neq k$). Dealer k optimizes according to this conjecture. In equilibrium, these conjectures coincide with actual equilibrium quantities:

$$q^e_{i,0} = q^*_{i,0}(p^*_0). \tag{3.3.7}$$

The certainty equivalent of wealth can be written (*cf.* lemma 18 in the appendix):

$$\widehat{W}_{k,i,0} = I_{i,0}v_0 + q_{k,0}(v_0 - p_0) - \frac{\gamma(\sigma_1^2 + \sigma_2^2)}{2}(I_{k,1})^2 + \widehat{S}^*_{k,1}(q_{k,0}) \tag{3.3.8}$$

where

$$\begin{aligned}
\widehat{S}^*_{k,1}(q_{k,0}) &= \frac{2N}{2N-2}\frac{1}{1+\alpha x}\frac{\gamma\sigma_2^2}{2}\left(\frac{\bar{\gamma}}{\gamma}\mathbb{E}_0\left[q^*_{i,1}\right]\right)^2 \\
&= \frac{\alpha}{1+\alpha x}\frac{\gamma\sigma_2^2}{2}\left(\frac{\bar{\gamma}}{\gamma}\mathbb{E}_0\left[\frac{Q}{N}\right] + \frac{\bar{I}^e_{-i,1}}{2} + \frac{N-1}{N}\frac{\bar{I}^e_{i,1}}{2} - \left(1 - \frac{1}{2N}\right)I_{k,1}\right)^2
\end{aligned}$$

and $x = \bar{\gamma}^2\sigma_2^2\sigma_q^2$ and $\alpha = \frac{2N(2N-2)}{(2N-2)^2}$ is increasing with N and is strictly between 0 and 1. The first order condition of the maximization of the above criterion, together with a consistency condition of price impacts $\lambda_{i,0}$ analogous to (3.3.2), leads to equilibrium demand schedules that are identical across dealers of class i (*cf.* lemma 19 in appendix):

$$\begin{aligned}
q^*_{i,0}(p_0) = \frac{2N-2}{2N-1}\Bigg[&\frac{v_0 - p_0}{\gamma(\sigma_1^2 + \delta\sigma_2^2)} - I_{i,0} \\
&- \frac{2N-2}{2N-1}\frac{1}{1+\alpha x}\frac{\gamma\sigma_2^2}{\gamma(\sigma_1^2+\delta\sigma_2^2)}\left(\frac{\bar{\gamma}}{\gamma}\mathbb{E}_0\left[\frac{Q}{N}\right] + \frac{\bar{I}^e_{-i,1}}{2} + \frac{N-1}{N}\frac{\bar{I}^e_{i,1}}{2}\right)\Bigg]
\end{aligned} \tag{3.3.9}$$

where $\delta = 1 - \frac{N-1}{N}\frac{1}{1+\alpha x} \in [0,1)$. There are a few differences between this imperfectly competitive demand schedule and the competitive one. Similarly to the date 1 equilibrium, all terms in the demand schedule are reduced by a factor by $(2N-2)/(2N-1) < 1$. In addition, consider the first term representing the hold-to-maturity component of demand: it is divided by a variance $\sigma_1^2 + \delta\sigma_2^2$ where $\delta \in [0,1)$ and is analogous to the competitive case. But unlike the competitive case, when the supply shock Q is known for sure ($\sigma_q^2 = 0$), $\delta = 1/N > 0$. It reflects the fact that dealer i has to keep part of his position built at date 0 until maturity, as risk sharing is limited by imperfect competition at date 1. Otherwise δ increases with σ_q^2 and converges to 1 as σ_q^2 becomes arbitrarily large, as in the competitive case: dealer k's effective horizon converges to $t = 2$.

Plugging (3.3.9) into the market clearing condition (3.2.2), and imposing consistency of conjectures on others' trades (3.3.7), I derive the following proposition, fully proven in appendix.

Proposition 12. *At date 0, the equilibrium quantity traded by class 1 traders is*

$$q_{1,0}^* = \frac{1}{1 + A(\sigma_q^2)} \frac{I_{2,0} - I_{1,0}}{2} + \frac{Q_c}{2N} \tag{3.3.10}$$

where $A(\cdot)$ is the demand reduction rate. It is positive and decreases with σ_q^2.

$A(\cdot)$ is the sum of a static demand reduction rate and of a positive dynamic demand reduction rate. The static demand reduction rate is the same as the one obtained in the static date 1 market. The dynamic demand reduction rate is positive and falls to zero as σ_q^2 goes to infinity.

The date 0 equilibrium price is:

$$p_0^* = v_0 - \gamma(\sigma_1^2 + \sigma_2^2) \frac{I_{1,0} + I_{2,0}}{2} - \gamma s \frac{Q_c}{2N} \tag{3.3.11}$$

where

$$\begin{aligned} s &= \frac{2N-1}{2N-2}(\sigma_1^2 + \sigma_2^2) - \frac{1}{2N} \frac{\sigma_2^2}{1 + \alpha x} \\ x &= \bar{\gamma}^2 \sigma_2^2 \sigma_q^2 \\ \bar{\gamma} &= \frac{\gamma}{2} \frac{2N-1}{2N-2} \end{aligned}$$

Equilibrium utilities. Plugging the equilibrium price and quantity traded in the centralized market in the certainty equivalent of wealth, one finds the following expression that is conditional on inventory $I_{i,0} = 0$ for dealers who have not traded in the centralized market, or $q_{i,d}$ for those who have. The lemma is fully proven in the appendix.

Lemma 12. *The equilibrium date 0 certainty equivalent of wealth of a dealer of class i arriving with inventory $I_{i,0}$ in the date 0 centralized market is:*

$$\begin{aligned} \widehat{W}_{i,0}^n(Q_c) =& I_{i,0} v_0 - \frac{\gamma(\sigma_1^2 + \sigma_2^2)}{2} (I_{i,0})^2 \\ &+ \gamma \theta_{cc}^n \left(\frac{Q_c}{2N} \right)^2 + \gamma \theta_{cd}^n \frac{Q_c}{2N} \frac{I_{-i,0} - I_{i,0}}{2} + \gamma \frac{\theta_{dd}^n}{2} \left(\frac{I_{-i,0} - I_{i,0}}{2} \right)^2 \end{aligned} \tag{3.3.12}$$

where, denoting $A \equiv A(\sigma_q^2)$,

$$\begin{aligned} \theta_{cc}^n &= s - \frac{\sigma_1^2 + \sigma_2^2}{2} = \left(1 + \frac{1}{N-1}\right) \frac{\sigma_1^2 + \sigma_2^2}{2} + \frac{1}{2N} \frac{\sigma_2^2}{1 + \alpha x} \\ \theta_{cd}^n &= \frac{(\sigma_1^2 + \sigma_2^2)A + s}{1 + A} \\ \theta_{dd}^n &= \sigma_1^2 + \sigma_2^2 - \left(\frac{A}{1+A} \right)^2 \left(\sigma_1^2 + \left(1 - \frac{\alpha}{1 + \alpha x}\right) \sigma_2^2 \right) \end{aligned}$$

The certainty equivalent of wealth is the sum of a mean variance criterion (first line of 3.3.12) and of the value of the surplus of date 0 and date 1 transactions (second line). The latter is itself the sum of a term corresponding to the surplus made from the customer transaction on the centralized market (in Q_c^2), a term corresponding to interdealer transactions (in $(I_{-i,1} - I_{i,1})^2$), which themselves come from transactions in the parallel market; the cross term (in $Q_c \times (I_{-i,0} - I_{i,0})$ comes from the fact that customer transactions in the interdealer market changes the interdealer terms of trade: if customers sell in the centralized market ($Q_c > 0$), then selling dealers (for which $I_{-i,0} - I_{i,0} < 0$) sell at a lower price, which decreases their utility.

3.3.3 Equilibrium in the parallel market

Customer-to-dealer market equilibrium price

If dealer k of class 1 has bought $q_{k,d}$ in the parallel market to the customers at price p_d, while dealers of class 2 have no inventory, the inventory difference is $\Delta I = q_d$ and the certainty equivalent of wealth for dealer k is:

$$\widehat{W}_{k,d} = q_{1,d}(v_0 - p_d) - \frac{\gamma(\sigma_1^2 + \sigma_2^2)}{2}(q_d)^2 + \gamma\theta_{cc}\left(\frac{Q_c}{2N}\right)^2 + \gamma\theta_{cd}\frac{Q_c}{2N}\frac{q_d}{2} + \gamma\frac{\theta_{dd}}{2}\left(\frac{q_d}{2}\right)^2$$

The first order condition for the maximization of W_k is

$$v_0 - p_d = \lambda_k q_{k,d} + \gamma\left(\sigma_1^2 + \sigma_2^2 - \frac{\theta_{dd}}{4}\right)q_{k,d} - \gamma\frac{\theta_{cd}}{2}\frac{Q_c}{2N}$$

Analogously to other equilibria (date 0, date 1), the equilibrium price impact is:

$$\lambda_k = \frac{\gamma(\sigma_1^2 + \sigma_2^2 - \theta_{dd}/4)}{N-2}$$

so that the equilibrium demand schedule of trader k is identical for all traders of class 1:

$$q_{1,d}^*(p_d) = \frac{2N-2}{2N-1}\left(\frac{v_0 - p_d}{\gamma(\sigma_1^2 + \sigma_2^2 - \theta_{dd}/4)} - \frac{\theta_{cd}}{2(\sigma_1^2 + \sigma_2^2 - \theta_{dd}/4)}\frac{Q_c}{2N}\right)$$

The market clearing condition is $Nq_{1,d} = Q_d$, *i.e.*

$$\frac{v_0 - p_d^*}{\gamma(\sigma_1^2 + \sigma_2^2 - \theta_{dd}/4)} = \frac{\theta_{cd}}{2(\sigma_1^2 + \sigma_2^2 - \theta_{dd}/4)}\frac{Q_c}{2N} + \frac{N-1}{N-2}\frac{Q_d}{N}$$

which can be rearranged to lead to the following proposition.

Proposition 13. *In the parallel market, the equilibrium price is*

$$p_d^* = v_0 - \gamma\frac{\theta_{cd}}{2}\frac{Q_c}{2N} - \gamma\left(\sigma_1^2 + \sigma_2^2 - \frac{\theta_{dd}}{4}\right)\frac{N-1}{N-2}\frac{Q_d}{N} \qquad (3.3.13)$$

while each dealer of class 1 gets the same quantity

$$q_{1,d} = \frac{Q_d}{N} \qquad (3.3.14)$$

The equilibrium price is the expected value of the asset minus a discount proportional to Q_c, and another proportional to Q_d. The first, proportional to Q_c, reflect the anticipation by dealers that they will trade at a low price p_0^* in the centralized market if customers sell there ($Q_c > 0$), which reduces the terms of trade in the interdealer market. The second discount is a risk premium for holding the asset until maturity $\gamma(\sigma_1^2 + \sigma_2^2)$ which is decreased by the surplus that dealers will make in later transactions, thus the θ_{dd}.

The centralized/parallel market spread

Buying a quantity Q_d/N in the parallel market, and re-selling in the centralized market, dealers of class 1 get the following spread, obtained by taking the difference between 3.3.11 with $I_{1,0} + I_{2,0} = \frac{Q_d}{N}$ and 3.3.13, one gets

$$p_0^* - p_d^* = \gamma \left(\frac{N-1}{N-2} \left(\sigma_1^2 + \sigma_2^2 - \frac{\theta_{dd}}{4} \right) - \frac{\sigma_1^2 + \sigma_2^2}{2} \right) \frac{Q_d}{N} + \gamma \left(\frac{\theta_{cd}}{2} - s \right) \frac{Q_c}{2N}$$

In the appendix I show that this boils down to

Lemma 13.

$$p_0^* - p_d^* = \gamma \left(s_{d1}\sigma_1^2 + s_{d2}\sigma_2^2 \right) \frac{Q_d}{2N} - \gamma \left(s_{c1}\sigma_1^2 + s_{c2}\sigma_2^2 \right) \frac{Q_c}{2N} \tag{3.3.15}$$

where $s_{d1}, s_{d2}, s_{c1}, s_{c2} > 0$.

Therefore, assuming first that $Q_c = 0$, dealers of class 1 make a positive spread on the quantity Q_d they buy. However, the spread decreases if $Q_c > 0$: if customers sell both in the centralized market and in the parallel market and in the same direction, dealers do not fully pass the decrease in centralized market terms of trade through the parallel market. This points towards a preference of dealers trading in the parallel market for having no trade in the centralized market.

Equilibrium utility for dealers trading in the parallel market.

