ABSTRACT

Title of dissertation: Essays on Market Microstructure and Asset Pricing
Wen Chen, Doctor of Philosophy, 2017

Dissertation directed by: Professor Albert S. Kyle
Department of Finance

This dissertation contains two essays exploring the asset pricing implications of asymmetric information, hedging and market making.

Chapter 1 studies position limits on strategic speculators in commodity futures. In this chapter I develop an equilibrium model with both spot and futures markets to evaluate the effects of speculative position limits proposed by commodity regulators. One of the main implications of this model is that the imperfectly competitive speculators can benefit from the limits at the expense of unconstrained market participants. Therefore, it is important to take into account the market competitiveness when setting position limits. I also find that the limits always reduce market liquidity and thereby increase the cost of hedging. Thus, position limits would benefit market makers but hurt hedgers. Moreover, the loss of liquidity due to the limits has a spillover effect on the spot market as futures prices reveal less information which makes all spot market participants worse off. Contrary to regulators' beliefs, the model suggests that an aggregate position limit may reduce speculators' competition and market liquidity even when the limit does not bind. The model provides an alternative explanation of magnet effect of position limits, which is imperfect competitive speculators tend to exert their market power to make the limits bind.

Chapter 2 (joint with Yajun Wang) studies dynamic of market making and asset pricing. In this chapter, we develop a dynamic model of market making with asymmetric information where imperfectly competitive market makers match offsetting trades and carry zero inventory over time. Our model captures key features of market making in many financial markets: market makers optimally facilitate trading in both bid and ask markets by adjusting bid and ask prices and they hold close-to-zero inventories at the end of the day. We solve for equilibrium bid/ask prices and market depths in closed-form, and examine how informed traders dynamically hedge liquidity shocks and reveal private information. We further study the dynamics of bid-ask spread and trading volume to understand how these may interact with each other in shaping asset prices and market liquidity.

ESSAYS ON MARKET MICROSTRUCTURE AND ASSET PRICING

by

Wen Chen

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2017

Advisory Committee:
Professor Albert S. Kyle, Chair
Professor Mark Loewenstein,
Professor Yajun Wang
Professor Julien Cujean
Professor Daniel R. Vincent, Dean's Representative

Acknowledgments

I am grateful for my committee members: Pete Kyle, Mark Loewenstein, Yajun Wang, Julien Cujean, and Daniel Vincent. This dissertation would not have been possible without their constant guidance and support. I would like to express my deepest and most sincere gratitude to my advisor Pete. He is such a great mentor that I always feel his unlimited energy and generosity. His particular combination of brilliance and devotion modeled for me how a premiere scholar should be. I am thankful for his considerable patience with me, and his wisdom and wide knowledge that guides me. I feel so lucky and honored to be one of his students. I am deeply grateful to Mark. He is always smiling and takes a lot of trouble to make me feel at ease. He has always encouraged me to take risks in my attempts at research. He respects my ideas however naive they are, and provides insightful discussions. His feedback is a steady resource at every stage of my research. I am very thankful to Yajun. She is always there when I have questions and helps me go through all the difficulties. Her enthusiasm and collaborative spirit help sustain my energy and push me to new levels of rigor and intellectual sophistication. She has been a dear friend and a continuous source of support. I want to thank Julien for chatting with me about research for the past a few years. He not only taught me many things about research but also taught me how make better presentations and how to write my paper. Special thanks goes to Professor Daniel Vincent for willing to participating in my dissertation committee and providing insightful suggestions.

I also thank all the other Ph.D. students in the program especially Xiaoyuan Hu, Danmo Lin, Jinming Xue and Don Bowen for their company, kindness, friendship and support.

I am deeply grateful to my husband Bo Hu and my in-laws for their devotion and support. I would like to thank my dear mother and farther for everything. I also want to thank my daughter YuanBao for bringing joys to my life. Their loves have given me strengthen to accomplish my second Ph.D. degree.

Contents

Chapter 1

Position Limits of Strategic Speculators in Commodity Futures

1.1 Introduction

Speculation in commodity futures has gained tremendous popularity over the last decade. Along with the entry of financial speculators into commodity futures, a sharp increase in price fluctuation raised concerns of policymakers. On November 5, 2013, the Commodity Futures Trading Commission (CFTC) approved proposed rules that would impose *speculative position limits* or *aggregate position limits* on physical commodity futures contracts under the mandate of the Dodd-Frank Act (see 16, 17). The Commission believes that the rules can reduce the cost of hedging by reducing price fluctuations and thereby benefit hedgers,[1] and can promote market competitiveness by preventing speculators from amassing market

[1] Testimony of Paul N. Cicio, the President of the Industrial Energy Consumers of America (IECA), in front of Senate Permanent Subcommittee on Investigations on November 3, 2011.

power.[2] However, whether position limits would reduce price volatility or not is debatable.[3] And it is unclear whether the rules will harm market liquidity and impede information discovery.[4] Moreover, several important questions remain unanswered: how would the limits affect the competition among speculators? What are the externalities on the underlying spot markets?

To address these questions, I develop a model with both futures and spot markets. To justify why regulators impose position limits on speculators, the model assumes that speculators have both speculating and hedging motives for trade and are imperfectly competitive. Speculators' trading can increase price fluctuations and their market power can increase the cost of hedging. However, this paper finds that position limits on speculators do not always reduce price volatility. More importantly, position limits always reduce market liquidity and thereby increase the cost of hedging, resulting in a welfare loss of hedgers. The illiquidity due to position limits has a spillover effect on the spot market because of reduced information revelation. The impacts of these rules on speculators depend on the degree of competition among speculators. Less competitive speculators can even benefit more from position limits. This is because position limits may not only protect speculators from competing with each

[2] Some legislation supporters (e.g., IECA) even suggest CFTC to set an aggregate position limit on all commodity index funds and commodity-related exchange traded funds. They argue that these passive funds follow some similar trading strategy and have the ability to amass market power.

[3] See 10, 26, 7, 54

[4] On September 28, 2012, a federal judge ruled in favor of the International Swaps and Derivatives Association (ISDA) and the Securities Industry and Financial Markets Association (SIFMA), that the position limits rule should not be imposed because the CFTC did not first take steps to determine whether such limits were "reasonable and appropriate".

other, but also mitigate the impact of adverse selection on speculators. Therefore, contrary to regulators' beliefs, position limits can lead imperfectly competitive speculators to take larger positions at the limits, and an aggregate position limit can encourage competitive speculators to amass their market power so that the market liquidity will be reduced even when the aggregate limit does not bind.

The model consists of both futures and spot markets in a two-period setup. In the futures market, there are three types of market participants: hedgers, market makers, and speculators. Take the commodity grain as an example. Hedgers are grain farmers or distributors. They harvest the grain at fall and sell the grain in the spot market after harvest. We can think of the first period as a pre-harvest contract of futures, and the second period as a post-harvest contract of grains. Before harvest, farmers trade in futures market to hedge their inherent risks which include both price risk and *quantity risk* (i.e., the uncertainty of commodity supply). This leads to seasonality in hedging demands of futures.[5] Farmers are *bona fide* hedgers as categorized by the Commission, and hence are exempted from position limits. The Commission argues that the rules will not harm market liquidity, because there are unconstrained traders who provide liquidity to hedgers when limits bind for speculators. The market makers in my model[6] denote such unconstrained traders who can be floor traders or grain consumers. Speculators in my model differ from hedgers and market makers in three important aspects. First, they correspond to those institution investors (ETFs, Hedge funds)

[5] Similar seasonality in hedging demands exists in many non-agricultural markets, e.g., natural gas.

[6] They are not designated market makers but correspond to floor traders and managed money traders. This paper distinguishes them from financial speculators, because from the view of regulators they always provide liquidity in futures and should not be subject to position limits.

who have no innate position in spot markets. Second, they are usually oligopolists as in real markets. Third, they are subject to position limits.

This paper shows that position limits, contrary to regulators' goals, can increase commodity price volatility. This is because speculators have dual trading motives: speculation and hedging. On one hand, speculators' trading brings their own noise into spot prices. On the other hand, their trading reduces spot price fluctuations due to the trades of hedgers. When hedgers face high quantity risk (large noise from hedgers) and low information asymmetry (small noise from speculators), speculators' trading tends to stabilize commodity price, i.e., provide liquidity.[7] In this situation, imposing position limits on speculators increases spot price volatility. Moreover, my model shows that no matter speculators provide or demand liquidity the rules always reduce *market liquidity*[8] in two ways. First, when limits bind the futures price becomes less informative, so the perceived uncertainty of spot price by hedgers and market makers increases. Second, when limits bind speculators' demands become price-inelastic and they lose the ability to absorb exogenous liquidity shocks in futures. Therefore, limiting speculators' trading positions always reduces market liquidity even though market makers are exempted from the position limits. In addition, the illiquidity in futures has a spillover effect on the underlying spot market through information channel.

---

[7] Speculators are said to stabilize prices if they buy when prices are low and sell when prices are high (27).

[8] Market liquidity is proportional to the dollar volume divided by the sum of perceived riskiness and market power. Dollar volume is measured by the reciprocal of risk aversion. Perceived riskiness is measured by the conditional variance of payoff. The idea that market liquidity is related to the dollar volume per unit of conditional payoff variance has the similar intuition with 38. Traders believe that transaction costs are high in markets with low dollar volume and high perceived riskiness.

Because the position limits increase the perceived uncertainty of less informed spot market participants, who use futures prices to infer information. Furthermore, the cost of hedging by hedgers can be decomposed into two components: market illiquidity and average hedging demands. Since position limits significantly increase market illiquidity and yet have only a marginal effect on the average hedging demands, the cost of hedging by hedgers always increases when position limits bind on speculators.

I undertake a welfare analysis of the proposed rules by the Commission.[9] Position limits exhibit quite different impacts on market participants, depending on speculators' competitiveness and other market characteristics such as the degree of information asymmetry and the level of quantity risk. Since the cost of hedging increases as discussed above, there is always a welfare loss for hedgers in the presence of position limits. Position limits have two effects on speculators (56): increasing risk premium (price effect) and decreasing trading quantities (quantity effect). As the number of speculators increases (i.e., market becomes more competitive and position limits become less likely to bind), the price effect dominates and makes speculators better off. The new insight revealed in my model is that any market power exerted by speculators will dampen the quantity effect and the increased information asymmetry due to position limits will exacerbate the price effect. Therefore, the presence of imperfect competition and information asymmetry tends to further improve speculators' welfare, relative to the case of perfect competition and symmetric information. Position limits also have two effects on market makers' trading quantity. First, the increase of risk premium due to position limits induces market makers to trade more in futures, since on

[9] "The Commission fails to undertake a comprehensive cost-benefit analysis by the statue" from Comment letter on position limits for derivatives from CME group on March 28, 2011.

average market makers trade in the same direction as speculators do. Second, the increased perceived uncertainty due to position limits makes market makers reluctant to trade in futures. Overall, when position limits are stringent, the first effect dominates and makes market makers better off. Two direct implications of my model are that it is inappropriate to apply a uniform position limit (which has been proposed by the Commission[10]) on markets with different competitiveness among speculators and it is important to incorporate the changes of market conditions when setting position limits.

In addition to risk sharing and liquidity provision, modern commodity futures markets also serve as a barometer for commodity spot markets and even equity markets (33). When position limits on speculators bind, information revealed by futures price is inevitably reduced. Previous literature (52) argues that information revealed by futures price can harm hedgers, because in spot market hedgers have information advantage over market makers. This argument implies that informational frictions in futures markets would benefit hedgers in spot markets. Nonetheless, this argument becomes questionable when there is quantity risk. For informed hedgers who are subject to quantity risk (as modeled here), the loss of information advantage due to public information (futures price) is always outweighed by the benefits of increased risk sharing due to public information. Consequently, position limits, which hamper information revealing of futures price and reduce the scope of risk sharing, generate negative externalities on both informed and less informed spot market participants. The more stringent the position limits are, the more welfare loss the hedgers will suffer in spot market, especially when quantity risk is high.

[10] The Commission proposed to set a uniform limit at 25 percent of the estimated supply for all 28 commodity futures and a time frame for review or change the limit at every two years.

The Commission believes that an aggregate position limit would promote market competitiveness. Contrary to this belief, I find that an aggregate position limit can disincentivize speculators to compete on their common information. There exists a trade-off for speculators to be an informational monopolist in the presence of risk averse market makers.[11] On one hand, acting as a monopolist allows speculators to extract more rent from their information advantage. On the other hand, being a monopolist incurs a higher adverse selection cost. There are cases where speculators are better off to stay competitive in the absence of an aggregate position limit, but once subject to the limit, their welfare from being an informed monopolist can be greater than that from being fully competitive. In other words, an aggregate position limit may hurt competitive speculators but benefit less competitive speculators. This is because position limits mitigate the impact of adverse selection on speculators. This result uncovers a potential counterproductive effect of implementing an aggregate position limit: it may reduce the competition of speculators, and thus reduce the market liquidity and the scope of risk sharing even when the limit does not bind.

In sum, this paper sheds light on this long-term debate on the CFTC proposed rules of speculative position limits in commodity futures. The paper studies how position limits interact with other market frictions such as information asymmetry and imperfect competition. It shows that position limits increase information asymmetry and cannot curb the

[11] Since there is a large number of uninformed market makers in this model, the competition among speculators is the competition on their common private information and liquidity shock. This trade-off is adjusted by market makers' risk aversion A_M. If $A_M \to \infty$, market makers trade like noise traders, then speculators are always better off if acting as a monopolist because there is no adverse selection from market makers. If market makers are risk neutral, i.e., $A_M = 0$, then speculators are always better off if competing on their information.

imperfect competition. The interaction of these market frictions leads to lower market liquidity in both futures and spot markets and a higher cost of hedging. Contrary to regulators' objectives, position limits can hurt hedgers but benefit speculators.

1.1.1 Contribution and Related Literature

To my best knowledge, this paper is the first one that examines how position limit on imperfectly competitive speculators affects both spot and futures markets in the presence of information asymmetry. In particular, I derive the existence condition for an imperfectly competitive equilibrium and show that as the number of speculators decreases (i.e., the market becomes less competitive) the equilibrium exists only when either speculators or hedgers have sufficiently strong trading motives. This is consistent with the stylized fact that it is rare to see a commodity futures market which has only a few financial speculators. Moreover, existing literature (55) focuses on the imperfectly competitive equilibrium in a one-period setup, and shows that market power always leads to a less liquid market. However, this model shows that in a two-period setup market power may leads to more aggressive trading, which increases market liquidity. In general, my model provides a framework that can be extended to study other financial markets with large traders under certain position restrictions.[12]

This paper is related to several bodies of literature. A widely accepted view of the function of commodity futures is associated most prominently with the names of Keynes

[12] For example, the European Union has banned naked credit default swap (CDS) positions on sovereign debt. Another example could be that exchange-traded funds only allow large institutional organizations to undertake the responsibility of obtaining the underlying assets.

(36) and Hicks (31). Their standard conception interprets speculation as a process for the transfer of price risks and the short positions in futures market as hedging pressure. Holbrook Working (58), in contrast, has argued that hedging is a form of arbitrage involving the purchase or sale of futures in the expectation of a favorable price change. This paper shows that the average endowments determine the average hedging demands and supplies. With a positive commodity supply, hedgers take short positions and speculators take long positions *on average*, which is consistent with empirical findings (53) and also in agreement with the hedging pressure theory of Keynes-Hicks. This model also shows that hedgers adjust their futures positions *on the margin* in response to futures price changes even when quantity risk (i.e., uncertainty of commodity supply) is low. This is consistent with Holbrook Working's interpretation of hedging and a series of recent empirical evidences (14, and 22). This paper reconciles these two seemingly conflicting theories of hedging in futures by showing that hedgers hedge on average but can speculate on the margin. In the presence of high quantity risk, hedgers under-hedge on average and use the information revealed by futures price more aggressively than speculators. Thus, this model presents a more nuanced view of speculation and hedging in commodity futures than existing models. It suggests that any regulatory measure based on the categorical identities of hedgers and speculators rather than the specific activities they engage in is ambiguous and questionable.