Lemma 14. *Dealers of class 1 have equilibrium utility*

$$\widehat{W}_1 = \frac{\gamma}{2} \frac{N}{N-2} \left(\sigma_1^2 + \sigma_2^2 - \frac{\theta_{dd}}{4} \right) \left(\frac{Q_d}{N} \right)^2 + \gamma \theta_{cc} \left(\frac{Q_c}{2N} \right)^2 + \gamma \theta_{cd} \frac{Q_c}{2N} \frac{Q_d}{N}$$

Proof. The certainty of wealth, plugging equilibrium price 3.3.13 and quantity 3.3.14, is

$$
\begin{aligned}
\widehat{W}_1 &= \frac{Q_d}{N}\left(\gamma\frac{\theta_{cd}}{2}\frac{Q_c}{2N} + \frac{N-1}{N-2}\left(\sigma_1^2+\sigma_2^2-\frac{\theta_{dd}}{4}\right)\frac{Q_d}{N}\right) - \frac{\gamma(\sigma_1^2+\sigma_2^2)}{2}\left(\frac{Q_d}{N}\right)^2 \\
&\quad + \gamma\theta_{cc}\left(\frac{Q_c}{2N}\right)^2 + \gamma\theta_{cd}\frac{Q_c}{2N}\frac{Q_d}{2N} + \gamma\frac{\theta_{dd}}{2}\left(\frac{Q_d}{2N}\right)^2 \\
&= \gamma\left[\frac{N-1}{N-2}\left(\sigma_1^2+\sigma_2^2-\frac{\theta_{dd}}{4}\right) - \frac{\sigma_1^2+\sigma_2^2}{2} + \frac{\theta_{dd}}{8}\right]\left(\frac{Q_d}{N}\right)^2 + \gamma\theta_{cc}\left(\frac{Q_c}{2N}\right)^2 \\
&\quad + \gamma\left[\theta_{cd}+\theta_{cd}\right]\frac{Q_c}{2N}\frac{Q_d}{2N} \\
&= \frac{\gamma}{2}\frac{N}{N-2}\left(\sigma_1^2+\sigma_2^2-\frac{\theta_{dd}}{4}\right)\left(\frac{Q_d}{N}\right)^2 + \gamma\theta_{cc}\left(\frac{Q_c}{2N}\right)^2 + \gamma\theta_{cd}\frac{Q_c}{2N}\frac{Q_d}{N}
\end{aligned}
$$

■

3.4 Dealers' preferred trading venue

In this section I examine whether dealers prefer customers to trade in the centralized or in the parallel market, which is the main question of the paper. Denote w the fractin of their total order Q_0 that customers trade in the parallel market, so that the fraction traded in the centralized market is $1-w$. To avoid customers buying in the centralized market and selling in the parallel market or conversely, I constrain w to be in $[0,1]$.

3.4.1 Dealers prefer all or nothing in the parallel market

The certainty equivalent of wealth is, from lemma 14:

$$
\widehat{W}_1(w) = \gamma\left[\frac{1}{2}\frac{N}{N-2}\left(\sigma_1^2+\sigma_2^2-\frac{\theta_{dd}}{4}\right)w^2 + \frac{\theta_{cc}}{4}(1-w)^2 + \frac{\theta_{cd}}{2}w(1-w)\right]\left(\frac{Q_0}{N}\right)^2 \tag{3.4.1}
$$

Given the constraint $w \in [0,1]$, I show in the appendix the following proposition.

Proposition 14. *Dealers trading in the parallel market prefer that all customer trading happens in the parallel market ($w=1$), thus force customers to trade with them in the parallel market.*

Dealers who do not trade in the parallel market prefer that all customer trading happens in the centralized market ($w=0$).

A sketch of proof is as follows. For traders trading in the parallel market, the certainty equivalent of wealth $\widehat{W}_1$ is a convex quadratic function of w, so that the maximum is attained either for $w=0$ or for $w=1$. Then I show that $\widehat{W}_1(1) > \widehat{W}_1(0)$. For traders not trading in the parallel market, $\widehat{W}_2$ is a quadratic concave function of w, and its unconstrained maximum (i.e. for w any real number) is negative: thus the constrained maximum ($w \in [0,1]$) is at $w = 0$.

Thus ex post, dealers who are not assigned the trade would vote for customers trading only in the centralized market.

3.4.2 Dealers' ex ante preference for opening the parallel market

However, before knowing if they belong to the class that is assigned the customer order, dealers may vote on opening the parallel market, or not. Their ex ante utility if they open the parallel market (thus imposing $w = 1$) is

$$\frac{c}{2}\left(-e^{-\gamma \widehat{W}_1(1)} - e^{-\gamma \widehat{W}_2(1)}\right)$$

for some positive constant c, while their utility if they do not open the parallel market ($w = 0$) is

$$\frac{c}{2}\left(-e^{-\gamma \widehat{W}_1(0)} - e^{-\gamma \widehat{W}_2(0)} 2\right) = -ce^{-\gamma \widehat{W}_2(0)}$$

One has the following theorem, which is the main result of the paper.

Theorem 3. *For σ_1^2/σ_2^2 not too large, dealers choose to open the parallel market. When σ_1^2/σ_2^2 becomes large, dealers prefer trading in the centralized market.*

Customers get a worse price when they trade only in the parallel market than if they trade in the centralized market. Thys they are worse off when dealers open the parallel market.

The fact that when $\sigma_1^2 >> \sigma_2^2$, dealers choose to trade in the centralized market suggests a link between over-the-counter trading dynamic trading. In fact, in this case, the market becomes close to static as date 1 trading surplus becomes small with respect to the cost of holding the asset from date 0 to date 1.

3.5 Simultaneous trading in the parallel and date 0 centralized market

In this section I change the timing assumption regarding trading in the parallel and n in the date 0 centralized market: in the baseline setting, trading in the parallel market occured before trading in the centralized market; now I assume that trading in the parallel and date 0 centralized market occur simultaneously.

The date 0 equilibrium changes only in the parallel market, where customers trade at a better price. Equilibrium in the centralized market is not changed.

Ultimately, dealers who are allowed to trade in the parallel market are still better off when all customer trading happens in the parallel market. Ex ante, dealers still prefer when

3.5.1 Equilibrium

Traders not trading in the parallel market. Traders of class 2 do not trade with customers. For a trader k in class 2, the certainty equivalent of wealth at date 0 is

$$\widehat{W}_{k,2,0} = I_{k,2,0}v_0 + q_{k,2,0}(v_0 - p_c) + \frac{1}{2}\frac{2N}{2N-2}\frac{\gamma\sigma_2^2}{1+\alpha x}\left(\mathbb{E}_0\left[q^*_{k,2,1}\right]\right)^2$$

where

$$\begin{aligned} q^*_{k,2,1} &= \frac{2N-2}{2N-1}\left(Q^* - I_{k,2,1}\right) \\ &= \frac{2N-2}{2N-1}\left(\frac{Q_1}{2N} + \frac{\bar{I}^e_{1,1}}{2} + \frac{\bar{I}^e_{2,1}}{2} - I_{k,2,1}\right) \end{aligned}$$

$\bar{I}^e_{i,1}$ is the rational expectation of the average inventory of dealers of class i after date 0 trades. These expectations are common to everyone. Dealers of class 2 are sensitive to the impact they have on this average inventory, which goes through the following identity:

$$\bar{I}^e_{2,1} = \frac{1}{N}\sum_{l=1,l\neq k}^{N} I^e_{k,2,1} + \frac{I_{k,2,1}}{N}$$

As all traders within the same class are expected to have symmetric equilibrium trades, the identity can be written

$$\bar{I}^e_{2,1} = \frac{N-1}{N}\bar{I}^e_{2,1} + \frac{I_{k,2,1}}{N}$$

Therefore traders of class 2 maximize

$$\begin{aligned} \widehat{W}_{k,2,0} &= I_{k,2,0}v_0 + q_{k,2,0}(v_0 - p_c) \\ &\quad + \frac{1}{2}\frac{2N(2N-2)}{(2N-1)^2}\frac{\gamma\sigma_2^2}{1+\alpha x}\left(\frac{\bar{I}^e_{1,1}}{2} + \frac{N-1}{N}\frac{\bar{I}^e_{2,1}}{2} - \frac{2N-1}{2N}I_{k,2,1}\right)^2 \end{aligned}$$

Their first order condition is

$$v_0 - p_c = \lambda_2\ q_{k,2,0} + \gamma(\sigma_1^2 + \delta\sigma_2^2)I_{k,2,1} + \frac{2N-2}{2N-1}\left(\frac{\bar{I}^e_{1,1}}{2} + \frac{N-1}{N}\frac{\bar{I}^e_{2,1}}{2}\right)$$

where $\delta = 1 - \frac{N-1}{N}\frac{1}{1+\alpha x}$ and $x = \bar{\gamma}^2\sigma_2^2\sigma_q^2$. It is easy to check that the problem is strictly concave in $q_{k,2,0}$.

Traders trading in the OTC market. Trader k in class 1 trades a quantity $q_{k,d}$ with customers in the fragmented market, and $q_{k,1,0}$ in the centralized market. The certainty equivalent of wealth is analogous to that of class 2 traders:

$$\widehat{W}_{k,1,0} = I_{k,1,0}v_0 + q_{k,d}(v_0 - p_d) + q_{k,1,0}(v_0 - p_c)$$
$$+ \frac{1}{2}\frac{2N(2N-2)}{(2N-1)^2}\frac{\gamma\sigma_2^2}{1+\alpha x}\left(\frac{\bar{I}^e_{2,1}}{2} + \frac{N-1}{N}\frac{\bar{I}^e_{1,1}}{2} - \frac{2N-1}{2N}I_{k,1,1}\right)^2$$

This trader optimizes simultaneously on both quantities $q_{k,d}$ and $q_{k,1,0}$ so that the first order conditions are

$$\begin{cases} v_0 - p_c &= \lambda_{cc}q_{k,1,0} + \lambda_{cd}q_{k,d} + \gamma(\sigma_1^2 + \delta\sigma_2^2)(I + q_{k,d} + q_{k,1,0}) + \frac{2N-2}{2N-1}\frac{\gamma\sigma_2^2}{1+\alpha x}\left(\frac{\bar{I}^e_{2,1}}{2} + \frac{N-1}{N}\frac{\bar{I}^e_{1,1}}{2}\right) \\ v_0 - p_d &= \lambda_{dc}q_{k,1,0} + \lambda_{dd}q_{k,d} + \gamma(\sigma_1^2 + \delta\sigma_2^2)(I + q_{k,d} + q_{k,1,0}) + \frac{2N-2}{2N-1}\frac{\gamma\sigma_2^2}{1+\alpha x}\left(\frac{\bar{I}^e_{2,1}}{2} + \frac{N-1}{N}\frac{\bar{I}^e_{1,1}}{2}\right) \end{cases}$$

which can be re-written in matrix form

$$\begin{pmatrix} v_0 - p_c \\ v_0 - p_d \end{pmatrix} = \left(\gamma(\sigma_1^2 + \delta\sigma_2^2)\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + \Lambda\right)\begin{pmatrix} q_{k,1,0} \\ q_{k,d} \end{pmatrix}$$
$$+ \left(\gamma(\sigma_1^2 + \delta\sigma_2^2)I + \frac{2N-2}{2N-1}\frac{\gamma\sigma_2^2}{1+\alpha x}\left(\frac{\bar{I}^e_{2,1}}{2} + \frac{N-1}{N}\frac{\bar{I}^e_{1,1}}{2}\right)\right)\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

where

$$\Lambda = \begin{pmatrix} \lambda_{cc} & \lambda_{cd} \\ \lambda_{dc} & \lambda_{dd} \end{pmatrix}$$

Equilibrium price impact matrix. The equilibrium λ_2 and Λ obey the following equations, denoting $\sigma^2 = \sigma_1^2 + \delta\sigma_2^2$ to ease notation:

$$\begin{cases} \Lambda &= \left(\frac{N}{\gamma\sigma^2+\lambda_2}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + (N-1)\left(\gamma\sigma^2\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + \Lambda\right)^{-1}\right)^{-1} \\ \lambda_2 &= \left[\left(\frac{N-1}{\gamma\sigma^2+\lambda_2}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + N\left(\gamma\sigma^2\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + \Lambda\right)^{-1}\right)^{-1}\right]_{1,1} \end{cases} \tag{3.5.1}$$

Lemma 15. *There is a unique solution to the system (3.5.1) such that* $\gamma\sigma^2\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + \Lambda$ *is invertible. It is:*

$$\Lambda = \frac{\gamma(\sigma_1^2 + \delta\sigma_2^2)}{2N-2}\begin{pmatrix} 1 & 1 \\ 1 & 1+y \end{pmatrix} \tag{3.5.2}$$

$$\lambda_2 = \frac{\gamma(\sigma_1^2 + \delta\sigma_2^2)}{2N-2} \tag{3.5.3}$$

where $y = \frac{N}{(N-1)(2N-1)}$.

The lemma is proven in appendix. The coefficient λ_{cc} in Λ is the same as if there was no parallel market. The off-diagonal terms λ_{cd} and λ_{dc} reflect an arbitrage of the same asset in different markets: buying more in the parallel market, for class 1 traders implies an increase in decentralized price p_d, which induces other class 1 traders to buy more of the same asset in the centralized market, and symmetrically for the other cross-market impact.