This paper is also related to the growing literature about the commodity financialization, both theoretical (3, 54, 6, 20) and empirical (4, 37, 9, 48). This paper is closely related to 51, 43, 21. The model in this paper differs from others in four key assumptions which are necessary ingredients to study position limits. First, this paper models large speculators who can exert market power and trade strategically, whereas other models assume competitive traders. Second, there are naturally three types of market participants (hedgers, speculators, and market makers) in futures market as assumed in this paper, while 43 and 21 consider only two types of participants (hedgers and speculators) in futures. By market clearing in their models, to constrain the trading of speculators is equivalent to constrain the trading of hedgers. Although 51 consider three types of traders in futures, they do not fully model the rational trading of speculators. Third, this paper assumes hedgers face quantity risk, i.e., the uncertainty of commodity supply, in addition to price risk. It is the interaction between price risk and quantity risk that risk averse market participants must respond to in their hedging-speculative commitments (32). The other models do not consider quantity risk but take commodity supply as a choice variable by commercial hedgers.[13] The commodity production, however, only affects the average hedging pressure instead of the market liquidity. Market liquidity is largely affected by the uncertainty of commodity supply (quantity risk). Therefore, quantity risk plays an important role in understanding how position limits affect the market liquidity. Lastly, this paper assumes that the futures and spot markets open sequentially and hedgers are non-myopic. Both models in 43 and 21 have one period, entangling the substitution effect of futures market to spot market with the informational effect of futures prices. Thus, it requires at least a two-period setup to identify the information externalities of position limits on spot market. Although 51 use a two-period setup, they assume commercial hedgers are myopic for tractability. When facing quantity risk,

[13] Commodity production can be incorporated into this model by endogenizing the mean of commodity supply or letting the supply mean correlate with commodity liquidation value. However, it is the quantity risk rather than commodity production that is essential to study position limits. Position limits have a significant impact on market liquidity but only a marginal effect on the average hedging pressure.

myopic hedgers would not engage in anticipatory hedging which accounts for a large portion of hedging demands in commodity futures.

This paper is also closely related to the literature studying the effects of position constraints on market participants (34, 1, 2, 56). However, the existing literature does not consider the impacts of position limits on imperfectly competitive traders. Moreover, the existing literature only studies how position constraints affect one market. This paper instead examines the effects of position limits on both the derivative market and the underlying spot market. Although 5 analyzes the underlying and derivative markets jointly when evaluating the impacts of trading restriction in one market, they focus on the role of derivative market as an substitution of the underlying market. In contrast, this paper emphasizes the informational role of derivative market on underlying market and highlights how trading restriction interrupts the information learning of less informed traders and reduces overall market liquidity. More importantly, the increased illiquidty has a spillover effect from a trading-constrained market to an unconstrained market as demonstrated in my model, similar to the notion of illiquidty contagion via the information channel (11, 19).

More generally, this paper shows that position limits on informed traders can be substituted by price limits (8, 40), and this substitution effect comes from the homogeneity of speculators. This model can be extended to compare the effectiveness of trading restrictions commonly used by regulators (such as price limits, position limits, and margin requirements).

1.2 The Model

This section presents the model setup and provides discussions on the main assumptions of the model. The model consists of both futures and spot markets in a two-period setup. Take the commodity grain as an example. Grain farmers harvest the grain at fall and sell the grain in the spot market after harvest. We can think of the first period as a pre-harvest contract of futures, and the second period as a post-harvest contract of grains. Before harvest, farmers trade in futures market to hedge their inherent risks.

1.2.1 Model Setup

Assets and Markets The economy has three assets: one is an underlying asset, which is a commodity, one is a futures contract, which is the derivative of the underlying asset, and one is a risk-free asset with zero rate of return (called cash), which serves as the numeraire. I assume that the commodity and its derivative can not be consumed. There are two markets: a spot market where the commodity is traded and a futures market where the derivative is traded. The economy has two periods and three dates, denoted as $t = 0, 1, 2$. To focus on the informational role of futures markets, I assume a centralized futures market first opens at $t = 0$ with a futures price F for futures contracts which expire at $t = 1$. The spot market opens at $t = 1$ with a spot price P, and all the futures contracts are cash-settled or unwound at P,[14] i.e., basis is assumed to be zero. The total supply of the futures contracts is zero and the total supply of the underlying asset is a random variable, denoted by X, which will

[14] I assume that no market participant takes the physical delivery. In reality, only less than 2% of commodity futures are settled with physical delivery.

be further discussed below. The liquidation value for each unit of the underlying asset is realized at $V + \eta$, where

$$V \sim \mathcal{N}(\bar{V}, \tau_V^{-1}), \tag{1.1}$$

$$\eta \sim \mathcal{N}(0, \tau_\eta^{-1}). \tag{1.2}$$

Here, V represents the fundamental value of this commodity, and η denotes the residual uncertainty that no one in the market knows. V and η are independent with each other.

Market Participants There are three types of market participants in the economy: commercial hedgers (**H** for short), market makers (**M** for short), and speculators (**S** for short). All the market participants are risk averse, as characterized by exponential (CARA) utility maximizers.

Commercial Hedgers (**H**) There are N_H identical commercial hedgers. Their market power is negligible by assuming $N_H >> 1$ so that they are effectively price takers. Hedgers trade in both futures and spot markets. In reality, they are usually commodity producers, processors, and/or inventory holders. At $t = 1$, they receive an aggregate endowment of X units of the commodity, i.e., each hedger receives X/N_H units, where

$$X \sim \mathcal{N}(\bar{X}, \tau_X^{-1}), \tag{1.3}$$

which is independent of all other random variables, and only known by the hedgers at $t = 1$. The distribution of X, however, is a common knowledge that all the market participants know at the beginning $t = 0$. As the net supply of the commodity, X can also be understood as the asynchronization of demand and supply of the commodity (23). The uncertainty of X represents the quantity risk (32) that interacts with the price risk.

Hedgers trade in both futures and spot markets. At $t = 0$, they know the priors of all the random variables, and observe the futures price F. Based on such information, each hedger submits his futures demand schedule y_H. At $t = 1$, they observe the spot price P and the fundamental value V, and receive the endowment shock X. Each hedger unwinds his futures position at P and submits his demand schedule, x_H, of the commodity in spot market. At $t = 2$, each hedger realizes a terminal wealth by liquidating his inventory of the commodity at $V + \eta$:

$$W^H = (P - F)y_H + (V + \eta - P)x_H + (V + \eta)X/N_H. \tag{1.4}$$

I assume that each hedger's risk aversion is $N_H A_H$ such that the total risk-bearing capacity of hedgers is constant and equal to $\frac{1}{A_H}$. This assumption ensures that the population only affects the degree of competition but not the aggregate risk tolerance. Then the optimal trading strategy for each hedger is determined by sequentially solving

$$J_1^H = \max_{x_H} \mathrm{E}_1^H[-e^{-N_H A_H W^H}], \tag{1.5}$$

$$J_0^H = \max_{y_H} \mathrm{E}_0^H[J_1^H(y_H)], \tag{1.6}$$

where $\mathrm{E}_t^H[\cdot] = \mathrm{E}[\cdot \mid \mathcal{F}_t^H]$. Hedgers' information sets are $\mathcal{F}_0^H = \{F\}$ and $\mathcal{F}_1^H = \{F, P, V, X\}$.

Market Makers (**M**) There are N_M identical market makers. Their market power is negligible by assuming $N_M >> 1$ so that they are also price takers. They have no private information or endowment shocks in neither market. In real world, they can be financial institutions (e.g. bank holding companies), or commodity users who have future purchases of the commodity. Market makers are rational uninformed traders who optimally trade to share risks.

Market makers can choose to trade in both futures and spot markets. At $t = 0$, they know the priors of all the random variables, and observe the futures price F. Each of them submits his futures demand schedule y_M. At $t = 1$, each market maker observes the spot price P at which he unwinds his futures position, and submits a demand schedule in spot market, x_M. At $t = 2$, every market maker realizes a terminal wealth by liquidating his inventory of commodity at the value $V + \eta$:

$$W^M = (P - F)y_M + (V + \eta - P)x_M. \tag{1.7}$$

Again, each market maker' risk aversion is assumed to be $N_M A_M$ so that their aggregate risk tolerance is constant and equal to $\frac{1}{A_M}$. The optimal trading strategy for every market maker is determined by sequentially solving

$$J_1^M = \max_{x_M} \mathrm{E}_1^M[-e^{-N_M A_M W^M}], \tag{1.8}$$

$$J_0^M = \max_{y_M} \mathrm{E}_0^M[-e^{J_1^M(y_M)}], \tag{1.9}$$

where $\mathrm{E}_t^M[\cdot] = \mathrm{E}[\cdot \mid \mathcal{F}_t^M]$. Their information sets are $\mathcal{F}_0^M = \{F\}$ and $\mathcal{F}_1^M = \{F, P\}$.

Speculators (**S**) There are N_S identical speculators. Since one of the goals of imposing speculative position limits on speculators is to curb their market power, I assume that they are imperfectly competitive and trade strategically. Their market power depends on their population N_S. If $N_S = 1$, there is only one speculator who is an informational monopolist. If $N_S >> 1$, then each speculator's market power becomes negligible and the whole economy is competitive. Speculators only trade in futures market. Different from hedgers and market makers, they do not hold any underlying commodity. In real world, most of them nowadays are institutional investors such as hedge funds and index traders.[15] As sophisticated portfolio

[15] In recent literature, they are often referred to as financial speculators. A large influx of these speculators

investors, they seek a risk exposure to the commodity's payoff to diversify their portfolios, and in general they have no ability or interest to hold the commodity. Empirical evidence (29, 30) shows that speculators have private information about the commodity fundamental value and they also trade in futures for hedging motives. For this reason, I assume that speculators at $t = 0$ receive a private information s about V: $s = V + \epsilon$, where

$$\epsilon \sim \mathcal{N}(0, \tau_\epsilon^{-1}), \tag{1.10}$$

is independent of all the other random variables. In addition, speculators have a total risk exposure z to the risk factor $V - \bar{V}$ where

$$z \sim \mathcal{N}(0, \tau_z^{-1}), \tag{1.11}$$

is independent of all the other random variables. So each speculator receives an endowment shock of $\frac{1}{N_S} z(V - \bar{V})$ and only speculators themselves know their risk exposure z at $t = 0$. As this endowment is correlated with V, it generates a hedging motive for speculators to trade in futures market.[16]

Unlike hedgers and market makers, each speculator is subject to a *speculative position limit* denoted by L. At $t = 0$, based on the private information s and z and public futures price F, each speculator submits a futures demand schedule y_S up to the limit L by taking into account his price impact. Then at $t = 1$, each speculator unwinds his futures position

into commodity futures since 2000 is referred to as commodity fictionalization.

[16] The dual-motive assumption is more realistic, though it is not the key assumption that drives the results. This prevents information from fully revealing in equilibrium. An alternative setup with exogenous liquidity traders will not alter the main results.

at the spot price P. At $t = 2$, each speculator realizes a terminal wealth:

$$W^S = (P - F)y_S + \frac{z}{N_S}(V - \bar{V}). \tag{1.12}$$

Again, each speculator's risk aversion is assumed to be $N_S A_S$ so that their aggregate risk tolerance is constant and equal to $\frac{1}{A_S}$. Then each speculator's optimal demand in futures market is determined by solving

$$J_0^S = \max_{y_S \leq L} \mathrm{E}_0^S[-e^{-N_S A_S W^S}], \tag{1.13}$$

where $\mathrm{E}_t^S[\cdot] = \mathrm{E}[\cdot \mid \mathcal{F}_t^S]$. Their information sets are $\mathcal{F}_0^S = \{F, s, z\}$ and $\mathcal{F}_1^S = \{F, P, s, z\}$.

Timeline I embed two sequentially opened markets into the model of 24 where speculators trade strategically as in 41. In each market, there is an auctioneer who collects the demand schedules and sets the publicly observable price to clear the market. Table 1.1 illustrates the timeline of the model.

Information Inference If the speculative position limit does not bind, the information revealing can be conjectured as follows. Since there is no other noise in the futures market, the futures price is conjectured to be a linear function of speculators' private information s and z, i.e., $F = \mathbb{L}[s, z]$. Since all the market participants are rational and non-myopic, both the current information in the spot market and the past information in the futures market will be incorporated into the spot price. Hence, the spot price is conjectured to a linear function of all the information available up to $t = 1$, i.e., $P = \mathbb{L}[s, z, V, X]$ (see Appendix 1.8.1). The linearity of pricing functions comes from the CARA-Normal setup.

If the speculative position limit binds, the information revealing in the futures market depends on speculators' trading strategy. If speculators continue submitting their demand

	$t = 0$	$t = 1$	$t = 2$
Market Participants: Population	Futures Price F	Spot Price P	Payoff $V + \eta$
Hedgers: N_H	submit y_H	observe V, X unwind y_H submit x_H	receive $X(V + \eta)$ W^H realized
Market Makers: N_M	submit y_M	unwind y_M submit x_M	W^M realized
Speculators: N_S	observe s, z submit $y_S \leq L$	unwind y_S	receive $z(V - \bar{V})$ W^S realized

Table 1.1: Timeline of the two sequentially opened markets. In each market, there is an auctioneer who collects the demand schedules and sets the publicly observable price to clear the market.

functions, i.e., a limit order, to the auctioneer, they still reveal a mixed information of s and z to the market. Speculators, however, may optimally submit a market order, i.e., a price-inelastic demand, at the position limit. In this case, the market will lose track of speculators' information. Under different market conditions, speculators react differently to the speculative position limit, as will be discussed in more detail later.

1.2.2 Discussions on Assumptions

For tractability, all the random variables in this model are assumed to be normally distributed. Under normality, however, the endowments $X(V+\eta)$ and $z(V-\bar{V})$ may take extremely large negative values which may drive the ex-ante expected utility to negative infinity (55). To ensure the ex-ante expected utility to be finite, this paper assumes that the variances satisfy the following regularity conditions:

$$\frac{A_H^2}{\tau_X}\left(\frac{1}{\tau_V}+\frac{1}{\tau_\eta}\right)<1, \tag{1.14}$$

$$\frac{A_S^2}{\tau_z}\frac{1}{\tau_V}<1. \tag{1.15}$$

All the futures contracts are cash settled or unwound in this model. In real world less than 2% futures contracts are physically delivered at the expiration date (51). Also, 42 shows that cash settlement is equivalent to physical delivery under the assumption that delivery mechanism does not affect the production market (13).

To focus on the informational role of futures market, this model assumes that the futures market opens before the spot market. This also allows one to disentangle the substitution effect of the futures market to the spot market. In contrast, 5 emphasizes the substitution role of the futures market by assuming both futures and spot markets open simultaneously. If both markets open at the same time, the futures price has a negative effect on the spot price due to the substitution effect.

The assumption of quantity risk, i.e., the uncertainty of X, is important. The presence of quantity risk has two effects on hedgers. First, higher quantity risk (lower values of τ_X) makes hedgers more likely to under-hedge on average since the commodity supply is more negatively correlated with the spot price. Second, higher quantity risk may make hedgers to use the information revealed by futures price more aggressively than speculators. When there is low information asymmetry between hedgers and speculators (i.e., high values of τ_z), the uncertainty of commodity supply makes hedgers' positions in futures change in the same direction as the change of futures prices.

Finally, whether exerting market power benefits or hurts speculators depends on market makers' risk aversion which adjusts the impact of adverse selection on speculators. If market makers are risk neutral, then speculators face the strongest adverse selection from market makers and thus they tend to stay competitive. If market makers are infinitely risk averse (like noise traders), then speculators face no adverse selection and have strong incentive to act as an informational monopolist.

1.3 Benchmark: No Speculative Position Limit

I first solve the benchmark model where there is no speculative position limit ($L \to \infty$). Each market participant's strategy is a demand schedule which is submitted to an auctioneer in either spot or futures market. The auctioneer then aggregates all the demand schedules submitted, determines a market clearing price, and allocates quantities to satisfy each market participant's demand. The equilibrium is a single price auction which, from the point of view of auctioneer, looks "Walrasian". The definition of an imperfectly competitive equilibrium follows the concept in 41. What makes the equilibrium concept imperfectly competitive is not the market-clearing rule itself, but rather the manner in which market participants exploit the rules in determining what demand schedules to submit. In this model, the market participants who need to take into account their price impacts are speculators. As

N_S increases, speculators become more and more competitive, and the equilibrium as defined below converges to the standard competitive equilibrium of rational expectations as in 24.

Definition 1 (Imperfectly Competitive Equilibrium)**.** An imperfectly competitive equilibrium (x_i, y_i, F, P) with a vector of speculators' strategies $Y_S = (y_{S,1}, ..., y_{S,N_S})^\top$ is such that

1. In the spot market, given the future price F, the demand schedule x_i as a function of the spot price P, where $i \in \{H, M\}$, solves market participant i's problem (1.5) and (1.8) respectively. At $t = 1$ the spot market clears at the equilibrium spot price P such that

$$N_H x_H + N_M x_M = 0, \tag{1.16}$$

   and the cash market is naturally cleared by Walras' law.