Plugging the solution into the first order condition yields

$$\begin{pmatrix} v_0 - p_c \\ v_0 - p_d \end{pmatrix} = \gamma(\sigma_1^2 + \delta\sigma_2^2)\frac{2N-1}{2N-2}\begin{pmatrix} 1 & 1 \\ 1 & 1 + \frac{y}{2N-1} \end{pmatrix}\begin{pmatrix} q_{k,1,0} \\ q_{k,d} \end{pmatrix} + \frac{2N-2}{2N-1}\frac{\gamma\sigma_2^2}{1+\alpha x}\left(\frac{\bar{I}^e_{2,1}}{2} + \frac{N-1}{N}\frac{\bar{I}^e_{1,1}}{2}\right)\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

which can be inverted to

$$\begin{aligned}
\begin{pmatrix} q^*_{k,1,0}(p_c, p_d) \\ q^*_{k,d}(p_c, p_d) \end{pmatrix} &= \frac{2N-2}{2N-1}\frac{1}{\gamma(\sigma_1^2 + \delta\sigma_2^2)}\frac{2N-1}{y}\begin{pmatrix} 1 + \frac{y}{2N-1} & -1 \\ -1 & 1 \end{pmatrix} \\
&\quad \times \left[\begin{pmatrix} v_0 - p_c \\ v_0 - p_d \end{pmatrix} - \frac{2N-2}{2N-1}\frac{\gamma\sigma_2^2}{1+\alpha x}\left(\frac{\bar{I}^e_{2,1}}{2} + \frac{N-1}{N}\frac{\bar{I}^e_{1,1}}{2}\right)\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right] \\
&= \frac{2N-2}{2N-1}\frac{1}{\gamma(\sigma_1^2 + \delta\sigma_2^2)}\begin{pmatrix} v_0 - p_c + \frac{2N-1}{y}(p_d - p_c) \\ \frac{2N-1}{y}(p_c - p_d) \end{pmatrix} \\
&\quad - \frac{2N-2}{2N-1}\left(\frac{2N-2}{2N-1}\frac{\sigma_2^2}{(\sigma_1^2 + \delta\sigma_2^2)(1+\alpha x)}\left(\frac{\bar{I}^e_{2,1}}{2} + \frac{N-1}{N}\frac{\bar{I}^e_{1,1}}{2}\right)\right)\begin{pmatrix} 1 \\ 0 \end{pmatrix}
\end{aligned}$$

The market clearing conditions apply in the parallel market, which write, given that all class 1 traders post identical demand schedules:

$$\frac{2N-2}{y}\frac{p_c^* - p_d^*}{\gamma(\sigma_1^2 + \delta\sigma_2^2)} = \frac{Q_d}{N}$$

and in the centralized market, where class 1 and class 2 traders meet:

$$\begin{aligned}
&2\frac{v_0 - p_c^*}{\gamma(\sigma_1^2 + \delta\sigma_2^2)} + \frac{2N-2}{y}\frac{p_d^* - p_c^*}{\gamma(\sigma_1^2 + \delta\sigma_2^2)} \\
&\quad - \frac{2N-2}{2N-1}\frac{\sigma_2^2}{(\sigma_1^2 + \delta\sigma_2^2)(1+\alpha x)}\left(\bar{I}^e_{2,1} + \bar{I}^e_{1,1}\right)\frac{2N-1}{2N} = \frac{2N-1}{2N-2}\frac{Q_c}{N}
\end{aligned}$$

Plugging the OTC market clearing condition, observing that $\bar{I}^e_{1,1} + \bar{I}^e_{2,1} = \frac{Q_d}{N} + \frac{Q_c}{N}$

$$\frac{v_0 - p_c^*}{\gamma(\sigma_1^2 + \delta\sigma_2^2)} - \frac{Q_d}{2N} - \frac{2N-2}{2N}\frac{\sigma_2^2}{(\sigma_1^2 + \delta\sigma_2^2)(1+\alpha x)}\left(\frac{Q_d}{2N} + \frac{Q_c}{2N}\right) = \frac{2N-1}{2N-2}\frac{Q_c}{2N}$$

This leads to the following proposition.

Proposition 15. *The equilibrium price in the centralized and parallel market are*

$$p_c^* = v_0 - \gamma(\sigma_1^2 + \sigma_2^2)\frac{Q_d}{2N} - \gamma s \frac{Q_c}{2N} \tag{3.5.4}$$

$$p_c^* - p_d^* = \gamma(\sigma_1^2 + \delta\sigma_2^2)\frac{N}{(N-1)^2(2N-1)}\frac{Q_d}{2N} \tag{3.5.5}$$

where

$$s = \frac{2N-1}{2N-2}(\sigma_1^2 + \sigma_2^2) - \frac{1}{2N}\frac{\sigma_2^2}{1+\alpha x}$$

The equilibrium quantity traded in the centralized market is

$$q_{1,0}^* = \frac{1}{1+A(\sigma_q^2)}\frac{Q_d}{2N} + \frac{Q_c}{2N} \tag{3.5.6}$$

where $A(\sigma_q^2)$ is positive and decreases as σ_q^2 increases.

With respect to the equilibrium with sequential trading in the parallel market, the centralized market price and quantity traded are identical. What changes is the parallel market price p_d^*. In particular, the spread $p_c^* - p_d^*$ does not depend on the quantity Q_c traded in the centralized market.

The spread $p_c^* - p_d^*$ is a measure of price dispersion in an OTC market. It increases with risk aversion: in periods of market stress, risk aversions increase (possibly because of tighter financing constraints) and price dispersion increases. Price dispersion also increases when uncertainty σ_q^2 on order flow increases (making δ increase).

Dealers' equilibrium utilities. As only the price in the parallel market changes with respect to the sequential trading case, dealers not trading in the parallel market have identical utilities as in the sequential trading case. The other dealers' utility is given in the following lemma.

Lemma 16. *Dealers trading in the parallel market have equilibrium utility*

$$\widehat{W}_1(Q_c, Q_d) = \left\{\left[\frac{\sigma_1^2+\sigma_2^2}{2}\left(1 + \frac{2N}{(N-1)^2(2N-1)}\right) + \frac{\sigma_2^2}{1+\alpha x}\frac{1}{(N-1)(2N-1)} + \frac{\theta_{dd}}{8}\right]w^2 \right.$$
$$\left. + \frac{\theta_{cc}}{4}(1-w)^2 + \frac{s+\theta_{cd}}{4}w(1-w)\right\}\gamma\left(\frac{Q_0}{N}\right)^2 \tag{3.5.7}$$

Proof. The certainty of wealth, plugging equilibrium price 3.5.5 and quantity 3.5.6, is

$$
\begin{aligned}
\widehat{W}_1 = & \frac{Q_d}{N}\left(\gamma\left((\sigma_1^2+\sigma_2^2)\left(1+\frac{N}{(N-1)^2(2N-1)}\right)+\frac{\sigma_2^2}{1+\alpha x}\frac{1}{(N-1)(2N-1)}\right)\frac{Q_d}{2N}+\gamma s\frac{Q_c}{2N}\right) \\
& -\frac{\gamma(\sigma_1^2+\sigma_2^2)}{2}\left(\frac{Q_d}{N}\right)^2+\gamma\theta_{cc}\left(\frac{Q_c}{2N}\right)^2+\gamma\theta_{cd}\frac{Q_c}{2N}\frac{Q_d}{2N}+\gamma\frac{\theta_{dd}}{2}\left(\frac{Q_d}{2N}\right)^2 \\
= & \gamma\left[(\sigma_1^2+\sigma_2^2)\left(1+\frac{N}{(N-1)^2(2N-1)}\right)+\frac{\sigma_2^2}{1+\alpha x}\frac{1}{(N-1)(2N-1)}-\frac{\sigma_1^2+\sigma_2^2}{2}+\frac{\theta_{dd}}{8}\right]\left(\frac{Q_d}{N}\right)^2 \\
& +\gamma\theta_{cc}\left(\frac{Q_c}{2N}\right)^2+\gamma\left[s+\theta_{cd}\right]\frac{Q_c}{2N}\frac{Q_d}{2N} \\
= & \gamma\left[\frac{\sigma_1^2+\sigma_2^2}{2}\left(1+\frac{2N}{(N-1)^2(2N-1)}\right)+\frac{\sigma_2^2}{1+\alpha x}\frac{1}{(N-1)(2N-1)}+\frac{\theta_{dd}}{8}\right]\left(\frac{Q_d}{N}\right)^2 \\
& +\gamma\theta_{cc}\left(\frac{Q_c}{2N}\right)^2+\gamma\frac{s+\theta_{cd}}{2}\frac{Q_c}{2N}\frac{Q_d}{N}
\end{aligned}
$$

Then plug $Q_d = wQ_0$ and $Q_c = (1-w)Q_0$. ■

3.5.2 Dealers' preferred trading venue

In the previous subsection, I showed that only the equilibrium price in the parallel market changes with respect to the equilibrium where trading in the parallel market occurs first. Therefore the preference of dealers who do not trade in the parallel market are unchanged: in particular, they still ex post prefer to have all customer trading in the centralized market.

I reconsider the preference of dealers who trade in the centralized market.

Proposition 16. *Dealers who trade in the parallel market prefer that all customer trading happens in the parallel market.*

The proposition is proven in the appendix. Now I revisit dealers' choice ex ante.

Theorem 4. *Before they learn whether they will trade in the parallel market, dealers' expected payoff is higher than the expected payoff from trading in the centralized market only.*

For $\gamma\sigma_2^2$ not too high, dealers ex ante choose to open the parallel market. Customers are worse off than when they trade only in the centralized market.

The theorem is proven in the appendix. Therefore the equilibrium and welfare conclusions remain qualitatively the same as in the case where dealers trade in the parallel market *before* they trade in the centralized market.

3.6 Conclusion

In this paper I provide a model of endogenous market fragmentation: dealers choose to open a market that is parallel to the all-to-all, centralized market. In doing this they extract rents to customers, who are worse off, while dealers who do not trade in the parallel market are worse off. However, if dealers do not know ex ante if they will be in the pool of dealers trading in the parallel market - capturing the idea that they do not know if their own pool of customers will have trading needs - dealers choose to open the parallel market as long as they expect trading opportunities with a cost of carrying the asset until the opportunity materializes is not too high. This paper thus contributes to the literature on endogenous networks.

3.7 Appendix: Proofs

3.7.1 Resolution of date 0 equilibrium without derivatives

Demand schedules

From proposition 11, the post-trade certainty equivalent of wealth at date 1 is given by the following lemma, proven in the appendix.

Lemma 17. *The interim expected utility for a trader k in class i is* $-\exp\left\{-\gamma_i \widehat{W}_{k,i,1}\right\}$, *where $\widehat{W}_{i,1}$ is the interim certainty equivalent of wealth given by:*

$$\widehat{W}_{k,i,1} = I_{k,0}v_1 + q_{k,0}(v_1 - p_0) - \frac{\gamma\sigma_2^2}{2}(I_{k,1})^2 + \alpha\frac{\gamma\sigma_2^2}{2}\left(\frac{\bar{\gamma}}{\gamma}Q^* - I_{k,1}\right)^2 \tag{3.7.1}$$

$$= I_{k,0}v_1 + q_{k,0}(v_1 - p_0) - \frac{\gamma\sigma_2^2}{2}(I_{k,1})^2 + \frac{2N}{2N-2}\frac{\gamma\sigma_2^2}{2}\left(q_{k,1}^*\right)^2 \tag{3.7.2}$$

where $\alpha = \frac{2N(2N-2)}{(2N-1)^2} = 1 - \frac{1}{(2N-1)^2}$.

Proof. Plugging equilibrium price (3.3.6) and quantities (3.3.4) into the date 1 certainty equivalent of wealth (2.2.4), one gets

$$\widehat{W}_{k,1} = I_{k,0}v_1 + q_{k,0}(v_1 - p_0) - \frac{\gamma\sigma_2^2}{2}(I_{k,1})^2 + q_{k,1}^*(v_1 - p_1^*) - \frac{\gamma\sigma_2^2}{2}\left(\left(I_{k,1} + q_{k,1}^*\right)^2 - (I_{k,1})^2\right)$$

Recognizing $\frac{v_1 - p_1^*}{\gamma\sigma_2^2} = \frac{2N-1}{2N-2}q_{k,1}^* + I_{k,1}$ and rearranging one get

$$\begin{aligned}\widehat{W}_{k,i,1} &= I_{k,0}v_1 + q_{k,0}(v_1 - p_0) - \frac{\gamma\sigma_2^2}{2}(I_{k,1})^2 \\ &\quad + q_{k,1}^*\gamma\sigma_2^2\left(\frac{2N-1}{2N-2}q_{k,1}^* + I_{k,1}\right) - \frac{\gamma\sigma_2^2}{2}\left(2I_{k,1} + q_{k,1}^*\right)q_{k,1}^* \\ &= I_{k,0}v_1 + q_{k,0}(v_1 - p_0) - \frac{\gamma\sigma_2^2}{2}(I_{k,1})^2 + \left(\frac{2N-1}{2N-2} - \frac{1}{2}\right)\gamma\sigma_2^2\left(q_{k,1}^*\right)^2\end{aligned}$$

which leads to the desired formulas. ■

It is then possible to compute the certainty equivalent of wealth at date 0.