2. In the futures market, given the pricing rule of P from Part 1, the demand schedule y_i as a function of F, where $i \in \{H, M\}$, solves market participant i's problem (1.6) and (1.9) respectively. For all speculators, $n = 1, ..., N_S$, and for any alternative strategy Y'_S differing from Y_S only in the n-th component $y_{S,n}$, the strategy Y_S yields a utility level no less than Y'_S:

$$\mathrm{E}_0^S[U_S((P - F(Y_S))y_{S,n}(Y_S))] \geq \mathrm{E}_0^S[U_S((P - F(Y'_S))y_{S,n}(Y'_S))], \tag{1.17}$$

   where U_S denotes speculators' negative exponential utility as in Eq. (1.13). At $t = 0$ the futures market clears at the equilibrium futures price F such that

$$N_H y_H + N_M y_M + \sum_{n=1}^{N_S} y_{S,n} = 0, \tag{1.18}$$

   and the cash market is naturally cleared by Walras' law.

In equilibrium, both prices and demands can be characterized by two state variables s_F and s_P, which are defined below. They the information reveal by the futures price F and the spot price P respectively. The state variable for futures market is $\{s_F\}$, and the state variables for spot market are $\{s_F, s_P\}$.

Proposition 1. *Without a speculative position limit, in equilibrium the spot price P reveals a mixed signal s_P to the market, which is a linear combination of fundamental value V and endowment quantity X:*

$$s_P = V - \frac{A_H}{\tau_\eta}(X - \bar{X}) \sim \mathcal{N}(\bar{V}, \tau_V^{-1} + \tau_P^{-1}), \tag{1.19}$$

where $\tau_P^{-1} = \frac{A_H^2}{\tau_\eta^2}\tau_X^{-1} = \mathrm{Var}(s_P - V)$ is the variance of the difference between s_P and V.

The futures price F reveals a mixed signal s_F to the market, which is a linear combination of speculators' private information s and their total endowment quantity z:

$$s_F = s - \frac{A_S}{\tau_\epsilon}z \sim \mathcal{N}(\bar{V}, \tau_V^{-1} + \tau_F^{-1}), \tag{1.20}$$

where $\tau_F^{-1} = \tau_\epsilon^{-1} + \frac{A_S^2}{\tau_\epsilon^2}\tau_z^{-1} = \mathrm{Var}(s_F - V)$ is the variance of the difference between s_F and V.

Proof. See Appendix 1.8.1. □

Since in equilibrium F is informationally equivalent to s_F and P is informationally equivalent to s_P, we can rewrite market participants' information sets: $\mathcal{F}_1^H = \{s_F, V, X\}$, $\mathcal{F}_1^M = \{s_F, s_P\}$, $\mathcal{F}_0^{H,M} = \{s_F\}$ and $\mathcal{F}_0^S = \{s, z\}$. To emphasize the dependence of the market-clearing price on the trading strategies of speculators, we write $F = F(Y_S)$ and define the price impact of the n-th speculator as $\lambda_{S,n} = \frac{\partial F}{\partial y_{S,n}}$ based on the intuition of 39. Then we can define the total price impact of all speculators as

$$\lambda = \sum_{n=1}^{N_S} \lambda_{S,n} = \sum_{n=1}^{N_S} \frac{\partial F}{\partial y_{S,n}}. \tag{1.21}$$

The optimal demands of each hedger, market maker and speculator respectively in futures are given by

$$
\begin{aligned}
y_H &= \frac{d_H(s_F - \bar{V}) - F}{N_H A_H G_{11}^H} + h_H, \\
y_M &= \frac{d_M(s_F - \bar{V}) - F}{N_M A_M G_{11}^M} + h_M, \\
y_{S,n} &= \frac{d_S(s_F - \bar{V}) - F}{N_S A_S \mathrm{Var}_0^S(P) + \lambda_{S,n}} + h_S,
\end{aligned}
$$

where the parameters d_j and the hedging demand h_j for $j \in \{H, M, S\}$ can be found in Appendix 1.8.1 and the matrices G^H and G^M are computed in Appendix 1.8.2.

Here, the matrix elements $G_{11}^{H,M} > 0$ denote hedgers' and market makers' effective conditional variances of spot price, respectively. In other words, $G_{11}^{H,M}$ can be viewed as the perceived risks of hedgers and market makers, respectively. In the presence of quantity risk, it can be shown that $G_{11}^H > \mathrm{Var}_0^H(P)$ and $G_{11}^M < \mathrm{Var}_0^M(P)$. It is intuitive that hedgers' perceived risk is higher in the presence of quantity risk. As quantity risk decreases to zero, i.e., τ_X increases to infinity, G_{11}^H decreases to $\mathrm{Var}_0^H(P)$ and G_{11}^M increases to $\mathrm{Var}_0^M(P)$. Since both hedgers and market makers are homogeneous and perfectly competitive, we do not need to distinguish their identity in the demand schedules. In equilibrium, the identity of any speculator n can also be dropped due to the homogeneity of speculators.

The dimensionless coefficients d_j for $j \in \{H, M, S\}$ represents the aggressiveness of speculative trading by each type of market participant. For the convenience of our discussion, we introduce and define

$$
\bar{d} = \mu_0\left(\frac{d_H}{A_H G_{11}^H} + \frac{d_M}{A_M G_{11}^M}\right), \tag{1.22}
$$

which is a weighted average of d_H and d_M, and where $\mu_0 = \left(\frac{1}{A_H G_{11}^H} + \frac{1}{A_M G_{11}^M}\right)^{-1}$. The parameter $\bar{d}$ can be interpreted as the effective aggressiveness of speculation by hedgers and market makers who are less informed than speculators in futures market. The following theorem characterizes the imperfectly competitive equilibrium described in Definition 1.

Theorem 1. *There exists a linear equilibrium* $(P, F, \{x_H, y_H\}, \{x_M, y_M\}, y_S)$ *if and only if*

$$\frac{d_S}{\bar{d}} > 1 + \frac{1}{N_S^2} \quad \text{or} \quad \frac{d_S}{\bar{d}} < \frac{N_S^2 - 1}{N_S^2 + \frac{2\mu_0}{A_S \mathrm{Var}_0^S(P)}}. \tag{1.23}$$

Without speculative position limit, the equilibrium futures and spot prices are

$$P = a(s_P - \bar{V}) + b(s_F - \bar{V}) + \bar{V} - \mu \bar{X}, \tag{1.24}$$

$$F = d(s_F - \bar{V}) + h, \tag{1.25}$$

where constants $0 < a < 1$, $0 < b < 1$, $a + b < 1$, $d > 0$ *and* $\mu > 0$. *The equilibrium demands in the spot market are*

$$x_H = \frac{\tau_\eta}{N_H A_H}[V - \bar{V} - a(s_P - \bar{V}) - b(s_F - \bar{V})] - \frac{1-\omega}{N_H}\bar{X}, \tag{1.26}$$

$$x_M = -\frac{\tau_\eta}{N_M A_M}[V - \bar{V} - a(s_P - \bar{V}) - b(s_F - \bar{V})] + \frac{1-\omega}{N_M}\bar{X}, \tag{1.27}$$

where $0 < \omega < 1$. *The equilibrium demands of each hedger, market maker and speculator in the futures market are respectively given by*

$$y_H = \frac{d_H - d}{N_H A_H G_{11}^H}(s_F - \bar{V}) + c_H, \tag{1.28}$$

$$y_M = \frac{d_M - d}{N_M A_M G_{11}^M}(s_F - \bar{V}) + c_M, \tag{1.29}$$

$$y_S = \frac{d_S - d}{N_S A_S \mathrm{Var}_0^S(P) + \lambda / N_S}(s_F - \bar{V}) + c_S, \tag{1.30}$$

where the coefficient d *above is a weighted average of* d_H, d_M, *and* d_S. *It can be shown that* $d_M < d$, *whereas* $d_H - d$ *and* $d_S - d$ *can be positive or negative. In equilibrium, the total price impact of speculators is given by*

$$\lambda = \frac{\mu_0 d_S + A_S \mathrm{Var}_0^S(P)\bar{d}}{d_S - \bar{d} - \bar{d}/N_S^2}. \tag{1.31}$$

Proof. See Appendix 1.8.1. □

According to Theorem 1, the futures demands of hedgers can be positive or negative, depending on the state variable s_F. Note that $c_j = \mathrm{E}[y_j]$ is the average futures position of market participant $j \in \{H, M, S\}$. To clear the market in equilibrium, the average hedging demands of hedgers and market makers must be equal to $N_S c_S = \bar{X} - \theta\bar{V}$, where $\theta = -\frac{G_{12}^H}{G_{11}^H} > 0$ represents the under-hedging by hedgers due to the quantity risk. The element $G_{12}^H < 0$ represents hedgers' effective conditional covariance between the spot price P and the commodity supply X. It can be shown that $\left|G_{12}^H\right|$ increases with the quantity risk and vanishes when there is no quantity risk. Due to symmetry, we can focus on the case of long-side position limit, i.e., $L > 0$. Without loss of generality, this paper assumes that

$$\bar{X} - \theta\bar{V} > 0. \tag{1.32}$$

Under this condition, one can verify that $c_H < 0$ and $c_M > 0$, that is, hedgers take the short position on average whereas market makers take long position as speculators do on average. The condition (1.32) is consistent with the stylized facts in modern commodity futures markets (15), where institutional investors (speculators) take long positions and hedgers take short positions. It is worth remarking that although on average hedgers and speculators take opposite positions in futures, on the margin they can still speculate in response to futures price.

In an environment where the maximal information about the fundamental available to each market is fixed (i.e., given τ_V, τ_η and τ_ϵ), the information asymmetry between speculators and other market participants only depends on τ_z, and the hedging motive of

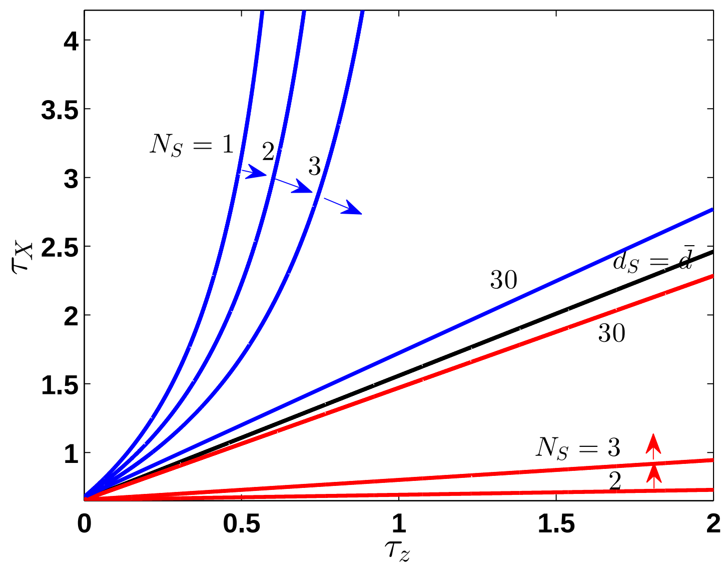


Figure 1.1: The equilibrium existence condition, Eq. (1.23), for different numbers of speculators. An imperfectly competitive equilibrium exists in the area either above the blue line or below the red line in the parameter space (τ_z, τ_X). Other parameters are $\tau_\epsilon/\tau_V = 0.3$, $\tau_\eta = 1$, $N_H = N_M = 30$, $\bar{X} = 1$ and $\bar{V} = 0$.

hedgers is determined by τ_X.[17] The equilibrium existence condition (1.23) in Theorem 1 indicates that an equilibrium exists only when there is sufficient difference of speculation intensities (as characterized by $d_S/\bar{d}$) between speculators and other market participants. When speculators are more and more competitive (as N_S increases), a slight deviation of d_S from $\bar{d}$ will motivate trading to support an equilibrium. This is confirmed in Figure 1.1 as we see that the area where the equilibrium exists in the parameter space (τ_z, τ_X) expands to the whole space excluding the line $d_S = \bar{d}$. When there is only one speculator in the market, there is no equilibrium for the space $d_S < \bar{d}$. The analytical expression for the condition $d_S = \bar{d}$ in terms of (τ_z, τ_X) is derived in Appendix 1.8.3. The line $d_S = \bar{d}$ divides

[17] By the *intertemporal price relation* in 57, hedgers under-hedge as τ_X is low.

the parameter space into two regions where the trading of speculators has different effects on futures price volatility.

Definition 2. For market participant $j \in \{H, M, S\}$, if $\text{Cov}(y_j, F) > 0$ then market participant j is said to **destabilize** the futures price. Otherwise, if $\text{Cov}(y_j, F) < 0$, then market participant j is said to **stabilize** the futures price.

Following the previous literature (27), this paper adopts the term of price destabilizing in the sense that trading makes the price more volatile.[18] Since we have $d_M < d$, it means that market makers buy when price is low and sell when price is high, that is, they are always stabilizing the futures price. The sign of $d_H - d$ can be positive or negative, depending on two effects of futures price: informational effect and cost effect. The standard cost effect suggests that a higher price leads to a lower demand. The informational effect means that a higher futures price signals a stronger fundamental value and thus may motivate hedgers to trade more aggressively on the information revealed from futures price. The term $d_H - d$ nets these two offsetting effects, and it turns out that informational effect can be stronger than the cost effect in the presence of quantity risk. This is similar for speculators. It can be proved that when $d_S > \bar{d}$, speculators destabilize futures price, and when $d_S < \bar{d}$ they stabilize futures price, as summarized in the following proposition.[19]

Proposition 2. *By Definition 2, speculators destabilize futures price if $d_S > \bar{d}$, and they stabilize futures price if $d_S < \bar{d}$.*

[18] 21 use a similar definition to define liquidity consumers. This paper uses the term of *destabilizing prices* because speculators in this model have dual trading motives (speculation and hedging) making it difficult to distinguish whether speculators are consuming or providing liquidity.

[19] We can easily prove that $d_S > d$ is equivalent to $d_S > \bar{d}$ and that $\text{Cov}(y_S, F) \propto d(d_S - d)$.

Proposition 3. *Speculators' price impact is positive, i.e.,* $\lambda > 0$*, if* $d_S > \bar{d}$*, and their price impact is negative, i.e.,* $\lambda < 0$*, if* $d_S < \bar{d}$*. The market power makes speculators trade less aggressively when* $d_S > \bar{d}$*, while the market power makes them trade more aggressively when* $d_S < \bar{d}$*.*

Intuitively, market power makes speculators trade less aggressively (55). This intuition, however, can be flipped when there are multiple trading periods and quantity risk. When $d_S < \bar{d}$, the information effect of futures prices dominates and hedgers utilize the information revealed by futures prices more aggressively. In this case, speculators' demands are downward sloping and thus their price impact is negative. This negative price impact leads to a more aggressive trading by speculators when their market power is not negligible.

1.4 Speculative Position Limit

Now we consider the case in which every speculator is subject to a position limit L:

$$y_S \leq L.$$

In the following, we use the subscript $\cdot_L$ to denote the equilibrium price or demand in the presence of speculative position limit L. As shown in Proposition 1, the state variables for the economy are $\{s_F, s_P\}$ when there is no speculative position limit. Intuitively, the tails of s_F can be affected by the position limit, while s_P is unaffected. When the state variable s_F is neither too large nor too small such that the speculative position limit does not bind, information inference from prices will not be affected. Because of the symmetry, we only focus on the positive position limit. This position limit can bind at a critical value of state variable $\overline{s}_F$ on the right tail of the distribution of s_F. In other words, when s_F is so large that the limit binds, information inference from futures price will be interrupted, leading to the following result:

Proposition 4. *Given a speculative position limit L, there exists a critical value of state variable $\overline{s}_F$ such that in equilibrium both futures and spot prices are independent of s_F for $s_F \geq \overline{s}_F$:*

$$F_L = \overline{F}, \tag{1.33}$$

$$P_L = \overline{P} \equiv \overline{a}(s_P - \bar{V}) + \overline{b} + \bar{V} - \overline{\mu}\bar{X}, \tag{1.34}$$

where $\overline{\mu} > \mu$, $\overline{a} > a$, $\overline{b} = b\mathrm{E}[s_F - \bar{V} \mid s_F \geq \overline{s}_F] > b(\overline{s}_F - \bar{V})$, and $\overline{F}$ are constants depending on L. The spot price P becomes more sensitive to the change of s_P.

Proof. See Appendix 1.8.4. □

The above proposition shows that the spot price P becomes more sensitive to the state variable s_P when the position limit binds since futures price becomes uninformative about the state variable s_F. Note that the futures price becomes a constant $\overline{F}$ when the limit binds. This suggests that the position limit may be substituted by a price limit in futures market.

Corollary 1.1. *The speculative position limit L can be effectively substituted by a price limit $\overline{F}$, and the information inference is*

$$\mathrm{E}[V - \bar{V} \mid F, P] = \begin{cases} [\frac{\tau_P}{\tau_V+\tau_F+\tau_P}(s_P - \bar{V}) + \frac{\tau_F}{\tau_V+\tau_F+\tau_P}(s_F - \bar{V}), & \text{if } F < \overline{F}, \\ [\frac{1}{\tau_V+\tau_F+\tau_P} + q_2(\overline{s}_F)]\tau_P(s_P - \bar{V}) + q_1(\overline{s}_F), & \text{if } F = \overline{F}, \end{cases}$$

$$\mathrm{Var}[V \mid F, P] = \begin{cases} \frac{1}{\tau_V+\tau_F+\tau_P}, & \text{if } F < \overline{F}, \\ \\ \frac{1}{\tau_V+\tau_F+\tau_P} + q_2(\overline{s}_F) > \frac{1}{\tau_V+\tau_F+\tau_P}, & \text{if } F = \overline{F}, \end{cases}$$

where q_1 *and* q_2 *are well-defined functions given in* 1.8.4. *To obtain* $\mathrm{E}[V - \bar{V} \mid F]$ *and* $\mathrm{Var}[V \mid F]$, *one can just take* τ_P *to zero.*

Proof. See Appendix 1.8.4. □

This substitution effect of the position limit results from the assumption that constrained traders is homogeneous and motivated by private information. Any heterogeneity among speculators is expected to weaken this substitution effect, because the position limit will bind at different prices for heterogeneous speculators. Another direct implication of Corollary 4.1 leads to the following statement:

Corollary 1.2. *The speculative position limit does not bias less informed market participants' expectation of future payoff, but it does increase their perceived uncertainty, as measured by their conditional variance about future payoff.*

For risk averse market participants, it is this adverse effect of position limits on their perceived uncertainty that will reduce futures market information efficiency. Here, the information efficiency is quantified by the less informed market participants' perceived variance of V conditional on the futures price F.

When speculators stabilize the futures price (i.e., $d_S < \bar{d}$), their demands are downward sloping so that the position limit may also bind on the left tail of the distribution of s_F. There exists a critical value of state variable $\hat{s}_F$ such that for $s_F \leq \hat{s}_F$ the position limit binds. However, this does not imply that the equilibrium futures price fails to be informative about s_F. In fact, when $d_S < \bar{d}$ it is hedgers who use information from futures price more aggressively than speculators, i.e., hedgers are selling when prices are low. Thus speculators still have incentives to reveal s_F to buy at a lower price. For $s_F \leq \hat{s}_F$, the market clearing price is such that it reveals s_F and clears the futures market with each speculator holding L shares. This result is stated below:

Proposition 5. *When speculators stabilize price, i.e., $d_S < \bar{d}$, given the speculative position limit L, there exists a critical value $\hat{s}_F$ such that for $s_F \leq \hat{s}_F$,*

$$F_L = \hat{F} \equiv \hat{d}(s_F - \bar{V}) + \hat{h}, \tag{1.35}$$

$$P_L = P, \tag{1.36}$$

where $\hat{d} > d$ and $\hat{d}(\hat{s}_F - \bar{V}) + \hat{h} = d(\hat{s}_F - \bar{V}) + h$.

Proof. See Appendix 1.8.5. □

When speculators destabilize price (i.e., $d_S > \bar{d}$), their demands are upward sloping, and the position limit also binds at $\hat{s}_F$ on the right tail of the distribution of s_F if they submit price-elastic demands. Both $\hat{s}_F$ and $\overline{s}_F$ depend on speculators' population N_S. We have $\hat{s}_F < \overline{s}_F$ when speculators are competitive, i.e, $N_S >> 1$. However, in this case, the position limit barely binds.[20] Thus, the case of large N_S is only reported in Appendix 1.8.5. In an oligopoly market with a small number of speculators, we have $\hat{s}_F \geq \overline{s}_F$ and thus $\hat{s}_F$ becomes irrelevant. The following theorem summarizes the equilibrium prices in this case:

[20] Both $\hat{s}_F$ and $\overline{s}_F$ are beyond five standard derivations for $N_S = 15$ and $L/\bar{X} = 0.25$. Hence this case only matters when there is an aggregate position limit on the entire group of speculators, as will be discussed in Section 1.5.

Theorem 2. *With the speculative position limit L in futures market,*

1. *When speculators stabilize price (i.e., $d_S < \bar{d}$), there exist $\hat{s}_F$ and $\overline{s}_F$ such that position limit binds for $s_F \leq \hat{s}_F$ and $s_F \geq \overline{s}_F$, and in equilibrium each speculator submits a price dependent limit order for $s_F \leq \hat{s}_F$ but executed at L, and submits a price independent market order at L for $s_F \geq \overline{s}_F$. The equilibrium prices are*

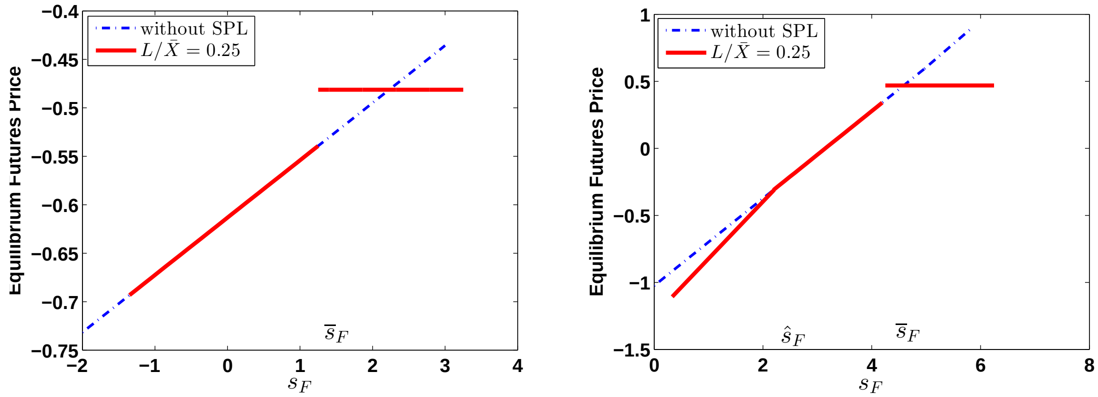
$$F_L = \begin{cases} \hat{F}, & \text{if } s_F \leq \hat{s}_F \\ F, & \text{if } \hat{s}_F < s_F < \overline{s}_F \\ \overline{F}, & \text{if } s_F \geq \overline{s}_F \end{cases} \qquad P_L = \begin{cases} P, & \text{if } s_F < \overline{s}_F \\ \overline{P}, & \text{if } s_F \geq \overline{s}_F \end{cases} \tag{1.37}$$

2. *When speculators destabilize price (i.e., $d_S > \bar{d}$), there exists $\overline{s}_F$ such that position limit binds for $s_F \geq \overline{s}_F$, and in equilibrium each speculator submits a price independent market order at L for $s_F \geq \overline{s}_F$. The equilibrium prices are*

$$F_L = \begin{cases} F, & \text{if } s_F < \overline{s}_F \\ \overline{F}, & \text{if } s_F \geq \overline{s}_F \end{cases} \qquad P_L = \begin{cases} P, & \text{if } s_F < \overline{s}_F \\ \overline{P}, & \text{if } s_F \geq \overline{s}_F \end{cases} \tag{1.38}$$

Proof. See Appendix 1.8.5. The equilibrium demands in futures and spot markets are also given in Appendix 1.8.5. □

One claimed objective of imposing speculative position limits is to prevent commodity price bubbles. However, as shown in Figure 1.2, futures price may jump up when the realized state s_F is close to the critical value $\overline{s}_F$. This corresponds to the situation when futures price suddenly becomes uninformative and less informed market participants perceive a sharp increase in uncertainty about the commodity fundamental. This price surge has a



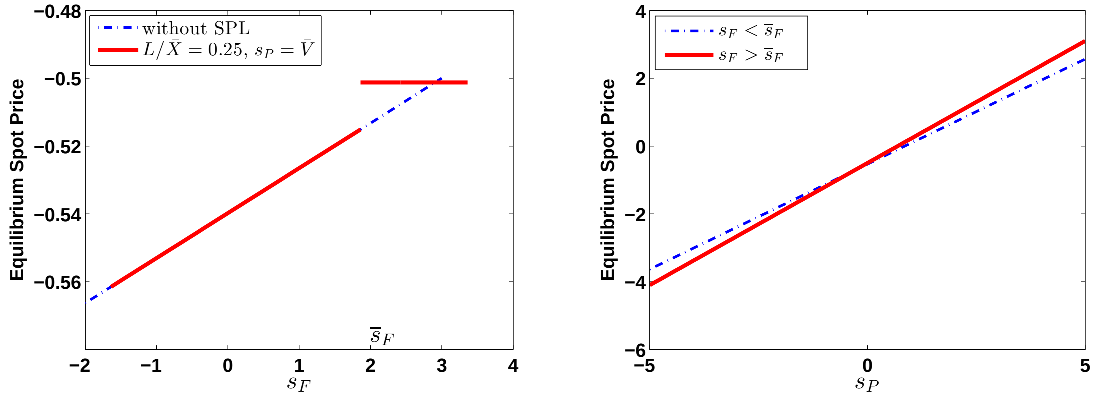
(a) Destabilize Price (b) Stabilize Price

Figure 1.2: Equilibrium futures price with respect to the state variable s_F for markets with two speculators $N_S = 2$. Dash-dot lines represent the equilibrium price without the speculative position limit (SPL), while solid lines denote the price with position limits $L/\bar{X} = 0.25$. (a) Speculators destabilize price: $d_S = 2.5\bar{d}$. (b) Speculators stabilize price: $d_S = 0.1\bar{d}$. Other parameters are $\tau_\epsilon/\tau_V = 0.3$, $\tau_\eta = 1$, $N_H = N_M = 30$, $\bar{X} = 1$, and $\bar{V} = 0$.

spillover effect on spot market as Figure 1.3(a) shows. Since futures price has a positive impact on spot price due to information learning, a sudden increase in futures price will also lead to a jump in spot price. Furthermore, when position limit binds, the spot price becomes more sensitive to the realized state s_P and hence amplifies the changes in s_P as Figure 1.3(b) shows.

1.4.1 Equilibrium Positions

In position-limit-rule proponents' mind, the limits will not affect speculators' trading strategies. In other words, when the limits do not bind, speculators will follow the same trading strategy as if there were no limits for them. As Figure 1.4 shows below, position-limit-rule



(a) (b)

Figure 1.3: Equilibrium spot price with respect to the state variables (a) s_F and (b) s_P for markets with two speculators $N_S = 2$. In panel (a), the dash-dot line represents the equilibrium price without the speculative position limit (SPL) and the solid line denotes the price with position limits $L/\bar{X} = 0.25$. In panel (b), the dash-dot line represents the price when SPL does not bind ($s_F < \overline{s}_F$) and the solid line is the price when SPL binds ($s_F > \overline{s}_F$). Other parameters are $d_S = 1.5\bar{d}$, $\tau_\epsilon/\tau_V = 0.3$, $\tau_\eta = 1$, $N_H = N_M = 30$, $\bar{X} = 1$ and $\bar{V} = 0$.

proponents expect that speculators' positions will not be affected by the position limits if the limits do not bind. Therefore, CFTC calculated the binding probability of the proposed rules based on the data without limits.

However, speculators' position demonstrated in Figure 1.4 is not an equilibrium. Because position limits can benefit constrained speculators in two ways. First, when the limits bind, speculators can submit price-independent market orders at the limits, which mitigates the adverse selection problem of speculators and thus reduces the cost of their price impacts. This effect is more prominent when speculators provide liquidity in the futures, because in this case their information is more valuable to uninformed hedgers. Second, the position

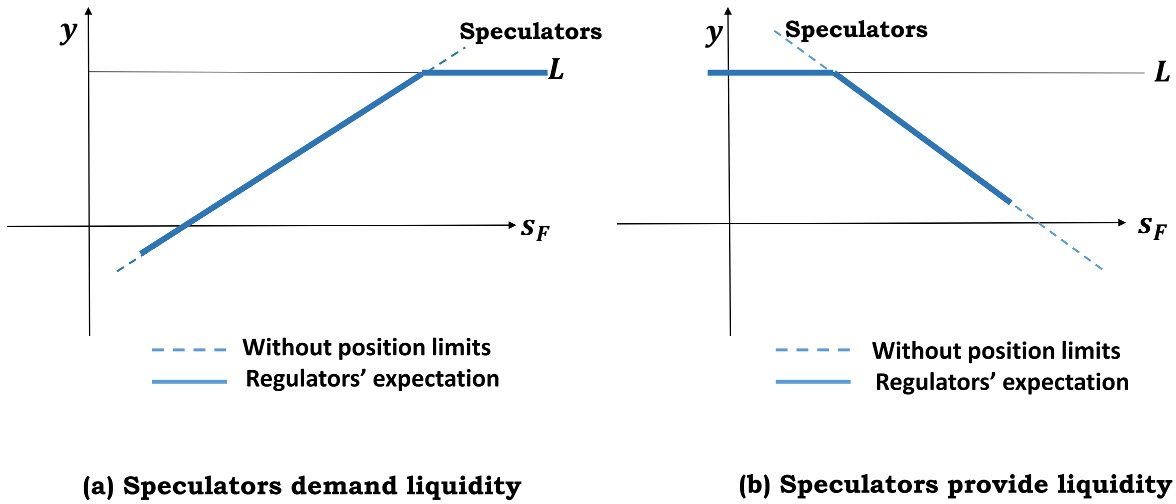


Figure 1.4: Position-limit-rule proponents expect that speculators' positions will not be affected by the position limits if the limits do not bind.

limits have a cartel effect. Without the limits, even though speculators know that they can get a better price by constraining their supply in the futures, there is no enforcement to do so. Position limits form a natural cartel for them and give them a higher return. Because of these two reasons, speculators have incentives to make the limits bind on them. Therefore, in equilibrium they will implement their market power to make the limits bind in the cases their positions will below the limits if there were no limits as Figure 1.5 shows. In other words, the limits are more likely to bind than CFTC expected, especially in the markets where speculators have significant market powers.

Proposition 6. *The binding probability of the limits is higher than the probability that*

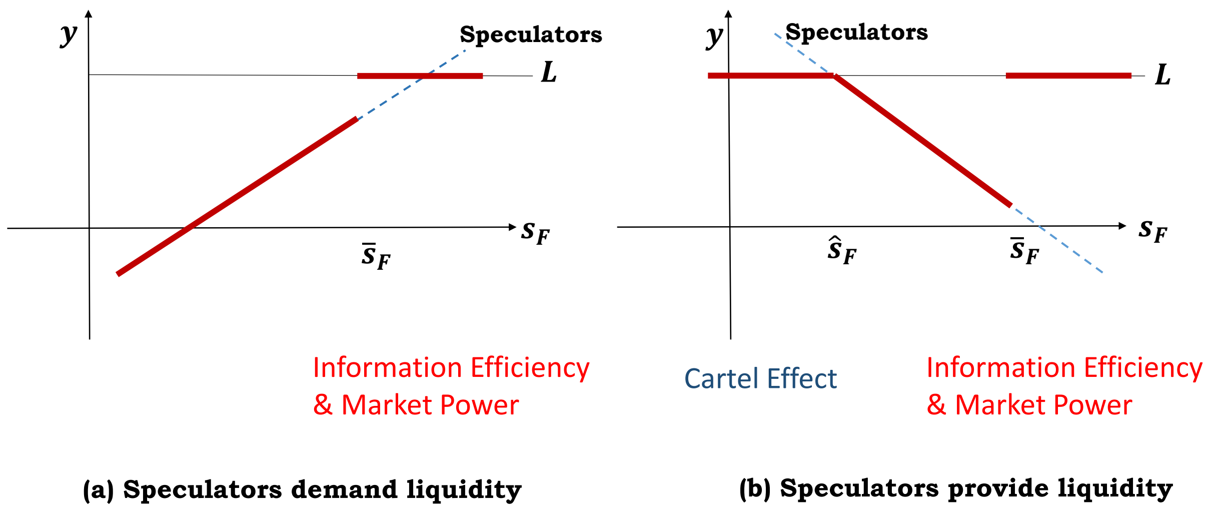


Figure 1.5: The equilibrium positions of speculators under individual position limits.

speculators' positions are higher than the limits if there were no limits, especially in the markets where market makers have significant market powers.