Lemma 18. *The date 0 certainty equivalent of wealth for trader k in class i is:*

$$\widehat{W}_{k,i,0} = I_{i,0}v_0 + q_{k,0}(v_0 - p_0) - \frac{\gamma(\sigma_1^2 + \sigma_2^2)}{2}(I_{k,1})^2 + \frac{1}{2}\frac{2N}{2N-2}\frac{\gamma\sigma_2^2}{1+\alpha x}\left(\mathbb{E}_0\left[q_{k,1}^*\right]\right)^2 \tag{3.7.3}$$

$$\begin{aligned}&= I_{k,0}v_0 + q_{k,0}(v_0 - p_0) - \frac{\gamma(\sigma_1^2 + \sigma_2^2)}{2}(I_{k,1})^2 \\ &\quad + \frac{\alpha}{2}\frac{\gamma\sigma_2^2}{1+\alpha x}\left(\frac{\bar{\gamma}}{\gamma}\mathbb{E}_0\left[\frac{Q}{N}\right] + \frac{\bar{I}^e_{-i,1}}{2} + \frac{N-1}{N}\frac{\bar{I}^e_{i,1}}{2} - \frac{2N-1}{2N}I_{k,1}\right)^2\end{aligned} \tag{3.7.4}$$

where $x = \bar{\gamma}^2\sigma_2^2\sigma_q^2$ *and* $\bar{\gamma} = \frac{1}{2}\frac{2N-1}{2N-2}$. $\bar{I}^e_{-i,1}$ *and* $\bar{I}^e_{i,1}$ *are the rational expectations of average dealer inventories after date 1 trade.*

Proof. Start from interim expected utility (3.7.2). Take the certainty equivalent with respect to ϵ_1 first, which gives

$$\widehat{W}_{k,0}|Q = I_{i,0}v_1 + q_{k,0}(v_1 - p_0) - \frac{\gamma(\sigma_1^2+\sigma_2^2)}{2}(I_{k,1})^2 + \alpha\frac{\gamma\sigma_2^3}{2}\left(\frac{\bar{\gamma}}{\gamma}Q^* - I_{k,1}\right)^2$$

Then take the certainty equivalent with respect to Q following lemma 11, which gives the desired formula.

For the rational expectations: Q^* is an outcome of date 0 trade, as it depends on dealers' average inventories in each class. Dealer k's trade has an impact on date his class' average inventory

$$\begin{aligned}\bar{I}^e_{i,1} &\equiv \frac{1}{N}\sum_{l=1,l\neq k}^{N} I^e_{l,i,1} + \frac{I_{k,1}}{N}\\ &= \frac{N-1}{N}\bar{I}^e_{i,1} + \frac{I_{k,1}}{N}\end{aligned}$$

where the second line follows from rational expectation of a symmetric equilibrium.

■

Lemma 19. *In equilibrium, all dealers within class i submit the same optimal demand schedules as follows:*

$$\begin{aligned}q^*_{i,0}(p_0) = \frac{2N-2}{2N-1}&\left[\frac{v_0-p_0}{\gamma(\sigma_1^2+\delta\sigma_2^2)} - I_{k,0}\right.\\ &\left.-\frac{2N-2}{2N-1}\frac{1}{1+\alpha x}\frac{\sigma_2^2}{\sigma_1^2+\delta\sigma_2^2}\left(\frac{\bar{\gamma}}{\gamma}\mathbb{E}_0\left[\frac{Q}{N}\right] + \frac{\bar{I}^e_{-i,1}}{2} + \frac{N-1}{N}\frac{\bar{I}^e_{i,1}}{2}\right)\right]\end{aligned} \tag{3.7.5}$$

It depends on trader k's expectation on other traders' equilibrium trades.

Proof. Differentiate the certainty equivalent of wealth (3.7.4) with respect to $q_{i,0}$, taking into account its price impact that is conjectured to be constant (and denoted $\lambda_{i,0}$). Equating to zero to get the first-order condition:

$$\begin{aligned}v_0 - p_0 &= q_{k,0}(\lambda_{k,0} + \gamma(\sigma_1^2+\sigma_2^2))q_{k,0} + \gamma(\sigma_1^2+\sigma_2^2)I_{k,0}\\ &\quad + \frac{2N-2}{2N-1}\frac{\gamma\sigma_2^2}{1+\alpha x}\left(\frac{\bar{\gamma}}{\gamma}\mathbb{E}_0\left[\frac{Q}{N}\right] + \frac{\bar{I}^e_{-i,1}}{2} + \frac{N-1}{N}\frac{\bar{I}^e_{i,1}}{2} - \frac{2N-1}{2N}I_{k,1}\right)\\ &= (\lambda_{k,0} + \gamma(\sigma_1^2+\delta\sigma_2^2))q_{i,0} + \gamma(\sigma_1^2+\delta\sigma_2^2)I_{i,0}\\ &\quad + \frac{2N-2}{2N-1}\frac{\gamma\sigma_2^2}{1+\alpha_i x_2}\left(\frac{\bar{\gamma}}{\gamma}\mathbb{E}_0\left[\frac{Q}{N}\right] + \frac{\bar{I}^e_{-i,1}}{2} + \frac{N-1}{N}\frac{\bar{I}^e_{i,1}}{2}\right)\end{aligned}$$

where

$$\delta = 1 - \frac{N-1}{N}\frac{1}{1+\alpha x} \in [0,1].$$

It is thus straightforward to check that the second derivative of $\widehat{W}_{i,0}$ is negative, so that the problem is strictly concave. Using proposition 1 of Malamud and Rostek (2017)

$$\lambda_{k,0} = \frac{\gamma(\sigma_1^2 + \delta\sigma_2^2)}{2N-2}$$

Plugging equilibrium price impacts $\lambda_{k,0}$ in the first order condition and rearranging, one gets the desired formula. ■

Equilibrium price and quantity

The date 0 market clearing condition can be written:

$$\frac{v_0 - p_0^*}{\gamma(\sigma_1^2 + \delta\sigma_2^2)} = \frac{I_{1,0} + I_{2,0}}{2} + \frac{2N-2}{2N-1}\frac{1}{1+\alpha x}\frac{\gamma\sigma_2^2}{\gamma(\sigma_1^2+\delta\sigma_2^2)}\left(\frac{2N-1}{2N}\frac{\bar{I}_{1,1}^e + \bar{I}_{2,1}^e}{2}\right) + \frac{2N-1}{2N-2}\frac{Q_c}{2N}$$

By market clearing at date 0, $\bar{I}_{1,1}^e + \bar{I}_{2,1}^e = I_{1,0} + I_{2,0} + \frac{Q_c}{N}$; in addition, recalling $\bar{\gamma} = \frac{1}{2}\frac{2N-1}{2N-2}$, one has

$$\begin{aligned}
v_0 - p_0^* &= \gamma(\sigma_1^2 + \delta\sigma_2^2)\frac{I_{1,0}+I_{2,0}}{2} + \frac{2N-2}{2N}\frac{\gamma\sigma_2^2}{1+\alpha x}\frac{I_{1,0}+I_{2,0}}{2} \\
&\quad + \gamma\left(\sigma_1^2 + \delta\sigma_2^2 + \frac{2N-2}{2N-1}\frac{2N-2}{2N}\frac{\sigma_2^2}{1+\alpha x}\right)\frac{2N-1}{2N-2}\frac{Q_c}{2N} \\
&= \gamma(\sigma_1^2+\sigma_2^2)\frac{I_{1,0}+I_{2,0}}{2} \\
&\quad + \gamma\left(\sigma_1^2 + \sigma_2^2 - \frac{1}{2N-1}\frac{2N-2}{2N}\frac{\sigma_2^2}{1+\alpha x}\right)\frac{2N-1}{2N-2}\frac{Q_c}{2N}
\end{aligned}$$

Recalling the definition of δ, the equilibrium price is therefore:

$$p_0^* = v_0 - \gamma(\sigma_1^2+\sigma_2^2)\frac{I_{1,0}+I_{2,0}}{2} - \gamma s^n \frac{Q_c}{2N} \tag{3.7.6}$$

where

$$s^n = \frac{2N-1}{2N-2}\left(\sigma_1^2+\sigma_2^2\right) - \frac{1}{2N}\frac{\sigma_2^2}{1+\alpha x}$$

Plugging 3.7.6 into the equilibrium demand schedule for class 1 traders:

$$
\begin{aligned}
q_{1,0}^* = \frac{2N-2}{2N-1}\Bigg[&\frac{I_{2,0}-I_{1,0}}{2} + \frac{2N-2}{2N-1}\frac{1}{1+\alpha x}\frac{\sigma_2^2}{\sigma_1^2+\delta\sigma_2^2}\left(\frac{\bar{\gamma}}{\gamma}\mathbb{E}_0\left[\frac{Q}{N}\right] + \frac{2N-1}{2N}\frac{\bar{I}_{1,1}^e+\bar{I}_{2,1}^e}{2}\right) \\
&-\frac{2N-2}{2N-1}\frac{1}{1+\alpha x}\frac{\sigma_2^2}{\sigma_1^2+\delta\sigma_2^2}\left(\frac{\bar{\gamma}}{\gamma}\mathbb{E}_0\left[\frac{Q}{N}\right] + \frac{\bar{I}_{2,1}^e}{2} + \frac{N-1}{N}\frac{\bar{I}_{1,1}^e}{2}\right)\Bigg] \\
= \frac{2N-2}{2N-1}\Bigg[&\frac{I_{2,0}-I_{1,0}}{2} + \frac{2N-2}{2N-1}\frac{1}{1+\alpha x}\frac{\sigma_2^2}{\sigma_1^2+\delta\sigma_2^2}\left(\frac{\bar{I}_{1,1}^e-\bar{I}_{2,1}^e}{4N}\right)\Bigg] \\
= \frac{2N-2}{2N-1}\Bigg[&\frac{I_{2,0}-I_{1,0}}{2} + \frac{1}{2N-1}\frac{N-1}{N}\frac{1}{1+\alpha x}\frac{\sigma_2^2}{\sigma_1^2+\delta\sigma_2^2}\left(\frac{I_{1,0}-I_{2,0}}{2} + q_{i,0}^*\right)\Bigg]
\end{aligned}
$$

where the third line used the equilibrium condition $q_{2,0}^e = q_{2,0}^*$ and market clearing 3.2.2. Thus

$$
\left(\frac{2N-1}{2N-2} - \frac{N-1}{N(2N-1)}\frac{1}{1+\alpha x}\frac{\sigma_2^2}{\sigma_1^2+\delta\sigma_2^2}\right)q_{i,0}^* = \left(1 - \frac{N-1}{N(2N-1)}\frac{1}{1+\alpha x}\frac{\sigma_2^2}{\sigma_1^2+\delta\sigma_2^2}\right)\frac{I_{2,0}-I_{1,0}}{2}
$$

Notice that with $\delta = 1 - \frac{N-1}{N}\frac{1}{1+\alpha x}$,

$$
1 - \frac{N-1}{N(2N-1)}\frac{1}{1+\alpha x}\frac{\sigma_2^2}{\sigma_1^2+\delta\sigma_2^2} = \frac{\sigma_1^2 + \left(1-\frac{2N-2}{2N-1}\frac{1}{1+\alpha x}\right)\sigma_2^2}{\sigma_1^2 + \left(1-\frac{2N-2}{2N}\frac{1}{1+\alpha x}\right)\sigma_2^2} > 0
$$

This and rearranging leads to the desired equilibrium quantity:

$$
q_{i,0}^* = \frac{1}{1+A(\sigma_q^2)}\frac{I_{2,0}-I_{1,0}}{2} \tag{3.7.7}
$$

where

$$
A(\sigma_q^2) = \frac{1}{2N-2}\frac{\sigma_1^2 + \left(1-\frac{N-1}{N}\frac{1}{1+\alpha x}\right)\sigma_2^2}{\sigma_1^2 + \left(1-\frac{2N-2}{2N-1}\frac{1}{1+\alpha x}\right)\sigma_2^2}
$$

where the dependence in σ_q^2 in the right-hand side goes through δ and x, to get formula 3.3.10. The properties of $A(\sigma_q^2)$ are derived in lemma 7 in appendix 2.9.3.

It is also possible to write

$$
A(\sigma_q^2) = \underbrace{\frac{1}{2N-2}}_{A^{static}} + \underbrace{\frac{1}{1+\alpha x}\frac{1}{(2N-2)(2N-1)}\frac{\sigma_2^2}{\sigma_1^2+\left(1-\frac{\alpha}{2}\frac{1}{1+\alpha x}\right)\sigma_2^2}}_{A^{dynamic}}
$$

The static demand reduction rate deserves its name because $1/(1+A^{static}) = \frac{2N-2}{2N-1}$, which is the same reduction factor as in the date 1 market which is a static game. It is straightforward to show that $A^{dynamic}$ converges to zero as σ_q^2, thus x, tends to infinity.

The date 1 quantity is computed straightforwardly from 3.7.7 and 3.3.4.

3.7.2 Proof of lemma 12

Proof. The certainty equivalent of wealth at date 0 is for trader 1

$$\widehat{W}_{1,0} = \underbrace{I_{1,0}v_0 + q^*_{1,0}(v_0 - p^*_0) - \frac{\gamma}{2}(\sigma_1^2+\sigma_2^2)(I^*_{1,1})^2}_{\widehat{W}^{HTM}_{1,0}} + \underbrace{\frac{2N}{2N-2}\frac{\gamma\sigma_2^2}{1+\alpha x}\left(\mathbb{E}_0\left[q^*_{1,1}\right]\right)^2}_{\widehat{S}^*_{1,1}}$$

Compute the HTM value. One has

$$\begin{aligned}
\widehat{W}^{HTM}_{1,0} &= I_{1,0}v_0 - \frac{\gamma(\sigma_1^2+\sigma_2^2)}{2}(I_{1,0})^2 + q^*_{1,0}(v_0-p^*_0) - \frac{\gamma}{2}(\sigma_1^2+\sigma_2^2)\left[(I^*_{1,1})^2-(I_{1,0})^2\right]\\
&= I_{1,0}v_0 - \frac{\gamma(\sigma_1^2+\sigma_2^2)}{2}(I_{1,0})^2 + q^*_{1,0}\left(v_0-p^*_0-\frac{\gamma}{2}(\sigma_1^2+\sigma_2^2)(2I_{1,0}+q^*_{1,0})\right)\\
&= I_{1,0}v_0 - \frac{\gamma(\sigma_1^2+\sigma_2^2)}{2}(I_{1,0})^2 + q^*_{1,0}\left(\gamma(\sigma_1^2+\sigma_2^2)\frac{I_{2,0}-I_{1,0}-q^*_{1,0}}{2} + \gamma s\frac{Q_c}{2N}\right)
\end{aligned}$$

Denoting $A \equiv A(\sigma_q^2)$ and $\Delta I = I_{2,0} - I_{1,0}$ to ease notation and plugging equilibrium quantity 3.7.7:

$$\begin{aligned}
\widehat{W}^{HTM}_{1,0} &= I_{1,0}v_0 - \frac{\gamma(\sigma_1^2+\sigma_2^2)}{2}(I_{1,0})^2\\
&\quad + \left(\frac{1}{1+A}\frac{\Delta I}{2} + \frac{Q_c}{2N}\right)\left(\gamma(\sigma_1^2+\sigma_2^2)\left(1-\frac{1}{2}\frac{1}{1+A}\right)\frac{\Delta I}{2} + \gamma\left(s-\frac{\sigma_1^2+\sigma_2^2}{2}\right)\frac{Q_c}{2N}\right)\\
&= I_{1,0}v_0 - \frac{\gamma(\sigma_1^2+\sigma_2^2)}{2}(I_{1,0})^2 + \frac{\gamma(\sigma_1^2+\sigma_2^2)}{2}\frac{1+2A}{(1+A)^2}\left(\frac{\Delta I}{2}\right)^2 + \gamma\left(s-\frac{\sigma_1^2+\sigma_2^2}{2}\right)\left(\frac{Q_c}{2N}\right)^2\\
&\quad + \frac{\gamma}{1+A}\left(\frac{\sigma_1^2+\sigma_2^2}{2}(1+2A) + s - \frac{\sigma_1^2+\sigma_2^2}{2}\right)\frac{Q_c}{2N}\frac{\Delta I}{2}
\end{aligned}$$

And

$$\widehat{S}^*_{1,1} = \frac{\alpha}{2}\frac{\gamma\sigma_2^2}{1+\alpha x}\left(\frac{A}{1+A}\frac{\Delta I}{2}\right)^2$$

Denote

$$\frac{\theta^n_{dd}}{2} = \frac{(\sigma_1^2+\sigma_2^2)}{2}\underbrace{\frac{1+2A}{(1+A)^2}}_{1-\left(\frac{A}{1+A}\right)^2} + \frac{\alpha}{2}\frac{\sigma_2^2}{1+\alpha x}\left(\frac{A}{1+A}\right)^2$$

$$\theta^n_{dd} = \sigma_1^2+\sigma_2^2 - \left(\sigma_1^2 + \left(1-\frac{\alpha}{1+\alpha x}\right)\sigma_2^2\right)\left(\frac{A}{1+A}\right)^2$$

and

$$\theta^n_{cd} = \frac{(\sigma_1^2+\sigma_2^2)A+s}{1+A}$$

■

3.7.3 Proof of lemma 13

$$
\begin{aligned}
p_0^* - p_d^* &= \gamma\left(\frac{1}{2}\frac{N}{N-2}\left(\sigma_1^2+\sigma_2^2\right) - \frac{N-1}{N-2}\frac{\theta_{dd}}{4}\right)\frac{Q_d}{N} \\
&\quad + \gamma\left(\frac{\sigma_1^2+\sigma_2^2}{2}\frac{A}{1+A} - \left(1-\frac{1}{2(1+A)}\right)s\right)\frac{Q_c}{2N} \\
&= \frac{\gamma N}{N-2}\left(\left(\sigma_1^2+\sigma_2^2\right) - \left(1-\frac{1}{N}\right)\frac{\theta_{dd}}{2}\right)\frac{Q_d}{2N} \\
&\quad + \gamma\left(\frac{\sigma_1^2+\sigma_2^2}{2}\frac{A}{1+A} - \frac{1+2A}{2(1+A)}\left(\frac{2N-1}{2N-2}(\sigma_1^2+\sigma_2^2) - \frac{1}{2N}\frac{\sigma_2^2}{1+\alpha x}\right)\right)\frac{Q_c}{2N} \\
&= \frac{\gamma N}{N-2}\left(\left(\sigma_1^2+\sigma_2^2\right) - \left(1-\frac{1}{N}\right)\frac{1}{2}\left(\sigma_1^2+\sigma_2^2-\left(\frac{A}{1+A}\right)^2\left(\sigma_1^2+\left(1-\frac{\alpha}{1+\alpha x}\right)\sigma_2^2\right)\right)\right)\frac{Q_d}{2N} \\
&\quad + \frac{\gamma}{2}\left(\left(\sigma_1^2+\sigma_2^2\right)\frac{A}{1+A} - \frac{1+2A}{1+A}\frac{2N-1}{2N-2}(\sigma_1^2+\sigma_2^2) + \frac{1+2A}{1+A}\frac{2N-1}{2N(2N-2)}\frac{\sigma_2^2}{1+\alpha x}\right)\frac{Q_c}{2N} \\
&= \frac{\gamma N}{N-2}\left(\left(\sigma_1^2+\sigma_2^2\right)\frac{N+1}{2N} + \frac{N-1}{2N}\left(\frac{A}{1+A}\right)^2\left(\sigma_1^2+\left(1-\frac{\alpha}{1+\alpha x}\right)\sigma_2^2\right)\right)\frac{Q_d}{2N} \\
&\quad + \frac{\gamma}{2}\left(-\left(\sigma_1^2+\sigma_2^2\right)\left(1+\frac{1}{2N-2}\frac{1+2A}{1+A}\right) + \frac{1+2A}{1+A}\frac{2N-1}{2N(2N-2)}\frac{\sigma_2^2}{1+\alpha x}\right)\frac{Q_c}{2N} \\
&= \frac{\gamma}{2(N-2)}\left(\left(\sigma_1^2+\sigma_2^2\right)(N+1) + (N-1)\left(\frac{A}{1+A}\right)^2\left(\sigma_1^2+\left(1-\frac{\alpha}{1+\alpha x}\right)\sigma_2^2\right)\right)\frac{Q_d}{2N} \\
&\quad - \frac{\gamma}{2}\left(\left(1+\frac{1}{2N-2}\frac{1+2A}{1+A}\right)\sigma_1^2 + \left(1+\frac{1}{2N-2}\frac{1+2A}{1+A} - \frac{1+2A}{1+A}\frac{2N-1}{2N(2N-2)}\frac{1}{1+\alpha x}\right)\sigma \right. \\
&= \frac{\gamma(N+1)}{2(N-2)}\left[\left(\frac{N+1}{N-1}+\left(\frac{A}{1+A}\right)^2\right)\sigma_1^2 + \left(\frac{N+1}{N-1}+\left(\frac{A}{1+A}\right)^2\left(1-\frac{\alpha}{1+\alpha x}\right)\right)\sigma_2^2\right]\frac{Q_d}{2N} \\
&\quad - \frac{\gamma}{2}\left(\left(1+\frac{1}{2N-2}\frac{1+2A}{1+A}\right)\sigma_1^2 + \left(1+\frac{1}{2N-2}\frac{1+2A}{1+A}\left(1-\frac{2N-1}{2N}\frac{1}{1+\alpha x}\right)\right)\sigma_2^2\right)\frac{Q_c}{2N}
\end{aligned}
$$

Then identify $s_{d1}, s_{d2}, s_{c1}, s_{c2}$ with the above expression. s_{d1} and s_{c1} are obviously positive. $s_{d2} > 0$ because $\alpha = 1 - \frac{1}{(2N-1)^2} < 1$, and $\frac{1}{1+\alpha x} < 1$. For s_{c2}, given $\frac{1}{1+\alpha x} < 1$,

$$
2s_{c2} > 1 + \frac{1}{2N-2}\frac{1+2A}{1+A}\left(1-\frac{2N-1}{2N}\right) = 1 + \frac{1}{2N-2}\frac{1+2A}{1+A}\frac{1}{2N} > 0
$$

3.7.4 Proof of proposition 14

Without loss of generality, assume that dealers of class 1 are assigned the customer order in the parallel market, and dealers of class 2 do not trade in the parallel market. I review the utility of class 1 dealers first, then the utility of class 2 dealers.

Class 1 dealers. I first prove that the certainty equivalent of wealth 3.4.1 is convex in w. As it is quadratic, this implies that the maximum is attained for $w = 0$ or 1. Then I show that the maximum is attained at $w = 1$.

Lemma 20. *The certainty equivalent of wealth 3.4.1 is a convex function of* w.

Proof. One has, from 3.4.1 and lemma 12:

$$\begin{aligned}
\frac{d^2\widehat{W}_1}{dw^2} &= \frac{N}{N-2}\left(\sigma_1^2+\sigma_2^2-\frac{\theta_{dd}}{4}\right)+\frac{\theta_{cc}}{2}-\theta_{cd}\\
&= \frac{N}{N-2}\left(1-\frac{1}{4}\left(1-\left(\frac{A}{1+A}\right)^2\right)\right)(\sigma_1^2+\sigma_2^2)+\frac{1}{4}\frac{N}{N-2}\left(\frac{A}{1+A}\right)^2\frac{\alpha\sigma_2^2}{1+\alpha x}\\
&\quad+\frac{s}{2}-\frac{\sigma_1^2+\sigma_2^2}{4}-\frac{A}{1+A}(\sigma_1^2+\sigma_2^2)-\frac{s}{1+A}
\end{aligned}$$

where $A = A(\sigma_q^2$. Notice that

$$\begin{aligned}
K &\equiv \frac{s}{2}-\frac{\sigma_1^2+\sigma_2^2}{4}-\frac{A}{1+A}(\sigma_1^2+\sigma_2^2)-\frac{s}{1+A}\\
&= s\left(\frac{A}{1+A}-\frac{1}{2}\right)+\left(\frac{1}{4}-\frac{A}{1+A}\right)(\sigma_1^2+\sigma_2^2)\\
&= \left(\frac{2N-1}{2N-2}\left(\frac{A}{1+A}-\frac{1}{2}\right)+\frac{1}{4}-\frac{A}{1+A}\right)(\sigma_1^2+\sigma_2^2)+\left(\frac{1}{2}-\frac{A}{1+A}\right)\frac{1}{2N}\frac{\sigma_2^2}{1+\alpha x}\\
&= \left(\frac{1}{2N-2}\frac{A}{1+A}-\frac{1}{2}\left(\frac{2N-1}{2N-2}-\frac{1}{2}\right)\right)(\sigma_1^2+\sigma_2^2)+\left(\frac{1}{2}-\frac{A}{1+A}\right)\frac{1}{2N}\frac{\sigma_2^2}{1+\alpha x}\\
&= \left(\frac{1}{2N-2}\frac{A}{1+A}-\frac{1}{2}\left(\frac{1}{2}+\frac{1}{2N-2}\right)\right)(\sigma_1^2+\sigma_2^2)+\left(\frac{1}{2}-\frac{A}{1+A}\right)\frac{1}{2N}\frac{\sigma_2^2}{1+\alpha x}\\
&= \left(\frac{1}{2N-2}\left(\frac{A}{1+A}-\frac{1}{2}\right)-\frac{1}{4}\right)(\sigma_1^2+\sigma_2^2)+\left(\frac{1}{2}-\frac{A}{1+A}\right)\frac{1}{2N}\frac{\sigma_2^2}{1+\alpha x}
\end{aligned}$$

Thus

$$\begin{aligned}
\frac{d^2\widehat{W}_1}{dw^2} &= \left[\frac{N}{N-2}\left(1-\frac{1}{4}\left(1-\left(\frac{A}{1+A}\right)^2\right)\right)-\frac{1}{2N-2}\left(\frac{1}{2}-\frac{A}{1+A}\right)-\frac{1}{4}\right](\sigma_1^2+\sigma_2^2)\\
&\quad+\left[\frac{\alpha}{4}\frac{N}{N-2}\left(\frac{A}{1+A}\right)^2+\left(\frac{1}{2}-\frac{A}{1+A}\right)\frac{1}{2N}\right]\frac{\sigma_2^2}{1+\alpha x}\\
&= \left[\frac{1}{2}-\left(\frac{A}{1+A}\right)^2+\frac{2}{N-2}\left(\frac{3}{4}-\left(\frac{A}{1+A}\right)^2-\frac{N-2}{4(N-1)}\left(\frac{1}{2}-\frac{A}{1+A}\right)\right)\right](\sigma_1^2+\sigma_2^2)\\
&\quad+\left[\frac{\alpha}{4}\frac{N}{N-2}\left(\frac{A}{1+A}\right)^2+\left(\frac{1}{2}-\frac{A}{1+A}\right)\frac{1}{2N}\right]\frac{\sigma_2^2}{1+\alpha x}
\end{aligned}$$

In another paper I show that

$$0 < \frac{1}{2N-1} \leq \frac{A}{1+A} \leq \frac{3}{2}\frac{1}{2N-1/2} \leq \frac{3}{7}$$