1.4.2 Market Liquidity and Risk Premia

One major concern of the position limit rule is that how it affects market liquidity. The Commission believes that market liquidity will not be harmed based on two arguments. First, speculators are liquidity consumers when they destabilize prices. Second, there exist a large number of unconstrained traders (market makers) who will keep providing liquidity to hedgers when position limits bind on speculators. In reality, speculators have dual trading motives: speculation and hedging. Speculation is traditionally viewed to provide liquidity while hedging is viewed to consume liquidity. The dual trading motives of speculators make it difficult to disentangle whether they are providing liquidity or not. Moreover, one trader's liquidity provider can be another trader's liquidity consumer. Thus, it is ambiguous to consider speculators as liquidity consumers without identifying from whose perspective. For this reason, this paper avoids to use the terms of liquidity consumer or liquidity provider. Instead, this paper directly focuses on the measure of market liquidity. There are different measures of market liquidity proposed in the literature. 55 suggests that market liquidity can be measured by the coefficient of regressing price on trading quantity. This approach is suitable for a linear equilibrium price. As shown in last section, the equilibrium price is not linear in trading quantity anymore when there is position limit. An alternative measure of market illiquidity in the literature is based on the intuition of 50 and 23 that the lower the autocorrelation in rates of return, the higher the equilibrium level of market liquidity.

Definition 3 (Return Autocorrelation)**.** Define γ as the autocorrelation between the price changes in futures market:

$$\gamma = \frac{\mathrm{Cov}(P - F, F - \mathrm{E}[F])}{\sqrt{\mathrm{Var}(P - F)\mathrm{Var}(F - \mathrm{E}[F])}}. \tag{1.39}$$

It is easy to show that the return autocorrelation γ is negative. We denote the return autocorrelation with speculative position limit L by γ_L. Then, as proved in Appendix 1.8.6, we have the following relationship

Proposition 7. *The return autocorrelation with the position limit L is less negative than that without the position limit, i.e.,*

$$\gamma < \gamma_L < 0. \tag{1.40}$$

The less autocorrelation is due to the fact that prices become less informative with position limit. It does not mean that the market liquidity will be improved after imposing position limits. Indeed, return autocorrelation is an appropriate measure of illiquidity when it quantifies price reversals after a major liquidity shock where trading is not driven by information.

Since a lot of trading in futures market is motivated by private information, here I use a marginal price impact measure similar to Kyle's λ, which is the marginal price change caused by a small change in the aggregate demand, to quantify market illiquidity. In its definition below, the small change δy (a perturbation) can be interpreted as some exogenous market orders (a liquidity event) coming to the market.

Definition 4 (Market Illiquidity Measure)**.** Market illiquidity measure is defined as the marginal price impact caused by an infinitesimal liquidity shock.

1. In futures market, by market clearing $N_H y_H + N_M y_M + N_S y_S + \delta y = 0$ where δy is the liquidity shock, we define the illiquidity measure of futures market as

$$\lambda^F \equiv \lim_{\delta y \to 0} \frac{\partial F}{\partial \delta y}. \tag{1.41}$$

2. In spot market, by market clearing $N_H x_H + N_M x_M + \delta x = 0$ where δx is the liquidity shock, we define the illiquidity measure of spot market as

$$\lambda^P \equiv \lim_{\delta x \to 0} \frac{\partial P}{\partial \delta x}. \tag{1.42}$$

It is then straightforward to calculate λ^F and λ^P in this model.

Proposition 8. *With a speculative position limit L:*

1. *The futures market illiquidity is given by*

$$\lambda^F = (\frac{1}{A_H G_{11}^H} + \frac{1}{A_M G_{11}^M} + \frac{N_S 1_{y_S < L}}{N_S A_S \mathrm{Var}_0^S(P) + \lambda / N_S})^{-1}, \tag{1.43}$$

where $G_{11}^{H,M}$ *denote the hedgers' and market makers' effective conditional variance of spot price respectively.*

2. *The spot market illiquidity is given by*

$$\lambda^P = (\frac{1}{A_H \mathrm{Var}_1^H(V + \eta)} + \frac{1}{A_M \mathrm{Var}_1^M(V + \eta)})^{-1}. \tag{1.44}$$

When the position limit binds, both λ^F *and* λ^P *increase.*

As can be seen from Eq. (1.43), λ^F can increase through two channels when the position limit binds. First, speculators stop contributing to market liquidity beyond position limit, i.e., $1_{y_S<L} = 0$. Second, the perceived uncertainty by hedgers and market makers rises, i.e., $G_{11}^{H,G}$ increases. When the limit binds, there is an illiquidity spillover effect on spot market. As Eq. (1.44) shows, λ^P increases due to less information learning by market makers. This illiqudity is generated by a sudden surge in uncertainty as perceived by less informed traders.

One of the rationales for developing commodity futures markets is to facilitate hedgers to unload their commodity risks to other economic agents. In Keynes-Hicks theory, speculators collect the risk premium by sharing risks with hedgers in the futures market. The risk premium increases with the net hedging demands (49). In Working theory (58), it is the speculators who shift risks to hedgers and pay hedgers the risk premium in futures market. The empirical evidence is mixed. 12 shows that the observed biases in futures prices, i.e., the futures basis, are related to the net positions of hedgers, which supports Keynes-Hicks theory. Nonetheless, a more recent study 22 shows that the observed biases reflect the state of inventory, which is more consistent with Working theory (58). My model reconciles these two seemingly conflicting theories by showing that hedgers can utilize the intertemporal price relation to under-hedge on average, and both hedgers and speculators engage in speculation on the margin. Therefore, the risk premium is determined by the average hedging demand (from Keynes-Hicks theory) and the speculative trading on the margin (from Working theory). It is the market liquidity that determines the speculative trading on the margin. In other words, liquidity is not free for hedgers.

Proposition 9. *The expected returns in futures and spot markets respectively are*

$$\mathrm{E}[P - F] = \lambda^F(\bar{X} - \theta\bar{V}), \tag{1.45}$$

$$\mathrm{E}[V + \eta - P] = \lambda^P \bar{X}. \tag{1.46}$$

where $\theta > 0$ represents the under-hedge by hedgers.

Corollary 2.1. *When the average hedging demand in futures is positive (i.e., $\bar{X} - \theta\bar{V} > 0$), the speculative position limit increases risk premia in both futures and spot markets as it reduces market liquidity.*

This result is consistent with empirical evidence that commodity financialization reduces risk premia in crude oil market (25). Imposing position limits on speculators can increase the risk premia in both futures and spot markets.[21] Position limits can not only significantly increase illiquidity measure λ^F and λ^P, but also marginally change θ. When

[21] The risk premia may be linked to the stochastic discount factor in this economy. Following 18, we conjecture that without the speculative position limit there may exist an *Equivalent Martingale Measure* (EMM) Q-measure, under which $F = \mathrm{E}^Q[P]$, $P = \mathrm{E}^Q[V + \eta]$. With position limit, if such a Q-measure exists, the expectation of payoff is conjectured to serve as a lower bound of the equilibrium price; see 47.

speculators stabilize price, position limit decreases θ, thus makes risk premia of futures even higher. When speculators destabilize price, position limit increases θ slightly. But the increase in λ^F dominates, thus risk premia are higher with position limits.

1.4.3 Welfare

In this section, I study how position limits affect the welfare of different market participants. The welfare and the certainty equivalent wealth (CEW) of market participant $j \in \{H, M, S\}$ are defined as

$$\mathcal{W}^j = \mathrm{E}[J_0^j], \tag{1.47}$$

$$\mathrm{CEW}^j = -log(-\mathcal{W}^j). \tag{1.48}$$

The calculation of the welfare and CEW can be found in Appendix 1.8.7.

Intuitively, when setting position limits regulators may expect a welfare transfer from constrained speculators to unconstrained hedgers. However, whether the welfare of constrained speculators are reduced or not is unclear. There are two opposite effects of position limits on speculators: price effect and quantity effect (56). The price effect comes from the increased risk premia, which benefit speculators. This is because position limit moves the price in the direction that favors speculators on average. The quantity effect comes from decreased trading quantity, which hurts speculators by restricting them from trading the optimal quantity. Thus, if the position limit is not too small, i.e., speculators are not too far from their optimal trading quantity, the price effect will dominate the quantity effect and speculators are better off with position limits. If the position limit L becomes too small, then the quantity effect dominates price effect and thus speculators are worse off with position limits. These two effects also depend on the market power of speculators and the information asymmetry between speculators and other market participants. The market power exerted by speculators dampens the quantity effect, and the increased information asymmetry due to the position limit exacerbate the price effect.

The two effects also depend on the competitiveness among speculators. The position limit binds in a market with a large number of speculators at a low probability, thus it should have little impact on a competitive market. For an oligopoly market with a few speculators, the position limit binds with a high probability. Thus, given a position limit L, whether speculators' welfare increases or decreases is affected by their competitiveness, i.e., their population N_S.

Proposition 10. *Given the information environment and market conditions, for a speculative position limit L, there exists a critical number of speculators N_S^L such that*

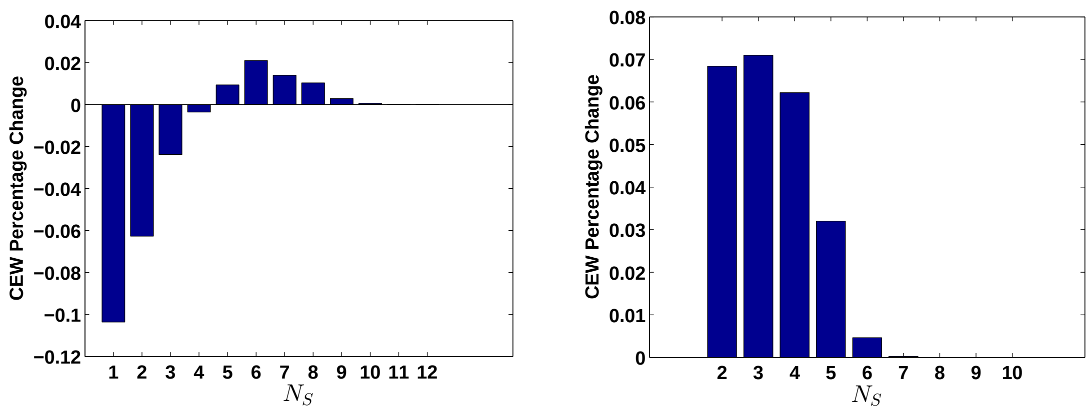
1. *If $N_S < N_S^L$, speculators are worse off with the position limit;*

2. *If $N_S \geq N_S^L$, speculators are better off with the position limit.*

N_S^L increases with $d_S/\bar{d}$.

Proof. See Appendix 1.8.7. □

Figure 1.6 illustrates how speculators' certainty equivalent wealth (CEW) change under a speculative position limit in two different markets: (a) when $d_S > \bar{d}$ ($N_S^L = 5$) and (b) when $d_S < \bar{d}$ ($N_S^L = 2$). In both markets, the position limit has small effects on competitive markets with a large population of speculators, but has significant effects on oligopoly markets with a small number of speculators. Figure 1.6(a) shows that when speculators destabilize prices, they are harmed by this rule if there are less than five speculators, they benefit from this rule if there are more than five speculators, and they are barely affected by this rule if there are more than ten speculators. Figure 1.6(b) shows that when speculators stabilize prices, they always benefit from this rule, especially when there are less than five speculators in the market.

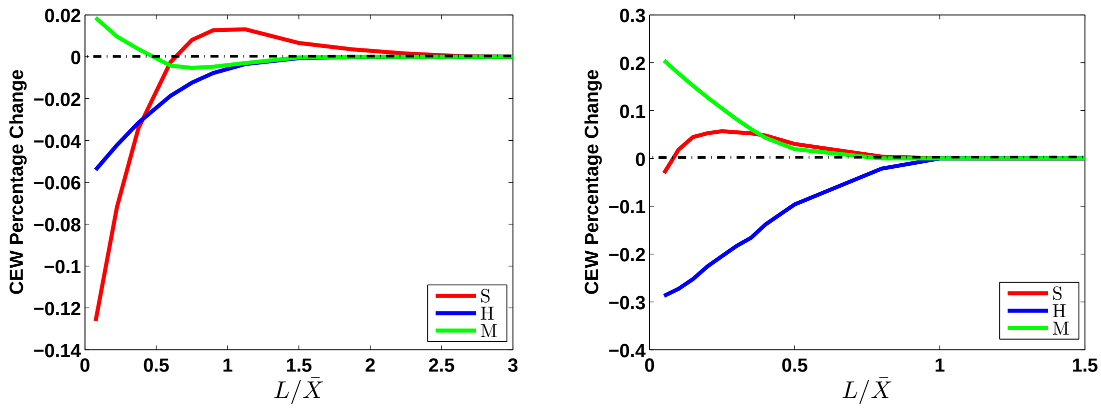


(a) Destabilize Prices

(b) Stabilize Prices

Figure 1.6: Percentage change of speculators' Certainty Equivalent Wealth (CEW) under the position limit $L/\bar{X} = 0.25$ with respect to their population N_S. (a) Speculators destabilize price: $d_S = 2.5\bar{d}$. (b) Speculators stabilize price: $d_S = 0.1\bar{d}$. Other parameters are $\tau_\epsilon/\tau_V = 0.3$, $\tau_\eta = 1$, $N_H = N_M = 30$, $\bar{X} = 1$, and $\bar{V} = 0$.

Position limits on speculators always reduce hedgers' welfare. As Figure 1.7 shows, in a duopoly market with only two speculators especially when they stabilize prices. Hedgers' CEW decreases as position limit L decreases. Position limits have a positive price effect on market makers, because on average market makers trade in the same direction as speculators. Position limits have two effects on the trading quantities of market makers, although they are not subject to the position limit. On one hand, the higher risk premia induce them to



(a) Destabilize Prices

(b) Stabilize Prices

Figure 1.7: Percentage changes of the Certainty Equivalent Wealth (CEW) for different market participants versus position limit $L/\bar{X}$. The symbols $\{S, H, M\}$ in the figure correspond to speculators, hedgers, and market makers. (a) Speculators destabilize price: $d_S = 2.5\bar{d}$. (b) Speculators stabilize price: $d_S = 0.1\bar{d}$. Other parameters are $\tau_\epsilon/\tau_V = 0.3$, $\tau_\eta = 1$, $N_S = 2$, $N_H = N_M = 30$, $\bar{X} = 1$ and $\bar{V} = 0$.

trade more. On the other hand, market makers lose valuable information if the limits bind on speculators. Such information disadvantage increases their perceived uncertainty and thus makes them trade less. As the position limit L decreases, the price effect dominates. Thus, market makers are better off with a small position limit and worse off with a large position limit, as shown Figure 1.7(a).

1.4.4 Information Externalities on Spot Market

If the position limit does not bind, then it has no impact on spot market. If it binds, however, then there should be information externalities on spot market. These externalities are more prominent when quantity risk is high, i.e τ_X is small. In cases that position limit binds at a point beyond the critical value $\overline{s}_F$, the futures price F (public information) becomes less informative. This reduced public information increases information asymmetry in spot market between market makers and hedgers, resulting in an increase of the hedging cost (i.e., market makers charging hedgers a higher risk premium). As a result, the trading volume in spot market decreases when the position limit binds in futures. In other words, the scope of risk sharing in spot market is reduced, and the welfare of both market makers and hedgers are adversely affected by the position limit.

Proposition 11. *If the speculative position limit binds in the futures market, the interim utilities of hedgers and market makers from spot market decreases, especially when quantity risk is high. Information revealed by futures price is more valuable for the market participants who bear the quantity risk.*

Proof. See 1.8.8. □

Previous literature (52) argues that the information revealed by futures price can harm hedgers, because in spot market hedgers have information advantage over market makers. This intuition is correct only if there is no quantity risk. For informed spot market participants with quantity risk (hedgers), the loss of information advantage due to public information is always outweighed by the increased scope of risk sharing. The position limit in futures hampers information revealing of futures price, thus it always has negative externalities on both informed and less informed spot market participants. In this model, hedgers are better informed than market makers. Thus, it seems that the information revealed by futures price is useless for hedgers, it is more valuable for market makers. Proposition 11, however, shows that it is the opposite: the information revealed by futures price is more valuable for hedgers who bear the quantity risk, no matter how informed they are.

1.5 Competition under an Aggregate Position Limit

Also approved at the November 5 meeting was a proposed rule on aggregation of positions, which the commission deems necessary in order to prevent affiliated companies from sidestepping position limits rules by distributing positions among affiliates. The Commission and the exchanges treat multiple positions subject to common ownership or control as if they were a single trader. Accounts are considered to be under a common ownership if there is a 10 percent or greater financial interest. The rules are applied in a manner calculated to aggregate related accounts. For example, each participant with a 10 percent or greater interest in a partnership account must aggregate the entire position of the partnership not just the participant's fractional share-together with each position they may hold separately from the partnership. Likewise, a pool comprised of many traders is allowed only to hold positions as if it were a single trader. The Commission also treats accounts that are not otherwise related, but are acting pursuant to an express or implied agreement, as a single aggregated position for purposes of applying the limits. The Commission believes that the aggregate position limit can promote market competition.[22]

This section examines how an aggregate position limit affects a pool of speculators compete on their common information. Therefore, in this section, there are N_S speculators who are allowed only to hold positions as if they were a single trader. In other words, each

[22] http://www.cftc.gov/IndustryOversight/MarketSurveillance/SpeculativeLimits/index.htm.

of them is subject to a position limit L/N_S:

$$y_S \leq L/N_S.$$

Surprisingly, the aggregate position limit may disincentive speculators to compete on their common information under certain market conditions. As Figure 1.8 shows, without an aggregate position limit, speculators' CEW is greater if they fully compete on their common information. However, with an aggregate position limit $L/\bar{X} < 2$, their CEW is greater if they act as an information monopolist. This result comes from the following trade-off. On one hand, speculators can extract more rent from their information advantage by acting as an information monopolist. On the other hand, speculators face more severe adverse selection as a monopolist. An aggregate position limit would mitigate the impacts of this adverse selection. Thus, under certain market conditions, the aggregate position limit makes speculators better off if they act as an informational monopolist. Proposition 12 summarizes such market conditions.