Since $3/7 < 1/2$, the term in $\frac{\sigma_2^2}{1+\alpha x}$ is positive. For the term in $\sigma_1^2 + \sigma_2^2$:

$$\begin{aligned}
\psi &\equiv \frac{1}{2} - \left(\frac{A}{1+A}\right)^2 + \frac{2}{N-2}\left(\frac{3}{4} - \left(\frac{A}{1+A}\right)^2 - \frac{N-2}{4(N-1)}\left(\frac{1}{2} - \frac{A}{1+A}\right)\right) \\
&\geq \frac{1}{2} - \left(\frac{3}{7}\right)^2 + \frac{2}{N-2}\left(\frac{3}{4} - \left(\frac{3}{7}\right)^2 - \frac{N-2}{4(N-1)}\frac{1}{2}\right) \\
&> \frac{1}{2} - \left(\frac{3}{7}\right)^2 + \frac{2}{N-2}\left(\frac{3}{4} - \left(\frac{3}{7}\right)^2 - \frac{1}{8}\right) \\
&> 0
\end{aligned}$$

Thus $\frac{\partial^2 \widehat{W}_1}{\partial w^2} > 0$. ■

Given that the certainty equivalent of wealth 3.4.1 is quadratic and convex in w, its maximum is either at $w = 0$ or $w = 1$. It remains to show that $\widehat{W}_1(1) > \widehat{W}_1(0)$. One has

$$\begin{aligned}
\gamma^{-1}\widehat{W}_1(1) &= \frac{1}{2}\frac{N}{N-2}\left(\sigma_1^2 + \sigma_2^2 - \frac{\theta_{dd}}{4}\right) \\
\gamma^{-1}\widehat{W}_1(0) &= \frac{\theta_{cc}}{4}
\end{aligned}$$

Thus

$$\begin{aligned}
\frac{\widehat{W}_1(1) - \widehat{W}_1(0)}{\gamma} &= \left(\frac{1}{2} + \frac{2}{N-2}\right)\left(\left(1 - \left(\frac{A}{1+A}\right)^2\right)(\sigma_1^2 + \sigma_2^2) + \left(\frac{A}{1+A}\right)^2 \frac{\alpha\sigma_2^2}{1+\alpha x}\right) \\
&\quad - \frac{1}{4}\left(\frac{2N-1}{2N-2}(\sigma_1^2 + \sigma_2^2) - \frac{1}{2N}\frac{\sigma_2^2}{1+\alpha x}\right) \\
&= \left(\left(\frac{1}{2} + \frac{2}{N-2}\right)\left(1 - \left(\frac{A}{1+A}\right)^2\right) - \frac{1}{4}\left(1 + \frac{1}{2N-2}\right)\right)(\sigma_1^2 + \sigma_2^2) \\
&\quad \left(\left(\frac{1}{2} + \frac{2}{N-2}\right)\left(\frac{A}{1+A}\right)^2 + \frac{1}{2N}\right)\frac{\sigma_2^2}{1+\alpha x}
\end{aligned}$$

The term in $\frac{\sigma_2^2}{1+\alpha x}$ is clearly positive. The term in $\sigma_1^2 + \sigma_2^2$ is

$$\begin{aligned}
\phi &\equiv \left(\frac{1}{2} + \frac{2}{N-2}\right)\left(1 - \left(\frac{A}{1+A}\right)^2\right) - \frac{1}{4}\left(1 + \frac{1}{2N-2}\right) \\
&\geq \left(\frac{1}{2} + \frac{2}{N-2}\right)\left(1 - \left(\frac{3}{7}\right)^2\right) - \frac{1}{4}\left(1 + \frac{1}{2N-2}\right) \\
&= \underbrace{\frac{1}{2} \times \left(1 - \frac{9}{49}\right) - \frac{1}{4}}_{>0} + \frac{2}{N-2}\underbrace{\left(1 - \frac{9}{49} - \frac{N-2}{4(N-1)}\frac{1}{4}\right)}_{\geq 1 - \frac{9}{49} - \frac{1}{16} > 0} > 0
\end{aligned}$$

Therefore $\widehat{W}_1(1) - \widehat{W}_1(0)$. Thus the maximum of $\widehat{W}_1(w)$ for $w \in [0,1]$ is attained for $w = 1$, which proves the first part of the proposition.

Class 2 dealers.

Lemma 21. *The certainty equivalent of wealth 3.3.12 for $I_{2,0} = 0$ is a concave function of w.*

Proof. The certainty equivalent of wealth for class 2 dealers is, from 3.3.12,

$$\widehat{W}_2(w) = \gamma \left[\frac{\theta_{cc}}{4}(1-w)^2 + \frac{\theta_{cd}}{4} w(1-w) + \frac{\theta_{dd}}{8} w^2 \right] \left(\frac{Q_0}{N} \right)^2$$

Thus

$$\frac{d\widehat{W}_2}{dw} = \gamma \left[-\frac{\theta_{cc}}{2}(1-w) + \frac{\theta_{cd}}{4}(1-2w) + \frac{\theta_{dd}}{4} w \right] \left(\frac{Q_0}{N} \right)^2$$

and

$$\frac{d^2\widehat{W}_2}{dw^2} = \frac{\gamma}{2} \left[\theta_{cc} - \theta_{cd} + \frac{\theta_{dd}}{2} \right] \left(\frac{Q_0}{N} \right)^2$$

One has

$$\begin{aligned}
B &\equiv \theta_{cc} - \theta_{cd} + \frac{\theta_{dd}}{4} \\
&= s - \frac{\sigma_1^2 + \sigma_2^2}{2} - \frac{A}{1+A}(\sigma_1^2 + \sigma_2^2) - \frac{s}{1+A} \\
&\quad + \frac{1}{4}\left(1 - \left(\frac{A}{1+A} \right)^2 \right)(\sigma_1^2 + \sigma_2^2) + \frac{\alpha}{4}\left(\frac{A}{1+A} \right)^2 \frac{\sigma_2^2}{1+\alpha x} \\
&= \frac{A}{1+A}s - \left(\frac{1}{2} + \frac{A}{1+A} \right)(\sigma_1^2 + \sigma_2^2) \\
&\quad + \frac{1}{4}\left(1 - \left(\frac{A}{1+A} \right)^2 \right)(\sigma_1^2 + \sigma_2^2) + \frac{\alpha}{4}\left(\frac{A}{1+A} \right)^2 \frac{\sigma_2^2}{1+\alpha x} \\
&= \frac{A}{1+A}s - \left(\frac{1}{2} + \frac{A}{1+A}\left(1 + \frac{A}{1+A} \right) \right)(\sigma_1^2 + \sigma_2^2) + \frac{\alpha}{4}\left(\frac{A}{1+A} \right)^2 \frac{\sigma_2^2}{1+\alpha x} \\
&= \left(-\frac{1}{2} + \frac{A}{1+A}\left(\frac{2N-1}{2N-2} - 1 - \frac{A}{1+A} \right) \right)(\sigma_1^2 + \sigma_2^2) + \left(\frac{\alpha}{4}\left(\frac{A}{1+A} \right)^2 - \frac{A}{1+A}\frac{1}{2N} \right) \frac{\sigma_2^2}{1+\alpha x} \\
&= -\frac{\sigma_1^2 + \sigma_2^2}{2} + \frac{A}{1+A}\left[\left(\frac{1}{2N-2} - \frac{A}{1+A} \right)(\sigma_1^2 + \sigma_2^2) + \left(\frac{\alpha}{2}\frac{A}{1+A} - \frac{1}{2N} \right)\frac{\sigma_2^2}{1+\alpha x} \right]
\end{aligned}$$

As $\frac{1}{2N-1} < \frac{A}{1+A} \leq \frac{3}{2}\frac{1}{2N-\frac{1}{2}}$, one has

$$\begin{aligned}
B &\leq \frac{A}{1+A}\left[\left(\frac{1}{2N-2} - \frac{1}{2N-1} \right)(\sigma_1^2 + \sigma_2^2) + \left(\frac{\alpha}{2}\frac{3}{2}\frac{1}{2N-1/2} - \frac{1}{2N} \right)\frac{\sigma_2^2}{1+\alpha x} \right] \\
&= \frac{A}{1+A}\left[\left(\frac{1}{2N-2} - \frac{1}{2N-1} \right)(\sigma_1^2 + \sigma_2^2) + \frac{1}{2N}\left(\alpha \times \frac{3}{4}\frac{2N}{2N-1/2} - 1 \right)\frac{\sigma_2^2}{1+\alpha x} \right]
\end{aligned}$$

One has $\frac{1}{2N-2} - \frac{1}{2N-1} < 0$ for $N \geq 2$, and since $\alpha < 1$ and $\frac{2N}{2N-1/2} \geq \frac{4}{7/2} = 8/7$, one has $\alpha \times \frac{3}{4}\frac{2N}{2N-1/2} - 1 < 0$. So $B < 0$, which finishes the proof. ■

Since $\widehat{W_2}(w)$ is concave and quadratic in w, it admits a global maximum. This global maximum w_2^u solves $\frac{d\widehat{W_2}}{dw} = 0$, *i.e.*

$$-\frac{\theta_{cc}}{2}(1 - w_2^u) + \frac{\theta_{cd}}{4}(1 - 2w_2^u) + \frac{\theta_{dd}}{4}w_2^u = 0$$

Rearranging leads to

$$w_2^u = \frac{\theta_{cc} - \frac{\theta_{cd}}{2}}{\frac{\theta_{dd}}{4} + \theta_{cc} - \theta_{cd}} \tag{3.7.8}$$

I successively compute the numerator and the denominator of 3.7.8.

$$\begin{aligned}
\theta_{cc} - \frac{\theta_{cd}}{2} &= s - \frac{\sigma_1^2 + \sigma_2^2}{2} - \frac{\sigma_1^2 + \sigma_2^2}{2}\frac{A}{1+A} - \frac{s}{2(1+A)} \\
&= \frac{1+2A}{1+A}\left(\frac{s}{2} - \frac{\sigma_1^2 + \sigma_2^2}{2}\right) \\
&= \frac{1}{2}\frac{1+2A}{1+A}\left(\frac{1}{2N-2}(\sigma_1^2 + \sigma_2^2) - \frac{1}{2N}\frac{\sigma_2^2}{1+\alpha x}\right) \\
&= \frac{1}{2}\frac{1+2A}{1+A}\left(\frac{\sigma_1^2}{2N-2} + \left(\frac{1}{2N-2} - \frac{1}{2N}\frac{1}{1+\alpha x}\right)\sigma_2^2\right)
\end{aligned}$$

and notice that $\frac{1}{2N-2} > \frac{1}{2N}\frac{1}{1+\alpha x}$ so that $\theta_{cc} - \frac{\theta_{cd}}{2}$ is strictly positive for all parameter values. The denominator is exactly equal to $\frac{d^2\widehat{W_2}}{dw^2}$, which has already been proven to be negative.

Therefore $w_2^u < 0$. It implies that the maximum of $\widehat{W_2}(w)$ for $w \in [0, 1]$ is attained at $w_2^* = 0$. QED.

3.7.5 Proof of theorem 3

The theorem has two parts: the ex ante welfare of dealers, and the welfare of customers.

Dealers' ex ante welfare. Dealers' ex ante welfare is improving when customers trade in the OTC market if and only if

$$-\frac{1}{2}\left(e^{-\gamma\widehat{W_1}(1)} + e^{-\gamma\widehat{W_2}(1)}\right) > -e^{-\gamma\widehat{W_2}(0)}$$

which is equivalent to

$$\frac{1}{2}\left(e^{-\gamma\left(\widehat{W_1}(1)-\widehat{W_2}(0)\right)} + e^{-\gamma\left(\widehat{W_2}(1)-\widehat{W_2}(0)\right)}\right) < 1 \tag{3.7.9}$$

One has

$$\begin{aligned}
\widehat{W}_1(1) - \widehat{W}_2(0) &= \frac{1}{2}\frac{N}{N-2}\left(\sigma_1^2 + \sigma_2^2 - \frac{\theta_{dd}}{4}\right) - \frac{\theta_{cc}}{4} \\
&= \frac{1}{2}\left[\left(1 + \frac{2}{N-2}\right)\left(\frac{3}{4} + \frac{1}{4}\left(\frac{A}{1+A}\right)^2\right)(\sigma_1^2 + \sigma_2^2) + \left(\frac{A}{1+A}\right)^2 \frac{\alpha}{4}\frac{\sigma_2^2}{1+\alpha x}\right] \\
&\quad - \frac{1}{4}\left(\left(1 + \frac{1}{2N-2} - \frac{1}{2}\right)(\sigma_1^2 + \sigma_2^2) - \frac{1}{2N}\frac{\sigma_2^2}{1+\alpha x}\right) \\
&= \frac{1}{2}\left[\left(1 + \frac{2}{N-2}\right)\left(\frac{3}{4} + \frac{1}{4}\left(\frac{A}{1+A}\right)^2\right) - \frac{1}{2}\left(\frac{1}{2} + \frac{1}{2N-2}\right)\right](\sigma_1^2 + \sigma_2^2) \\
&\quad + \frac{1}{8}\left(\alpha\left(\frac{A}{1+A}\right)^2 + \frac{1}{N}\right)\frac{\sigma_2^2}{1+\alpha x} \\
&= \frac{1}{2}\left[\frac{1}{2} + \frac{1}{4}\left(\frac{A}{1+A}\right)^2 + \frac{2}{N-2}\left(\frac{3}{4} + \frac{1}{4}\left(\frac{A}{1+A}\right)^2 - \frac{1}{8}\frac{N-2}{N-1}\right)\right](\sigma_1^2 + \sigma_2^2) \\
&\quad + \frac{1}{8}\left(\alpha\left(\frac{A}{1+A}\right)^2 + \frac{1}{N}\right)\frac{\sigma_2^2}{1+\alpha x}
\end{aligned}$$

and

$$\begin{aligned}
4(\widehat{W}_1(1) - \widehat{W}_2(0)) = &\left[1 + \frac{1}{2}\left(\frac{A}{1+A}\right)^2 + \frac{2}{N-2}\left(\frac{3}{2} + \frac{1}{2}\left(\frac{A}{1+A}\right)^2 - \frac{1}{4}\frac{N-2}{N-1}\right)\right](\sigma_1^2 + \sigma_2^2) \\
&+ \frac{1}{2}\left(\alpha\left(\frac{A}{1+A}\right)^2 + \frac{1}{N}\right)\frac{\sigma_2^2}{1+\alpha x}
\end{aligned}$$

The coefficients in $(\sigma_1^2 + \sigma_2^2)$ and in $\frac{\sigma_2^2}{1+\alpha x}$ are clearly positive, given in particular that $\frac{N-2}{N-1} < 1$.

Now consider the second payoff differential:

$$
\begin{aligned}
4(\widehat{W}_2(1) - \widehat{W}_2(0)) &= \frac{\theta_{dd}}{2} - \theta_{cc} \\
&= \left(1 - \left(\frac{A}{1+A}\right)^2\right) \frac{\sigma_1^2 + \sigma_2^2}{2} + \frac{\alpha}{2} \left(\frac{A}{1+A}\right)^2 \frac{\sigma_2^2}{1+\alpha x} \\
&\quad - \left(\left(\frac{1}{2} + \frac{1}{2N-2}\right)(\sigma_1^2 + \sigma_2^2) - \frac{1}{2N} \frac{\sigma_2^2}{1+\alpha x}\right) \\
&= -\left(\frac{1}{2}\left(\frac{A}{1+A}\right)^2 + \frac{1}{2N-2}\right)(\sigma_1^2 + \sigma_2^2) \\
&\quad + \frac{1}{2}\left(\alpha\left(\frac{A}{1+A}\right)^2 + \frac{1}{N}\right) \frac{\sigma_2^2}{1+\alpha x} \\
&= -\left(\left(\frac{A}{1+A}\right)^2 + \frac{1}{N-1}\right) \frac{\sigma_1^2}{2} \\
&\quad - \left(\left(\frac{A}{1+A}\right)^2 \left(1 - \frac{\alpha}{1+\alpha x}\right) + \frac{1}{N-1} - \frac{1}{N} \frac{1}{1+\alpha x}\right) \frac{\sigma_2^2}{2}
\end{aligned}
$$

The term in σ_1^2 is clearly negative, and the term in σ_2^2 is also negative because $\alpha/(1+\alpha x) < 1$ and $\frac{1}{N-1} > \frac{1}{N}\frac{1}{1+\alpha x}$.

The expected payoff from opening the OTC market is positive:

$$
\begin{aligned}
\varphi &\equiv \frac{\widehat{W}_1(1) + \widehat{W}_2(1)}{2} - \widehat{W}_2(0) \\
&= \left[1 - \frac{1}{2N-2} + \frac{2}{N-2}\left(\frac{3}{2} + \frac{1}{2}\left(\frac{A}{1+A}\right)^2 - \frac{1}{4}\frac{N-2}{N-1}\right)\right](\sigma_1^2 + \sigma_2^2) \\
&\quad + \left(\alpha\left(\frac{A}{1+A}\right)^2 + \frac{1}{N}\right) \frac{\sigma_2^2}{1+\alpha x}
\end{aligned}
$$

The term in $\sigma_1^2 + \sigma_2^2$ is positive because $1 - \frac{1}{2N-2} > 0$ and $\frac{3}{3} - \frac{1}{4}\frac{N-2}{N-1} > 0$. The second term is clearly positive.

Now using these expressions to plot the function

$$
\left(\frac{\sigma_1^2}{\sigma_2^2}, N\right) \longmapsto 1 - \frac{1}{2}\left(e^{-\gamma\left(\widehat{W}_1(1) - \widehat{W}_2(0)\right)} + e^{-\gamma\left(\widehat{W}_2(1) - \widehat{W}_2(0)\right)}\right)
$$

for various values of $z \in [0,1]$ and of γ, one sees that given N, there is a y_N for which $y < y_N$ implies that dealers' expected utility is higher when they open the parallel market.

Customers' welfare. If dealers trade in the centralized market ($w = 0$), they get, from 3.3.11:

$$Q_0 \; p_0^*(w=0) = Q_0 \left(v_0 - \gamma \frac{s}{2} \frac{Q_0}{N} \right)$$

If they trade in the parallel market ($w = 1$), they get, from 3.3.13:

$$Q_0 \; p_d^*(w=1) = Q_0 \left(v_0 - \gamma \frac{N-1}{N-2} \left(\sigma_1^2 + \sigma_2^2 - \frac{\theta_{dd}}{4} \right) \frac{Q_0}{N} \right)$$

Then compute, denoting to ease notation $p_0^*(w=0) = p_0^*$ and $p_d^*(w=1) = p_d^*$,

$$\begin{aligned}
\frac{N}{\gamma Q_0} (p_d^* - p_0^*) &= \frac{1}{2} \left(\frac{2N-1}{2N-2} (\sigma_1^2 + \sigma_2^2) - \frac{1}{2N} \frac{\sigma_2^2}{1+\alpha x} \right) - \frac{N-1}{N-2} \left(\sigma_1^2 + \sigma_2^2 - \frac{\theta_{dd}}{4} \right) \\
&\leq \frac{1}{2} \left(\frac{2N-1}{2N-2} (\sigma_1^2 + \sigma_2^2) - \frac{1}{2N} \frac{\sigma_2^2}{1+\alpha x} \right) - \left(\sigma_1^2 + \sigma_2^2 - \frac{\theta_{dd}}{4} \right) \\
&= \left(\frac{1}{2} \frac{1}{2N-2} - \frac{1}{2} \right) (\sigma_1^2 + \sigma_2^2) - \frac{1}{4N} \frac{\sigma_2^2}{1+\alpha x} \\
&\quad + \frac{1}{4} \left(\left(1 - \left(\frac{A}{1+A} \right)^2 \right) (\sigma_1^2 + \sigma_2^2) - \left(\frac{A}{1+A} \right)^2 \frac{\alpha \sigma_2^2}{1+\alpha x} \right) \\
&= \frac{1}{4} \left(\frac{1}{N-1} - 1 - \left(\frac{A}{1+A} \right)^2 \right) (\sigma_1^2 + \sigma_2^2) + \left(\frac{\alpha}{4} \left(\frac{A}{1+A} \right)^2 - \frac{1}{2N} \right) \frac{\sigma_2^2}{1+\alpha x} \\
&= \frac{1}{4} \left(\frac{1}{N-1} - 1 - \left(\frac{A}{1+A} \right)^2 \right) \sigma_1^2 \\
&\quad + \frac{1}{4} \left(\frac{1}{N-1} - \left(1 - \frac{2}{N} \frac{1}{1+\alpha x} \right) - \left(1 - \frac{\alpha}{1+\alpha x} \right) \left(\frac{A}{1+A} \right)^2 \right) \sigma_2^2
\end{aligned}$$

Given that $N \geq 2$ and $A \geq 0$, it is easy to see that the coefficient in σ_1^2 is negative. It follows, also from observation that $\alpha < 1$ and $\frac{1}{1+\alpha x} < 1$, that the coefficient in σ_2^2 is negative. Thus the unit spread that customers get from trading in the parallel market is higher that the spread they get from trading in the centralized market.

3.7.6 Proof of proposition 16

I first show that dealers' equilbrium utility is convex, which together with the fact that it is quadratic implies that its maximum is either at $w = 0$ or at $w = 1$. Then I show that $\widehat{W}_1(1) > \widehat{W}_1(0)$.

Utility is convex. I have to show that the equilibrium certainty equivalent of wealth from lemma 16 has a positive second derivative:

$$\frac{d\widehat{W}_1}{dw} \propto \left((\sigma_1^2+\sigma_2^2)(1+2\zeta) + \frac{2}{(N-1)(2N-1)}\frac{\sigma_2^2}{1+\alpha x} + \frac{\theta_{dd}}{4}\right) w \\ - \frac{\theta_{cc}}{2}(1-w) + \frac{s+\theta_{cd}}{4}(1-2w)$$

where $\zeta = \frac{N}{(N-1)^2(2N-1)}$, so that

$$\frac{d^2\widehat{W}_1}{dw^2} \propto (\sigma_1^2+\sigma_2^2)(1+2\zeta) + \frac{2}{(N-1)(2N-1)}\frac{\sigma_2^2}{1+\alpha x} + \frac{\theta_{dd}}{4} + \frac{\theta_{cc}}{2} - \frac{s+\theta_{cd}}{2}$$

One has

$$\begin{aligned}\frac{\theta_{cc}}{2} - \frac{s+\theta_{cd}}{2} &= \frac{1}{2}\left[s - \frac{\sigma_1^2+\sigma_2^2}{2} - \frac{s}{2} - \frac{\sigma_1^2+\sigma_2^2}{2}\frac{A}{1+A} - \frac{s}{2}\frac{1}{1+A}\right] \\ &= \frac{1}{2}\left[\left(s - \frac{\sigma_1^2+\sigma_2^2}{2}\right)\frac{A}{1+A} - \frac{\sigma_1^2+\sigma_2^2}{2}\right] \\ &= -\frac{1}{2}\left[\frac{1-\frac{1}{N-1}A}{1+A}(\sigma_1^2+\sigma_2^2) + \frac{1}{2N}\frac{\sigma_2^2}{1+\alpha x}\right]\end{aligned}$$

Thus

$$\begin{aligned}\frac{d^2\widehat{W}_1}{dw^2} &\propto \left(1+2\zeta+\frac{1}{4}\left(1-\left(\frac{A}{1+A}\right)^2\right) - \frac{1}{2}\frac{1-\frac{1}{N-1}A}{1+A}\right)(\sigma_1^2+\sigma_2^2) \\ &\quad + \left(\frac{\alpha}{4}\left(\frac{A}{1+A}\right)^2 + \frac{2}{(N-1)(2N-1)} - \frac{1}{2N}\frac{A}{1+A}\right)\frac{\sigma_2^2}{1+\alpha x} \\ &= \left[1+2\zeta+\frac{1}{4}\left(1-\left(\frac{A}{1+A}\right)^2\right) - \frac{1}{2}\frac{1-\frac{1}{N-1}A}{1+A}\right]\sigma_1^2 \\ &\quad + \left[1 - \frac{1}{2}\frac{1-\frac{1}{N-1}A}{1+A} + 2\zeta + \frac{1}{4}\left(1-(1-\alpha z)\left(\frac{A}{1+A}\right)^2\right)\right. \\ &\quad\quad \left. + z\left(\frac{2}{(N-1)(2N-1)} - \frac{1}{2N}\frac{A}{1+A}\right)\right]\sigma_2^2 \\ &\geq \left[1+2\zeta+\frac{1}{4}\left(1-\left(\frac{A}{1+A}\right)^2\right) - \frac{1}{2}\frac{1-\frac{1}{N-1}\frac{3}{7}}{1+A}\right]\sigma_1^2 \\ &\quad + \left[1 - \frac{1}{2}\frac{1-\frac{1}{N-1}A}{1+A} + 2\zeta + \frac{1}{4}\left(1-(1-\alpha z)\left(\frac{A}{1+A}\right)^2\right)\right. \\ &\quad\quad \left. + z\left(\frac{2}{(N-1)(2N-1)} - \frac{3/4}{N(2N-1/2)}\right)\right]\sigma_2^2\end{aligned}$$

where $z = \frac{1}{1+\alpha z} < 1$, and the inequality follows from $\frac{A}{1+A} \leq \frac{3}{2}\frac{1}{2N-1/2} \leq \frac{3}{7}$. It is easy to see that

$$1 - \frac{1}{2}\frac{1 - \frac{3}{7(N-1)}}{1+A} > 0$$

and to show that $\frac{2}{(N-1)(2N-1)} - \frac{3/4}{N(2N-1/2)} > 0$. Thus the certainty equivalent of wealth is convex for dealers trading in the parallel market.

Dealers trading in the parallel market prefer $w = 1$. Consider

$$\widehat{W}_1(0) = \frac{\theta_{cc}}{4} = \frac{1}{8}\left[\left(1 + \frac{1}{N-1}\right)(\sigma_1^2 + \sigma_2^2) - \frac{4}{N}\frac{\sigma_2^2}{1+\alpha x}\right]$$

and

$$\widehat{W}_1(1) = \frac{1}{8}\left[\left(4 + 8\zeta + 1 - \left(\frac{A}{1+A}\right)^2\right)(\sigma_1^2 + \sigma_2^2) + \left(\frac{1}{(N-1)(2N-1)} + \frac{\alpha}{8}\left(\frac{A}{1+A}\right)^2\right)\frac{8\sigma_2^2}{1+\alpha x}\right]$$

Comparing the terms in $(\sigma_1^2 + \sigma_2^2)$ and in $\sigma_2^2/(1+\alpha x)$ in the two expressions, given $\zeta > 0$ and $\left(\frac{A}{1+A}\right)^2 < 1$, it is straightforward to see that $\widehat{W}_1(1) > \widehat{W}_1(0)$. QED.