Proposition 12. *For $\tau_\epsilon/\tau_V < \frac{\sqrt{5}-1}{2}$ and $N_S >> 1$, given an aggregate position limit L, there exits a critical value $\kappa > 0$ such that for $2\bar{d} < d_S < (\kappa+2)\bar{d}$,*

1. *with the aggregate position limit, speculators' total CEW of acting as an information monopolist is greater than their total CEW of fully competing on their common information;*

2. *without the aggregate position limit, it is the opposite;*

3. *κ increases as L decreases.*

Proof. See Appendix 1.8.9. □

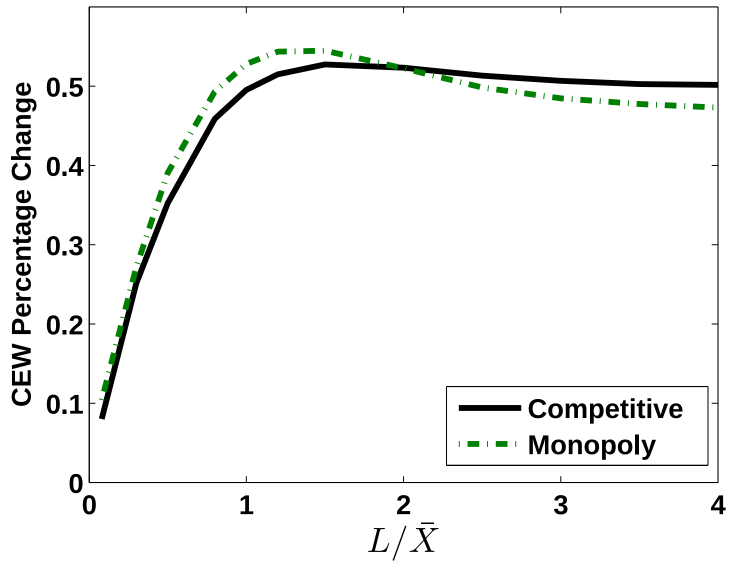
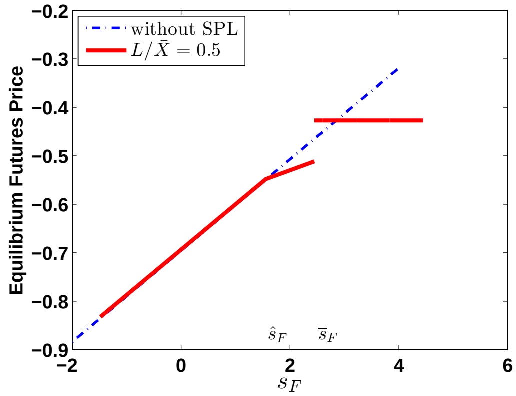
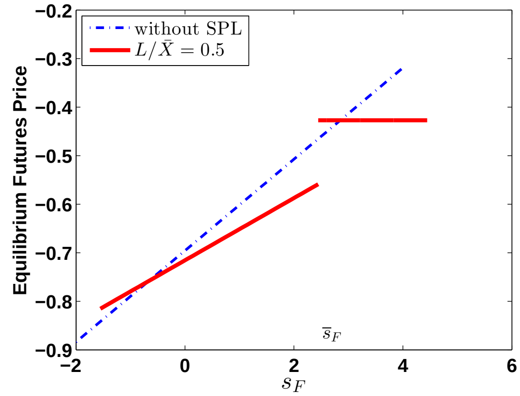


Figure 1.8: Percentage changes of speculators' CEW relative to no trading versus the position limit $L/\bar{X}$. The solid line represents the case in which speculators fully compete on their common information. The dash-dot line represents the case in which speculators act as an information monopolist. Other parameters are $d_S = 2.5\bar{d}$, $\tau_\epsilon/\tau_V = 0.3$, $\tau_\eta = 1$, $N_S = 15$, $N_H = N_M = 30$, $\bar{X} = 1$ and $\bar{V} = 0$.

In markets that satisfy the conditions in Proposition 12, an aggregate position limit may potentially reduce speculators' competition even when the limit does not bind. As Figure 1.9 shows, in the presence of an aggregate position limit the futures price becomes less sensitive to state variable s_F due to the market power exerted by speculators acting as an informational monopolist. As a result, a potential equilibrium position of speculators is shown in Figure 1.10. When speculators demand liquidity in futures, their imperfect competition makes them trade less aggressively, and in this case an aggregate position limit may lead them to trade even lesser aggressively. When speculators provide liquidity in futures, their imperfect competition makes them trade more aggressively, and in this case an



(a) Competitive



(b) Monopoly

Figure 1.9: The equilibrium futures price with respect to the state variable s_F for markets with $N_S = 15$ and an aggregate position limit $L/\bar{X} = 0.5$. Dash-dot lines represent the equilibrium futures prices without the aggregate speculative position limit (SPL) and solid lines denote the prices with the aggregate SPL $L/\bar{X} = 0.5$ (a) Speculators fully compete on their common information. (b) Speculators act as an information monopolist. Other parameters are $d_S = 2.5\bar{d}$, $\tau_\epsilon/\tau_V = 0.3$, $\tau_\eta = 1$, $N_H = N_M = 30$, $\bar{X} = 1$, and $\bar{V} = 0$.

aggregate position limit may lead them to trade even more aggressively. Therefore, another potential unintended consequence can be concluded that an aggregate position limit may exacerbate the imperfect competition among speculators. In markets that do not satisfy the conditions in Proposition 12 (e.g., when speculators are stabilizing price so that $d_S < \bar{d}$), an aggregate position limit may improve speculators' welfare (Figure 1.6) at the expense of hedgers (as shown in Figure 1.7).

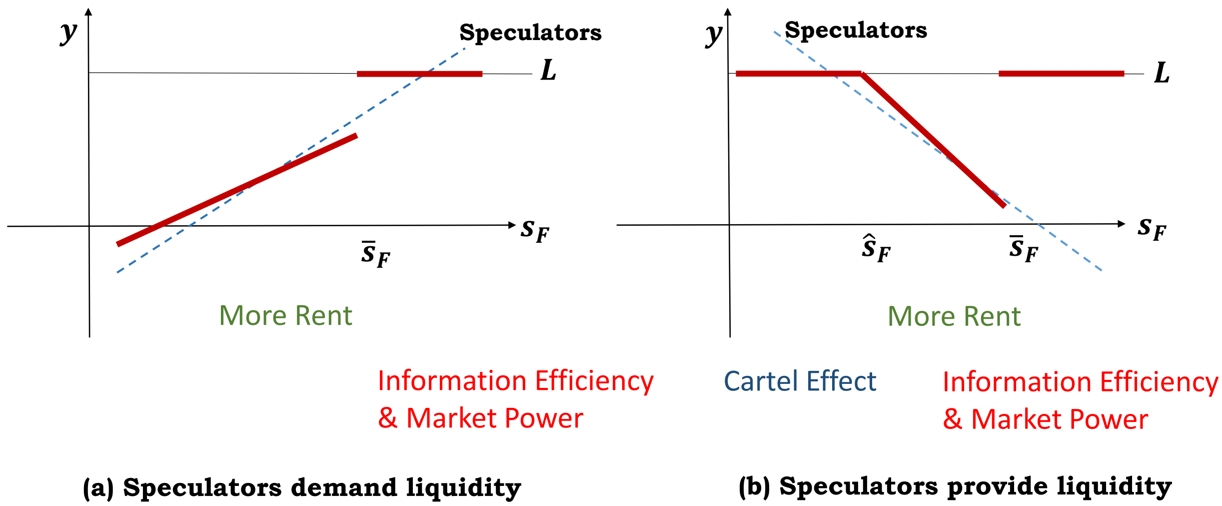


Figure 1.10: A potential equilibrium position of speculators under an aggregate position limit.

1.6 Impacts on Price Volatility and Skewness

From Corollary 2.1, it is straightforward to conclude that the ex-ante average futures and spot price levels are lower in the presence of position limits. However, the realized prices can be higher than those in the case of no position limit. Moreover, both futures prices and spot prices exhibit discontinuity with respect to the state variable s_F.

One purpose of implementing position limits is to curb excessive fluctuations in commodity prices. By the law of total variance, the variance of the spot price can be written as

$$\mathrm{Var}(P) = \mathrm{E}[\mathrm{Var}(P \mid F)] + \mathrm{Var}(\mathrm{E}[P \mid F]). \tag{1.49}$$

The volatility of the commodity price can be decomposed into two parts. The first part denotes the variance of spot price predicted by the futures price F, which comes from speculators' trades in futures. The second part denotes the residual variance, which comes from hedgers' trades in spot market. The position limit has opposite effects on these two parts. It reduces the first part by directly limiting speculators' trades in futures. It, however, increases the second part by making the spot price P more responsive to s_P as a result of hedgers' trades in spot market. Therefore, position limit can increase or decrease the commodity price volatility by netting the two effects. When information asymmetry in futures is low (i.e., large τ_z) and quantity risk is high (i.e., small τ_X), the first part is small and the second part is large. This is the case of $d_S < \bar{d}$, in which speculators' trading stabilizes futures price, (i.e., speculators buy when price is low and sell when price is high). Thus, in this case imposing speculative position limits should destabilize both futures and spot prices.

Proposition 13. *When speculators stabilize futures price, i.e., $d_S < \bar{d}$, the speculative position limits increase the volatilities of both futures and spot prices,*

$$\mathrm{Var}(F_L) > \mathrm{Var}(F), \quad \mathrm{Var}(P_L) > \mathrm{Var}(P).$$

Thus, contrary to the policymakers' belief, imposing position limits can increase price volatility under certain market conditions (when speculators buy low and sell high). In order to study price skewness, I define the average skewness of short-horizon returns in futures market by

$$\mathrm{SKEW} = \frac{1}{2}\mathrm{Skew}[F] + \frac{1}{2}\mathrm{Skew}[P - F], \tag{1.50}$$

where $\mathrm{Skew}[x] = \mathrm{E}[(x - \mathrm{E}[x])^3]$. Without position limits, the average skewness is zero by normality: $\mathrm{Skew}[F] = \mathrm{Skew}[P - F] = 0$. With position limits, the first term is negative:

Skew$[F_L] < 0$, and the second term is positive: Skew$[P_L - F_L] > 0$. The average skewness with position limit can be positive or negative, depending on market conditions.

1.7 Conclusion

To evaluate the impacts of speculative position limits in futures market, I develop a tractable and flexible two-period rational expectation model, which jointly characterizes the equilibrium of commodity futures and spot markets with strategic speculators. I find that the equilibrium for an oligopoly market (with a few speculators) only exists when hedging motive for speculators is high or hedgers' quantity risk is high, while the equilibrium for a competitive market (with a large number of speculators) alway exists. Speculators destabilize prices when their hedging motive outweighs hedgers' hedging motive, and they are stabilize prices otherwise. When speculators stabilize prices, the imperfect competition makes speculators trade more aggressively due to a negative price impact.

This paper shows that, when the speculative position limit binds, market liquidity will be harmed for two reasons. First, the perceived riskiness by less informed market participants increases, because the binding of position limits interrupts the information revealing. Second, speculators' contribution to market liquidity vanishes when they submit price-inelastic orders at position limits. The reduction of market liquidity raises the cost of hedging and makes hedgers worse off.

This paper further examines how speculative position limits work in markets with different competitiveness among speculators. I show that position limits have a marginal effect on competitive markets with a large number of speculators. In oligopoly markets with a few speculators, however, position limits can significantly improve speculators' welfare at the expense of hedgers, especially when speculators stabilize futures prices. Thus, it is inappropriate for regulators to implement a uniform position limit on all commodity futures without considering the heterogeneity in market competitiveness. This paper also examines how an aggregate position limit affects a competitive market. I find that an aggregate position limit may hurt competitive speculators but benefit less competitive speculators. This implies that an aggregate position limit can potentially reduce speculators' competition even when the limit does not bind.

Finally, the binding of position limits can dramatically increase the levels of both futures and spot prices, as it increases the uncertainties perceived by less informed market participants. Moreover, when speculators stabilize futures prices, imposing position limits on them increases both futures and spot price volatilities.

1.8 Appendix

1.8.1 Benchmark Model: No Speculative Position Limit

In this section, I solve the benchmark model in the absence of position limits. I assume that there is a finite number of speculators, i.e., $1 \leq N_S << N_{H,M}$. In other words, both hedgers and market makers have negligible market power, while speculators are imperfectly competing on their common information. The model can be solved backward, and the following lemma will be used repeatedly.

Lemma 1. *Let u be an $n \times 1$ normal vector with mean $\bar{u}$ and covariance matrix Σ, A a scalar, B an $n \times 1$ vector, C an $n \times n$ symmetric matrix, I the $n \times n$ identity matrix, and $|M|$ the determinant of a matrix M. Then,*

$$\mathrm{E}_u \exp\{-\rho[A + B^\top u + \frac{1}{2}u^\top C u]\} = \frac{1}{\sqrt{|I + \rho C\Sigma|}} \exp\Big\{ -\rho[A + B^\top \bar{u} + \frac{1}{2}\bar{u}^\top C\bar{u} \\ -\frac{1}{2}\rho(B + C\bar{u})^\top(\Sigma^{-1} + \rho C)^{-1}(B + C\bar{u})]\Big\}.$$

I first solve the benchmark model with no position limits on speculators. For every commercial hedger, her terminal wealth at $t = 2$ is

$$W^H = (P - F)y_H - Px_H + (V + \eta)(x_H + X/N_H), \tag{1.51}$$

which is normally distributed. So her expected utility at $t = 1$ is

$$\mathrm{E}_1^H[-e^{-N_H A_H W^H}] = -\exp\Big\{ - N_H A_H[(P - F)y_H + PX/N_H + (V - P)(x_H + X/N_H) \\ -\frac{1}{2}N_H A_H \mathrm{Var}_1^H(\eta)(x_H + X/N_H)^2]\Big\}. \tag{1.52}$$

Taking P as given, each commercial hedger maximizes her expected utility at $t = 1$ by choosing

$$x_H = \frac{V - P}{N_H A_H \mathrm{Var}_1^H(\eta)} - \frac{X}{N_H}. \tag{1.53}$$

The second order condition to ensure the optimality requires that $N_H A_H \mathrm{Var}_1^H(\eta) > 0$, which is satisfied automatically.

For every market maker, the terminal wealth at $t = 2$ is

$$W^M = (P - F)y_M + (V + \eta - P)x_M, \tag{1.54}$$

which is normally distributed. So each market maker's expected utility at $t = 1$ is

$$\begin{aligned} \mathrm{E}_1^M[-e^{-N_M A_M W^M}] = -\exp\big\{ -N_M A_M[(P - F)y_M + (\mathrm{E}_1^M[V] - P)x_M \\ -\frac{1}{2}N_M A_M \mathrm{Var}_1^M(V + \eta)x_M^2]\big\}. \end{aligned} \tag{1.55}$$

To maximize this expected utility, each market maker chooses

$$x_M = \frac{\mathrm{E}_1^M[V] - P}{N_M A_M \mathrm{Var}_1^M(V + \eta)}. \tag{1.56}$$

The second order condition to ensure the optimality requires that $N_M A_M \mathrm{Var}_1^M(V + \eta) > 0$ which is satisfied automatically.