3.7.7 Proof of theorem 4

The expected payoff is higher for . The expected payoff of opening the parallel market is, denoting $z = (1+\alpha x)^{-1}$,

$$\frac{\widehat{W}_1(1) + \widehat{W}_2(1)}{2} = \frac{1}{4}\left[\left(1 + 2y + \frac{1}{2}\left(1 - \left(\frac{A}{1+A}\right)^2\right)\right)(\sigma_1^2 + \sigma_2^2) + \left(\frac{2}{(N-1)(2N-1)} + \frac{\alpha}{2}\left(\frac{A}{1+A}\right)^2\right)\sigma_2^2 z\right]$$

This is to be compared with the expected payoff of shutting the parallel market:

$$\widehat{W}_2(0) = \frac{\theta_{cc}}{4} = \frac{1}{4}\left[\frac{1}{2}\left(1 + \frac{1}{N-1}\right)(\sigma_1^2 + \sigma_2^2) + \frac{1}{2N}\sigma_2^2 z\right]$$

Now compare the coefficients in $\sigma_1^2 + \sigma_2^2$ in each expression: as $1 - \left(\frac{A}{1+A}\right)^2$, one has

$$1 + 2y + 1/2\left(1 - \left(\frac{A}{1+A}\right)^2\right) > 1$$
$$\leq \frac{1}{2}\left(1 + \frac{1}{N-1}\right)$$

As far as the coefficients in $\sigma_2^2 z$ are concerned, the coefficient is positive in the first expression and negative in the second expression. Thus

$$\frac{\widehat{W}_1(1) + \widehat{W}_2(1)}{2} > \widehat{W}_2(0)$$

Expected utility effect of opening the parallel market. The statement that when $\gamma\sigma_2^2$ is not too high, dealers prefer to open the parallel market comes from the previous point: as the expected payoff of opening the parallel market is higher than that of leaving it closed, by continuity, the inequality is preserved as $\gamma\sigma_2^2$ is not too high.

Customers' welfare. Given that $p_d^* < p_0^*$ when customers sell (and conversely if they buy), as shown by proposition 15, customers are worse off if they are forced to trade in the parallel market.

Bibliography

Adrian, T., N. Boyarchenko, and O. Shachar (2017). Dealer balance sheets and bond liquidity provision. *Journal of Monetary Economics 89*, 92 – 109.

Adrian, T., E. Etula, and T. Muir (2014). Financial intermediaries and the cross-section of asset returns. *The Journal of Finance 69*(6), 2557–2596.

Adrian, T. and H. S. Shin (2009). Money, liquidity, and monetary policy. *American Economic Review Papers and Proceedings 99*(2), 600 – 605.

Adrian, T. and H. S. Shin (2010). Liquidity and leverage. *Journal of Financial Intermediation 19*(3), 418 – 437.

Allaz, B. and J.-L. Vila (1993). Cournot competition, forward markets and efficiency. *Journal of Economic Theory 59*(1), 1 – 16.

An, Y. (2019). Competing with Inventory in Dealership Markets. Working paper.

Antill, S. and D. Duffie (2018). Augmenting Markets with Mechanisms. Working paper.

Atkeson, A. G., A. L. Eisfeldt, and P.-O. Weill (2015). Entry and Exit in OTC Derivatives Markets. *Econometrica 83*(6), 2231–2292.

Bao, J., M. O'Hara, and X. A. Zhou (2018). The Volcker Rule and corporate bond market making in times of stress. *Journal of Financial Economics 130*(1), 95 – 113.

Bao, J., J. Pan, and J. Wang (2011). The Illiquidity of Corporate Bonds. *The Journal of Finance 66*(3), 911–946.

Bessembinder, H., S. Jacobsen, W. Maxwell, and K. Venkataraman (2018). Capital Commitment and Illiquidity in Corporate Bonds. *The Journal of Finance 73*(4), 1615–1661.

Biais, B., F. Heider, and M. Hoerova (2016). Risk-Sharing or Risk-Taking? Counterparty Risk, Incentives, and Margins. *The Journal of Finance 71*(4), 1669–1698.

Biais, B., F. Heider, and M. Hoerova (2019). Variation Margins, Fire Sales, and Information-Constrained Optimum. Working paper.

Biais, B., J. Hombert, and P.-O. Weill (2019). Incentive Constrained Risk sharing, Segmentation, and Asset Pricing. Working paper.

Brogaard, J., T. Hendershott, and R. Riordan (2014, 06). High-Frequency Trading and Price Discovery. *The Review of Financial Studies 27*(8), 2267–2306.

Brogaard, J., T. Hendershott, and R. Riordan (2019). Price Discovery without Trading: Evidence from Limit Orders. *The Journal of Finance 74*(4), 1621–1658.

Brunnermeier, M. K. and L. H. Pedersen (2008, 11). Market Liquidity and Funding Liquidity. *The Review of Financial Studies 22*(6), 2201–2238.

Cameron, A. C., J. B. Gelbach, and D. L. Miller (2011). Robust inference with multiway clustering. *Journal of Business & Economic Statistics 29*(2), 238–249.

Choi, J. and Y. Huh (2019). "customer liquidity provision: Implications for corporate bond transaction costs". Working Paper.

Collin-Dufresne, P., B. Junge, and A. Trolle (2018). Market Structure and Transaction Costs of Index CDSs. Working Paper.

Copeland, A., A. Martin, and M. Walker (2010). The Tri-Party Repo Market before the 2010 Reforms. Federal Reserve Bank of New York Staff Reports no. 477.

Di Maggio, M., A. Kermani, and Z. Song (2017). The value of trading relations in turbulent times. *Journal of Financial Economics 124*(2), 266 – 284.

Dick-Nielsen, J. (2014). How to clean Enhanced TRACE data. Unpublished manuscript.

Dick-Nielsen, J. and M. Rossi (2018). The Cost of Immediacy for Corporate Bonds. *The Review of Financial Studies 32*(1), 1–41.

Du, S. and H. Zhu (2017). Bilateral Trading in Divisible Double Auctions. *Journal of Economic Theory 167*, 285 – 311.

Du, W., A. Tepper, and A. Verdelhan (2018). Deviations from Covered Interest Rate Parity. *The Journal of Finance 73*(3), 915–957.

Duffie, D. (2012). Market making under the proposed volcker rule. Rock Center for Corporate Governance at Stanford University Working Paper.

Duffie, D. (2018). Post-crisis bank regulations and financial market liquidity. Thirteenth Paolo Baffi Lecture on Money and Finance, Banca d'Italia, Eurosystem.

Duffie, D., N. Garleanu, and L. H. Pedersen (2005). Over-the-Counter Markets. *Econometrica 73*(6), 1815–1847.

Duffie, D. and H. Zhu (2017). Size Discovery. *The Review of Financial Studies 30*(4), 1095–1150.

Dugast, J., S. Üslü, and P.-O. Weill (2019). A Theory of Participation in OTC and Centralized Markets. Working paper.

Evans, M. D. D. and R. K. Lyons (2002). Order flow and exchange rate dynamics. *Journal of Political Economy 110*(1), 170–180.

Foucault, T., M. Pagano, and A. Roell (2014). *Market Liquidity. Theory, Evidence and Policy.* Oxford University Press.

Friewald, N. and F. Nagler (2019). Over-the-counter market frictions and yield spread changes. *The Journal of Finance 74*(6), 3217–3257.

Gilchrist, S. and B. Mojon (2017, 04). Credit Risk in the Euro Area. *The Economic Journal 128*(608), 118–158.

Gilchrist, S. and E. Zakrajsek (2012). Credit spreads and business cycle fluctuations. *The American Economic Review 102*(4), 1692–1720.

Glosten, L. R. and L. E. Harris (1988). Estimating the components of the bid/ask spread. *Journal of Financial Economics 21*(1), 123 – 142.

Glosten, L. R. and P. R. Milgrom (1985). Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics 14*(1), 71 – 100.

Goldstein, M. A. and E. S. Hotchkiss (2020). Providing liquidity in an illiquid market: Dealer behavior in us corporate bonds. *Journal of Financial Economics 135*(1), 16 – 40.

Gromb, D. and D. Vayanos (2002). Equilibrium and welfare in markets with financially constrained arbitrageurs. *Journal of Financial Economics 66*(23), 361 – 407.

Gromb, D. and D. Vayanos (2010). A model of financial market liquidity based on intermediary capital. *Journal of the European Economic Association 8*(2-3), 456–466.

Gromb, D. and D. Vayanos (2018). The dynamics of financially constrained arbitrage. *The Journal of Finance 73*(4), 1713–1750.

Grossman, S. J. and M. H. Miller (1988). Liquidity and market structure. *The Journal of Finance 43*(3), 617–633.

Gurkaynak, R. S., B. Sack, and J. H. Wright (2007). The U.S. Treasury yield curve: 1961 to the present. *Journal of Monetary Economics 54*(8), 2291 – 2304.

Hart, O. D. (1975). On the optimality of equilibrium when the market structure is incomplete. *Journal of Economic Theory 11*(3), 418 – 443.

Hasbrouck, J. (1991). Measuring the Information Content of Stock Trades. *The Journal of Finance 46*(1), 179–207.

He, Z., B. Kelly, and A. Manela (2017). Intermediary asset pricing: New evidence from many asset classes. *Journal of Financial Economics 126*(1), 1 – 35.

Hendershott, T. and A. Madhavan (2015). Click or Call? Auction versus Search in the Over-the-Counter Market. *The Journal of Finance 70*(1), 419–447.

Ho, T. and H. R. Stoll (1981). Optimal dealer pricing under transactions and return uncertainty. *Journal of Financial Economics 9*(1), 47 – 73.

Huang, R. D. and H. R. Stoll (2015, 06). The Components of the Bid-Ask Spread: A General Approach. *The Review of Financial Studies 10*(4), 995–1034.

Hull, J. C. (2003). *Options, Futures, and other Derivatives.* Prentice Hall Finance Series.

Klemperer, P. D. and M. A. Meyer (1989). Supply Function Equilibria in Oligopoly under Uncertainty. *Econometrica 57*(6), 1243–1277.

Kyle, A. S. (1985). Continuous Auctions and Insider Trading. *Econometrica 53*(6), 1315–1335.

Kyle, A. S. (1989). Informed Speculation with Imperfect Competition. *The Review of Economic Studies 56*(3), 317–355.

Li, D. and N. Schuerhoff (2019). Dealer Networks. *The Journal of Finance 74*(1), 91–144.

Lopez-Salido, D., J. C. Stein, and E. Zakrajsek (2017, 05). Credit-Market Sentiment and the Business Cycle. *The Quarterly Journal of Economics 132*(3), 1373–1426.

Lorenzoni, G. (2008, 07). Inefficient Credit Booms. *The Review of Economic Studies 75*(3), 809–833.

Lyons, R. K. (1995). Tests of microstructural hypotheses in the foreign exchange market. *Journal of Financial Economics 39*(2), 321 – 351.

Madhavan, A. and S. Smidt (1993). An analysis of changes in specialist inventories and quotations. *The Journal of Finance 48*(5), 1595–1628.

Malamud, S. and M. Rostek (2017, November). Decentralized Exchange. *American Economic Review 107*(11), 3320–62.

McDonald, R. and A. Paulson (2015, May). AIG in Hindsight. *Journal of Economic Perspectives 29*(2), 81–106.

Oehmke, M. and A. Zawadowski (2015, 08). Synthetic or Real? The Equilibrium Effects of Credit Default Swaps on Bond Markets. *The Review of Financial Studies 28*(12), 3303–3337.

Oehmke, M. and A. Zawadowski (2016, 08). The Anatomy of the CDS Market. *The Review of Financial Studies 30*(1), 80–119.

Philippon, T. (2009). The Bond Market's q. *The Quarterly Journal of Economics 124*(3), 1011–1056.

Rapp, A. C. (2016). Middlemen matter: Corporate Bond Market Liquidity and Dealer Inventory Funding. Working paper.

Rostek, M. and M. Weretka (2015). Dynamic Thin Markets. *The Review of Financial Studies 28*(10), 2946–2992.

Saar, G., S. Jian, R. Yang, and H. Zhu (2019). From Market Making to Matchmaking: Does Bank Regulation Harm Market Liquidity? Working Paper.

Schultz, P. (2017). Inventory management by corporate bond dealers. Working Paper.

Siriwardane, E. N. (2019). Limited Investment Capital and Credit Spreads. *The Journal of Finance 74*(5), 2303–2347.

Stein, J. C. (2012). Monetary Policy as Financial Stability Regulation. *The Quarterly Journal of Economics 127*(1), 57–95.

Stoll, H. R. (1978). The Supply of Dealer Services in Securities Markets. *The Journal of Finance 33*(4), 1133–1151.

Stulz, R. M. (2004, September). Should We Fear Derivatives? *Journal of Economic Perspectives 18*(3), 173–192.

Vayanos, D. (1999). Strategic Trading and Welfare in a Dynamic Market. *The Review of Economic Studies 66*(2), 219–254.

Volcker, P. (2010). How to reform our financial system. *The New York Times Jan. 30.*