For commercial hedgers, we plug Eq. (1.53) into Eq. (1.52) and obtain

$$J_1^H = -e^{-N_H A_H[(P-F)y_H + P\frac{X}{N_H} + \frac{1}{2}N_H A_H \mathrm{Var}_1^H(\eta)(x_H + \frac{X}{N_H})^2]}. \tag{1.57}$$

To compute the expectation $\mathrm{E}_0^H[J_1^H]$, I use Lemma 7 and set $\rho = A_H$, $A = -FN_H y_H$,

$\bar{u} = \mathrm{E}_0^H[u]$, $\Sigma = \mathrm{Cov}_0^H(u)$, and

$$u = \begin{pmatrix} P \\ \\ X \\ \\ N_H x_H + X \end{pmatrix}, \quad B = \begin{pmatrix} N_H y_H \\ \\ 0 \\ \\ 0 \end{pmatrix}, \quad C = \begin{pmatrix} 0 & 1 & 0 \\ \\ 1 & 0 & 0 \\ \\ 0 & 0 & A_H \mathrm{Var}_1^H(\eta) \end{pmatrix}. \tag{1.58}$$

Let $G^H = (\Sigma^{-1} + A_H C)^{-1}$, then

$$\begin{aligned} \mathrm{E}_0^H[J_1^H] = \quad & -\frac{|G|}{|\Sigma|} \exp\Big\{ - A_H \big\{ (\mathrm{E}_0^H[P] - F) N_H y_H + \frac{1}{2} \bar{u}^\top C \bar{u} - \frac{1}{2} A_H [G_{11}(N_H y_H + \mathrm{E}_0^H[X])^2 \\ & + 2G_{12}(N_H y_H + \mathrm{E}_0^H[X]) \mathrm{E}_0^H[P] + 2G_{13}(N_H y_H + \mathrm{E}_0^H[X]) A_H \mathrm{Var}_1^H(\eta) \mathrm{E}_0^H[N_H x_H + X] \\ & + G_{22}(\mathrm{E}_0^H[P])^2 + G_{33}(A_H \mathrm{Var}_1^H(\eta) \mathrm{E}_0^H[N_H x_H + X])^2 \\ & + 2G_{23} A_H \mathrm{Var}_1^H(\eta) \mathrm{E}_0^H[N_H x_H + X] \mathrm{E}_0^H[P]] \big\} \Big\}. \end{aligned} \tag{1.59}$$

The first order condition with respect to y_H gives

$$y_H = \frac{\mathrm{E}_0^H[P] - F}{N_H A_H G_{11}^H} - \frac{G_{12}^H}{N_H G_{11}^H} \mathrm{E}_0^H[P] - \frac{G_{13}^H}{G_{11}^H} A_H \mathrm{Var}_1^H(\eta) \mathrm{E}_0^H[x_H + \frac{X}{N_H}] - \frac{1}{N_H} \mathrm{E}_0^H[X], \tag{1.60}$$

and the second order condition is $A_H G_{11}^H > 0$ which is satisfied automatically.

For market makers, we plug Eq. (1.56) into Eq. (1.55) to obtain

$$J_1^M \quad = \quad -e^{-N_M A_M [(P-F) y_M + \frac{1}{2} N_M A_M \mathrm{Var}_1^M (V+\eta) x_M^2]}. \tag{1.61}$$

To compute the expectation $\mathrm{E}_0^M[J_1^M]$, I use Lemma 7 and set $\rho = A_M$, $A = -F N_M y_M$, $\bar{u} = \mathrm{E}_0^M[u]$, $\Sigma = \mathrm{Cov}_0^M(u)$, and

$$u = \begin{pmatrix} P \\ \\ N_M x_M \end{pmatrix}, \quad B = \begin{pmatrix} N_M y_M \\ \\ 0 \end{pmatrix}, \quad C = \begin{pmatrix} 0 & 0 \\ \\ 0 & A_M \mathrm{Var}_1^M(V+\eta) \end{pmatrix}. \tag{1.62}$$

Let $G^M = (\Sigma^{-1} + A_M C)^{-1}$, then

$$\begin{aligned} \mathrm{E}_0^M[J_1^M] = -\frac{|G|}{|\Sigma|} \exp\Big\{ - A_M \big\{ - F y_M + \bar{u}_1 N_M y_M + \frac{1}{2} A_M \mathrm{Var}_1^M(V+\eta) \bar{u}_2^2 \\ -\frac{1}{2} A_M [G_{11} N_M^2 y_M^2 + 2G_{12} A_M \mathrm{Var}_1^M(V+\eta) \bar{u}_2 N_M y_M + G_{22}(A_M \mathrm{Var}_1^M(V+\eta) \bar{u}_2)^2] \big\} \Big\}. \end{aligned} \tag{1.63}$$

The first order condition with respect to y_M gives

$$y_M = \frac{\mathrm{E}_0^M[P] - F}{N_M A_M G_{11}^M} - \frac{G_{12}^M}{G_{11}^M} A_M \mathrm{Var}_1^M(V+\eta)\mathrm{E}_0^M[x_M], \tag{1.64}$$

and the second order condition requires that $A_M G_{11}^M > 0$ which is satisfied automatically.

We assume that the total price impact of speculators is $\lambda = \frac{dF}{d(N_S y_S)}$. Since $N_S << N_{H,M}$, each speculator is able to exert a finite market power λ/N_S. Thus, each speculator's demand of futures is

$$y_S = \frac{\mathrm{E}_0^S[P] - F - A_S \mathrm{Cov}_0^S(P,V)z}{N_S A_S \mathrm{Var}_0^S(P) + \lambda/N_S}. \tag{1.65}$$

The market clearing conditions give that

$$\mathrm{E}_1^H[V+\eta] - A_H \mathrm{Var}_1^H(V+\eta)X = \mathbb{L}(P,F), \tag{1.66}$$

$$\mathrm{E}_0^S[P] - A_S \mathrm{Cov}_0^S(P,V)z = \mathbb{L}(F). \tag{1.67}$$

This is because under CARA-Normal setup $\mathrm{E}_1^M[.]$ is a linear function of market makers' information set at $t=1$: $\mathcal{F}_1^M = \{P,F\}$. Similarly, $\mathrm{E}_0^i[.]$ is a linear function of i's ($i \in \{H,M\}$) information set at $t=0$: $\mathcal{F}_0^i = \{F\}$. Hence, the information revealed by futures price is unaffected by N_S. In other words, the competitiveness among speculators does not affect their information revealing[23]. Therefore, the dependency of spot prices on the state variable s_F, which is defined in Proposition 1, does not change with the number of speculators N_S.

By market clearing condition in the spot market, the equilibrium spot price has the

[23]This is because there is no noise trading in this model. If there is noise trading, then the information revealing will increase by a factor of $\frac{N_S}{N_S+1}$ which increases with the population of speculators (44). Nonetheless, adding noise trading in my model does not alter the main results.

form

$$P = \bar{V} - \mu\bar{X} + a(s_P - \bar{V}) + b(s_F - \bar{V}), \tag{1.68}$$

where $\mu = A_H\tau_\eta^{-1}\omega$, $a = \omega + (1-\omega)\frac{\tau_P}{\tau_V+\tau_F+\tau_P}$, $b = (1-\omega)\frac{\tau_F}{\tau_V+\tau_F+\tau_P}$, and $\omega = (1 + \frac{A_H}{A_M}\frac{\tau_V+\tau_F+\tau_P}{\tau_\eta+\tau_V+\tau_F+\tau_P})^{-1}$. It is easy to see that

$$\omega \geq \frac{1}{1+\frac{A_H}{A_M}}, \quad 0 < a < 1, \quad 0 < b < 1, \quad a + b < 1. \tag{1.69}$$

which means the sensitivity of spot price to commercial hedgers' private information is reduced because of the information brought by the speculators in futures market.

Thus, the equilibrium demands in spot market are

$$\begin{aligned} x_H &= \frac{\omega(1-\omega)}{\mu}[\frac{\tau_V+\tau_F}{\tau_V+\tau_F+\tau_P}(s_P - \bar{V}) - \frac{\tau_F}{\tau_V+\tau_F+\tau_P}(s_F - \bar{V})] - (1-\omega)\bar{X}, \\ x_M &= -\frac{\omega(1-\omega)}{\mu}[\frac{\tau_V+\tau_F}{\tau_V+\tau_F+\tau_P}(s_P - \bar{V}) - \frac{\tau_F}{\tau_V+\tau_F+\tau_P}(s_F - \bar{V})] + (1-\omega)\bar{X}, \end{aligned}$$

which are not affected by the number of speculators (N_S) in futures market.

The equilibrium demand schedules in futures market are

$$\begin{aligned} y_H &= \frac{d_H(s_F - \bar{V}) - F}{N_H A_H G_{11}^H} + h_H, \\ y_M &= \frac{d_M(s_F - \bar{V}) - F}{N_M A_M G_{11}^M} + h_M, \\ y_S &= \frac{d_S(s_F - \bar{V}) - F}{N_S A_S \mathrm{Var}_0^S(P) + \lambda/N_S} + h_S, \end{aligned}$$

where

$$
\begin{aligned}
h_H &= \frac{1 - A_H G_{12}^H}{N_H A_H G_{11}^H}(\bar{V} - \mu \bar{X}) - \mu \frac{G_{13}^H}{G_{11}^H}\frac{\bar{X}}{N_H} - \frac{\bar{X}}{N_H}, \\
h_M &= \frac{1}{N_M A_M G_{11}^M}(\bar{V} - \mu \bar{X}) - \mu \frac{G_{12}^M}{G_{11}^M}\frac{\bar{X}}{N_M}, \\
h_S &= \frac{1}{N_S A_S \mathrm{Var}_0^S(P) + \lambda / N_S}(\bar{V} - \mu \bar{X}), \\
d_H &= (1 - A_H G_{12}^H)\frac{\tau_F}{\tau_V + \tau_F}, \\
d_M &= \frac{\tau_F}{\tau_V + \tau_F}, \\
d_S &= \frac{\tau_\epsilon}{\tau_V + \tau_\epsilon} a + b.
\end{aligned}
$$

As $G_{12}^H < 0$, we can see that $0 < d_M < d_S$ and $0 < d_M < d_H$, and whether d_H is larger or smaller than d_S depends on G_{12}^H. In equilibrium, the market clearing condition is $N_H y_H + N_M y_M + \sum_{n=1}^{N_S} y_{S,n} = 0$, which gives the futures price

$$
F = d(s_F - \bar{V}) + h,
$$

where

$$
d = \omega_H d_H + \omega_M d_M + \omega_S d_S,
$$

$$
h = \mu_1 (N_H h_H + N_M h_M + N_S h_S),
$$

and

$$
\begin{aligned}
\mu_1 &= \left(\frac{1}{A_H G_{11}^H} + \frac{1}{A_M G_{11}^M} + \frac{N_S}{N_S A_S \mathrm{Var}_0^S(P) + \lambda / N_S}\right)^{-1}, \\
\omega_H &= \frac{\mu_1}{A_H G_{11}^H}, \\
\omega_M &= \frac{\mu_1}{A_M G_{11}^M}, \\
\omega_S &= 1 - \omega_H - \omega_M.
\end{aligned}
$$

Since d is a weighted average of d_H, d_M and d_S, then we have $d > d_M$ which means the market makers' demand in the futures market is always downward sloping. However, both the hedgers' and the speculators' demands can be upward or downward sloping. We also define

$$\bar{d} = \mu_0 \left(\frac{d_H}{A_H G_{11}^H} + \frac{d_M}{A_M G_{11}^M} \right), \tag{1.70}$$

which is a weighted average of d_H and d_M, and where $\mu_0 = (\frac{1}{A_H G_{11}^H} + \frac{1}{A_M G_{11}^M})^{-1}$. In equilibrium,

$$\lambda = \frac{\mu_0 d_S + A_S \mathrm{Var}_0^S(P) \bar{d}}{d_S - \bar{d} - \bar{d}/N_S^2}, \tag{1.71}$$

which needs to satisfy the second order condition that

$$\lambda > -\frac{1}{2} N_S^2 A_S \mathrm{Var}_0^S(P). \tag{1.72}$$

Thus, the equilibrium exists if and only if

$$\frac{d_S}{\bar{d}} > 1 + \frac{1}{N_S^2} \quad \text{or} \quad \frac{d_S}{\bar{d}} < \frac{N_S^2 - 1}{N_S^2 + \frac{2\mu_0}{A_S \mathrm{Var}_0^S(P)}}. \tag{1.73}$$

Now plugging λ back in, we obtain the equilibrium futures price and demands. If $d_S > (1 + \frac{1}{N_S^2})\bar{d}$, then d increases with N_S towards d_S and h decreases. If $d_S < \frac{N_S^2 - 1}{N_S^2 + \frac{2\mu_0}{A_S \mathrm{Var}_0^S(P)}} \bar{d}$, then as N_S increases d decreases towards d_S and h decreases. As N_S increases, the existence area on the parameter space (τ_z, τ_X) expands. Note that there is no equilibrium for $d_S = \bar{d}$, no matter how many speculators are present.

1.8.2 Computation of Gs

In this section, I compute the values of Gs.

$$
\begin{aligned}
&|I + A_C C\Sigma| = 1 + \frac{\tau_\eta}{\tau_V + \tau_F}(1-a)^2 - \frac{A_H^2(\tau_\eta + \tau_V + \tau_F)}{\tau_X \tau_\eta(\tau_V + \tau_F)} a(2-a), \\
&G_{11}^H = \frac{1}{|I + A_H C\Sigma|}\left(\frac{A_H^2}{\tau_X \tau_\eta(\tau_V + \tau_F)} + \frac{A_H^2}{\tau_X \tau_\eta^2} + \frac{1}{\tau_V + \tau_F}\right)a^2, \\
&G_{12}^H = -\frac{A_H}{|I + A_H C\Sigma|}\frac{\tau_\eta + \tau_V + \tau_F}{\tau_X \tau_\eta(\tau_V + \tau_F)} a, \\
&G_{13}^H = \frac{\frac{1}{A_H}\frac{\tau_\eta}{\tau_V + \tau_F} a(1-a) - \frac{A_H}{\tau_X}\frac{\tau_\eta + \tau_V + \tau_F}{\tau_\eta(\tau_V + \tau_F)} a^2}{|I + A_H C\Sigma|}, \\
&\mathrm{Var}_0^S(P) = a^2\left(\frac{1}{\tau_V + \tau_\epsilon} + \frac{A_H^2}{\tau_\eta^2 \tau_X}\right),
\end{aligned}
$$

the second order condition requires that $G_{11}^H > 0$, i.e., $|I + A_H C\Sigma^C| > 0$, otherwise the commercial hedgers will not do anticipatory hedging in the futures market. This requires that

$$
\left(\frac{A_H^2}{\tau_X \tau_\eta} + \frac{\tau_\eta}{\tau_\eta + \tau_V + \tau_F}\right)a(2-a) < 1. \tag{1.74}
$$

As $0 < a(2-a) < 1$, the sufficient condition for the above condition is

$$
\tau_X > A_H^2\left(\frac{1}{\tau_\eta} + \frac{1}{\tau_V + \tau_F}\right) = A_H^2 \mathrm{Var}(V + \eta | F) \Leftrightarrow A_H^2 \mathrm{Var}(V + \eta | F)\mathrm{Var}(X) < 1,
$$

this condition comes from the argument that under normality the endowment value $(V - \bar{V})X$ may take large negative values to generate an infinitely negative expected utility for commercial hedgers. Therefore, to ensure that the commercial hedgers' utility is finite, the sufficient condition is

$$
A_H^2 \mathrm{Var}(V + \eta)\mathrm{Var}(X) = \frac{A_H^2}{\tau_X}\left(\frac{1}{\tau_V} + \frac{1}{\tau_\eta}\right) < 1, \tag{1.75}
$$

which is guaranteed by Equation (1.14). And

$$G_{11}^M = \frac{\Sigma_{11}^M + A_M^2 \mathrm{Var}_1^M(V+\eta)[\Sigma_{11}^M\Sigma_{22}^M - (\Sigma_{12}^M)^2]}{A_M^2 \mathrm{Var}_1^M(V+\eta)\Sigma_{22}^M + 1} > 0,$$

$$G_{12}^M = \frac{\Sigma_{12}^M}{A_M^2 \mathrm{Var}_1^M(V+\eta)\Sigma_{22}^M + 1} = \frac{-\frac{\omega(1-\omega)}{\mu}\frac{\tau_V+\tau_F}{\tau_V+\tau_F+\tau_P}(\frac{A_H^2}{\tau_X\tau_\eta^2} + \frac{1}{\tau_V+\tau_F})a}{1 + A_M\frac{1-\omega}{\mu}\omega^2(\frac{\tau_V+\tau_F}{\tau_V+\tau_F+\tau_P})^2(\frac{A_H^2}{\tau_X\tau_\eta^2} + \frac{1}{\tau_V+\tau_F})},$$

$$G_{22}^M = \frac{\Sigma_{22}^M}{A_M^2 \mathrm{Var}_1^M(V+\eta)\Sigma_{22}^M + 1} = \frac{(\frac{\omega(1-\omega)}{\mu}\frac{\tau_V+\tau_F}{\tau_V+\tau_F+\tau_P})^2(\frac{A_H^2}{\tau_X\tau_\eta^2} + \frac{1}{\tau_V+\tau_F})}{1 + A_M\frac{1-\omega}{\mu}\omega^2(\frac{\tau_V+\tau_F}{\tau_V+\tau_F+\tau_P})^2(\frac{A_H^2}{\tau_X\tau_\eta^2} + \frac{1}{\tau_V+\tau_F})}.$$

The second order condition requires that $G_{11}^M > 0$, which is satisfied by Equation (1.15). There is a relation between G^Hs and G^Ms:

$$\frac{G_{12}^H - G_{13}^H}{G_{11}^H} = \frac{G_{12}^M}{G_{11}^M}. \tag{1.76}$$

1.8.3 Market Condition for $d_S = \bar{d}$

In the parameter space (τ_z, τ_X), the boundary function $f(\tau_X)$ that satisfies $d_S = \bar{d}$ can be determined by solving the equation

$$\frac{A_M + \alpha A_H}{A_M}\tau_X + \frac{\alpha A_H^3}{A_M\tau_\eta}\frac{\tau_X}{\tau_X + \frac{A_H^2}{\tau_\eta^2}(\tau_V+\tau_F)} = A_H^2\frac{(\frac{1}{\tau_\epsilon} + \frac{1}{\tau_\eta} + \frac{\tau_V}{\tau_\epsilon\tau_\eta})\tau_z + \frac{A_S^2}{\beta^2\tau_\epsilon^2}(1+\frac{\tau_V}{\tau_\eta})}{\frac{A_S^2\tau_V}{\beta^2\tau_\epsilon^2}}.$$

One can check that $f(\tau_X)$ is a monotone increasing convex function. Thus, f is invertible, and $f^{-1}(\tau_z)$ is a monotone increasing concave function. The asymptotic formula of f^{-1} is derived and given by

$$f^{-1}(\tau_z) \approx \frac{A_M}{A_M + \alpha A_H}[A_H^2\frac{(\frac{1}{\tau_\epsilon} + \frac{1}{\tau_\eta} + \frac{\tau_V}{\tau_\epsilon\tau_\eta})\tau_z + \frac{A_S^2}{\beta^2\tau_\epsilon^2}(1+\frac{\tau_V}{\tau_\eta})}{\frac{A_S^2\tau_V}{\beta^2\tau_\epsilon^2}} - \frac{\alpha A_c^3 H}{A_M\tau_\eta}],$$

and for small τ_X we have

$$f^{-1}(\tau_z) \approx \frac{\beta^2 A_H^2\tau_\epsilon^2}{A_S^2\tau_V}\frac{(\frac{1}{\tau_\epsilon} + \frac{1}{\tau_\eta} + \frac{\tau_V}{\tau_\epsilon\tau_\eta})\tau_z + \frac{A_S^2}{\beta^2\tau_\epsilon^2}(1+\frac{\tau_V}{\tau_\eta})}{\frac{A_M+\alpha A_H}{A_M} + \frac{\alpha A_H}{A_M}\frac{\frac{\tau_\eta}{\tau_\epsilon} + \frac{A_S^2}{\beta^2\tau_\epsilon^2}\frac{\tau_\eta}{\tau_z}}{1+\frac{\tau_V}{\tau_\epsilon} + \frac{A_S^2}{\beta^2\tau_\epsilon^2}\frac{\tau_V}{\tau_z}}},$$

and for large τ_X,

$$f^{-1}(\tau_z) \approx \frac{A_M}{A_M + \alpha A_H}[A_H^2 \frac{(\frac{1}{\tau_\epsilon} + \frac{1}{\tau_\eta} + \frac{\tau_V}{\tau_\epsilon \tau_\eta})\tau_z + \frac{A_s^2}{\beta^2 \tau_\epsilon^2}(1 + \frac{\tau_V}{\tau_\eta})}{\frac{A_S^2 \tau_V}{\beta^2 \tau_\epsilon^2}} - \frac{\alpha A_H^3}{A_M \tau_\eta}].$$

1.8.4 Proof of Proposition 4

Proof. The market makers' expectation $E_1^M[V] = E[V|F,P]$ and variance $Var_1^M(V) = Var(V|F,P)$ at $t = 1$ are:

$$\mathrm{E}[V - \bar{V}|F,P] = \begin{cases} [\frac{\tau_P}{\tau_V + \tau_F + \tau_P}(s_P - \bar{V}) + \frac{\tau_F}{\tau_V + \tau_F + \tau_P}(s_F - \bar{V}), & \text{if } F < \overline{F}, \\ [\frac{1}{\tau_V + \tau_F + \tau_P} + q_2(\overline{s}_F)]\tau_P(s_P - \bar{V}) + q_1(\overline{s}_F), & \text{if } F = \overline{F}, \end{cases}$$

$$\mathrm{Var}[V|F,P] = \begin{cases} \frac{1}{\tau_V + \tau_F + \tau_P}, & \text{if } F < \overline{F}, \\ \frac{1}{\tau_V + \tau_F + \tau_P} + q_2(\overline{s}_F) > \frac{1}{\tau_V + \tau_F + \tau_P}, & \text{if } F = \overline{F}, \end{cases}$$

where

$$q_1(\overline{s}_F) = \frac{\tau_F \sigma_F}{\tau_V + \tau_P + \tau_F} \frac{\phi(\frac{\overline{s}_F - \bar{V}}{\sigma_F})}{1 - \Phi(\frac{\overline{s}_F - \bar{V}}{\sigma_F})} > \frac{\tau_F}{\tau_V + \tau_P + \tau_F}(\overline{s}_F - \bar{V}), \tag{1.77}$$

$$q_2(\overline{s}_F) = (\frac{\tau_F \sigma_F}{\tau_V + \tau_P + \tau_F})^2 [1 + \frac{\overline{s}_F - \bar{V}}{\sigma_F} \frac{\phi(\frac{\overline{s}_F - \bar{V}}{\sigma_F})}{1 - \Phi(\frac{\overline{s}_F - \bar{V}}{\sigma_F})} - \Big(\frac{\phi(\frac{\overline{s}_F - \bar{V}}{\sigma_F})}{1 - \Phi(\frac{\overline{s}_F - \bar{V}}{\sigma_F})}\Big)^2], \tag{1.78}$$

where $\sigma_F^2 = \tau_V^{-1} + \tau_F^{-1}$, and ϕ and Φ denote the standard normal PDF and CDF respectively.

To compute $\mathrm{E}[V - \bar{V}|F]$ and $\mathrm{Var}(V|F)$, one can simply take the limit $\tau_P \to 0$:

$$\mathrm{E}[V - \bar{V}|F] = \begin{cases} \frac{\tau_F}{\tau_V + \tau_F}(s_F - \bar{V}), & \text{if } F < \overline{F}, \\ q_1'(\overline{s}_F), & \text{if } F = \overline{F}, \end{cases}$$

$$\mathrm{Var}[V|F] = \begin{cases} \frac{1}{\tau_V+\tau_F}, & \text{if } F < \overline{F}, \\ \\ \frac{1}{\tau_V+\tau_F} + q_2'(\overline{s}_F) > \frac{1}{\tau_V+\tau_F}, & \text{if } F = \overline{F}, \end{cases}$$

where

$$q_1'(\overline{s}_F) = \frac{\tau_F \sigma_F}{\tau_V+\tau_F} \frac{\phi(\frac{\overline{s}_F-\bar{V}}{\sigma_F})}{1-\Phi(\frac{\overline{s}_F-\bar{V}}{\sigma_F})} > \frac{\tau_F}{\tau_V+\tau_P+\tau_F}(\overline{s}_F-\bar{V}), \tag{1.79}$$

$$q_2'(\overline{s}_F) = (\frac{\tau_F \sigma_F}{\tau_V+\tau_F})^2 [1 + \frac{\overline{s}_F-\bar{V}}{\sigma_F} \frac{\phi(\frac{\overline{s}_F-\bar{V}}{\sigma_F})}{1-\Phi(\frac{\overline{s}_F-\bar{V}}{\sigma_F})} - \big(\frac{\phi(\frac{\overline{s}_F-\bar{V}}{\sigma_F})}{1-\Phi(\frac{\overline{s}_F-\bar{V}}{\sigma_F})}\big)^2]. \tag{1.80}$$

See 35 for more mathematical details.

If speculators stop revealing their state by submitting market orders beyond $\overline{s}_F$, then for $s_F \geq \overline{s}_F$

$$\overline{P} = \overline{a}(s_P - \bar{V}) + \overline{b} + \bar{V} - \overline{\mu}\bar{X}, \tag{1.81}$$

where $\overline{\omega} = \frac{\tau_\eta}{A_H}\overline{\mu}$, and

$$\overline{\mu} = (\frac{1}{A_H \tau_\eta^{-1}} + \frac{\alpha}{A_M} \frac{1}{\frac{1}{\tau_V+\tau_P+\tau_F} + q_2(\overline{s}_F) + \tau_\eta^{-1}})^{-1} > \mu, \tag{1.82}$$

$$\overline{a} = \overline{\omega} + (1-\overline{\omega}) \frac{\tau_P}{\tau_V+\tau_P+\tau_F}(1+q_2(\overline{s}_F)) > a, \tag{1.83}$$

$$\overline{b} = (1-\overline{\omega}) q_1(\overline{s}_F) > b(\overline{s}_F - \bar{V}). \tag{1.84}$$

The equilibrium spot demands for $s_F \geq \overline{s}_F$ are

$$\overline{x}_H = \frac{\tau_\eta}{A_H}[V - \bar{V} - \overline{a}(s_P - \bar{V}) - \overline{b}] - (1-\overline{\omega})\bar{X}, \tag{1.85}$$

$$\overline{x}_M = -\frac{\tau_\eta}{A_H}[V - \bar{V} - \overline{a}(s_P - \bar{V}) - \overline{b}] + (1-\overline{\omega})\bar{X}. \tag{1.86}$$

The average trading quantity $(1-\overline{\omega})\bar{X}$ is reduced when limit binds. In futures market, the equilibrium price for $s_F \geq \overline{s}_F$ is

$$\overline{F} = \overline{c} + \overline{c_0}, \tag{1.87}$$

where

$$\overline{c} = (1 - \frac{\overline{G}_{11}^H}{\overline{G}_{12}^H}\overline{\mu_0})(\overline{a}q_1'(\overline{s}_F) + \overline{b}), \tag{1.88}$$

$$\overline{c_0} = (1 - \frac{\overline{G}_{11}^H}{\overline{G}_{12}^H}\overline{\mu_0})(\bar{V} - \overline{\mu}\bar{X}) - (1 + \frac{\overline{G}_{13}^H}{\overline{G}_{11}^H}\overline{\mu} + \frac{\overline{G}_{12}^M}{\overline{G}_{11}^M}\alpha\overline{\mu})\overline{\mu_0}\bar{X} + \overline{\mu_0}L, \tag{1.89}$$

where $\overline{\mu_0} = (\frac{1}{A_H\overline{G}_{11}^H} + \frac{\alpha}{A_M\overline{G}_{11}^M})^{-1}$ and all the $\overline{G}$s are functions of $\overline{s}_F$. Thus the equilibrium futures demands for $s_F \geq \overline{s}_F$ are

$$\overline{y}_H = \frac{\mathrm{E}_0^H[\overline{P}] - \overline{F}}{A_H\overline{G}_{11}^H} - \frac{\overline{G}_{12}^H}{\overline{G}_{11}^H}\mathrm{E}_0^H[\overline{P}] - \frac{\overline{G}_{13}^H}{\overline{G}_{11}^H}A_H\mathrm{Var}_1^H(\eta)\mathrm{E}_0^H[\overline{x}_H + X] - \mathrm{E}_0^H[X], \tag{1.90}$$

$$\overline{y}_M = \frac{\mathrm{E}_0^M[\overline{P}] - \overline{F}}{A_M\overline{G}_{11}^M} - \frac{\overline{G}_{12}^M}{\overline{G}_{11}^M}A_M\mathrm{Var}_1^M(V + \eta)\mathrm{E}_0^M[\overline{x}_M], \tag{1.91}$$

$$\overline{y}_S = \frac{\mathrm{E}_0^S[\overline{P}] - \overline{F} - A_S\mathrm{Cov}_0^S(\overline{P}, V)z/\beta}{A_S\mathrm{Var}_0^S(\overline{P})}. \tag{1.92}$$

Now I compute all the $\overline{G}$s. Fist I need to compute $\overline{\Sigma}$s, since $\overline{G}^H = (\overline{\Sigma}^H + A_H C)^{-1}$ where $C = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & A_C\tau_\eta^{-1} \end{pmatrix}$, all the elements of $\overline{G}^H$ can be calculated by

$$\begin{aligned}
\mathrm{Var}_0^H(V) &= \frac{1}{\tau_V + \tau_F} + q_2(\overline{s}_F), \\
\mathrm{Var}_0^H(\overline{P}) &= \overline{a}^2(\frac{1}{\tau_V + \tau_F} + q_2(\overline{s}_F) + \frac{A_H^2\tau_\eta^{-2}}{\tau_X}), \\
\mathrm{Cov}_0^H(V, \overline{P}) &= \overline{a}(\frac{1}{\tau_V + \tau_F} + q_2(\overline{s}_F)), \\
\mathrm{Cov}_0^H(X, \overline{P}) &= -\overline{a}\frac{A_H\tau_\eta^{-1}}{\tau_X}.
\end{aligned}$$

Similarly,

$$\overline{\Sigma}_{11}^M = \mathrm{Var}_0^M(\overline{P}), \quad \overline{\Sigma}_{22}^M = (\frac{\overline{\omega}(1 - \overline{\omega})}{\overline{\mu}}\frac{\tau_V + \tau_F}{\tau_V + \tau_F + \tau_P})^2(\frac{1}{\tau_V + \tau_F} + q_2(\overline{s}_F) + \frac{A_H^2\tau_\eta^{-2}}{\tau_X}),$$

$$\overline{\Sigma}_{12}^M = -\frac{\overline{\omega}(1 - \overline{\omega})}{\overline{\mu}}\frac{\tau_V + \tau_F}{\tau_V + \tau_F + \tau_P}\overline{a}(\frac{1}{\tau_V + \tau_F} + q_2(\overline{s}_F) + \frac{A_H^2\tau_\eta^{-2}}{\tau_X}).$$

Since $\overline{G}^M = (\overline{\Sigma}^M + A_M\overline{C})^{-1}$ where $\overline{C} = \begin{pmatrix} 0 & 0 \\ 0 & \frac{\mu}{1-\overline{\omega}} \end{pmatrix}$, all elements of $\overline{G}^M$ can be computed and $\overline{G}^M_{11} > G^M_{11}$, i.e., the perceived riskiness by market makers increases when limits bind.

$$\begin{aligned} |I + A_H C\overline{\Sigma}| &= 1 + \frac{\tau_\eta}{\tau_V + \tau_F}(1-\overline{a})^2 - \frac{A_H^2(\tau_\eta + \tau_V + \tau_F)}{\tau_X\tau_\eta(\tau_V + \tau_F)}\overline{a}(2-\overline{a}) \\ &+ \tau_\eta(1-\overline{a})^2 q_2 - \frac{A_H^2}{\tau_X}\overline{a}(2-\overline{a})q_2, \\ \overline{G}^H_{11} &= \frac{\overline{a}^2}{|I + A_H C\overline{\Sigma}|}[\frac{A_H^2}{\tau_X\tau_\eta(\tau_V+\tau_F)} + \frac{A_H^2}{\tau_X\tau_\eta^2} + \frac{1}{\tau_V+\tau_F} + (1 + \frac{A_H^2}{\tau_X\tau_\eta})q_2]. \end{aligned}$$

Note that $\overline{G}^H_{11}$ depends on $\overline{s}_F$, and $\overline{G}^H_{11} > G^H_{11}$, i.e., the perceived riskiness by hedgers increases when limits bind,

$$\begin{aligned} & q_2 > 0, \quad \lim_{\overline{s}_F \to \infty} q_2 = 0, \quad \lim_{\overline{s}_F \to -\infty} q_2 = \frac{\tau_F}{\tau_V(\tau_V + \tau_F)}, \\ & \overline{a} > a, \quad \lim_{\overline{s}_F \to \infty} \overline{a} = a, \\ & |I + A_H C\overline{\Sigma}| < |I + A_H C\Sigma|, \quad \lim_{\overline{s}_F \to \infty} = |I + A_H C\Sigma|. \end{aligned}$$

□

1.8.5 Equilibrium with Speculative Position Limit

If speculators stop revealing their state by submitting market orders beyond $\overline{s}_F$, then for $s_F \geq \overline{s}_F$

$$\overline{P} = \overline{a}(s_P - \bar{V}) + \bar{b} + \bar{V} - \overline{\mu}\bar{X}, \tag{1.93}$$

where $\overline{\omega} = \frac{\tau_\eta}{A_H}\overline{\mu}$, and

$$\overline{\mu} = (\frac{1}{A_H\tau_\eta^{-1}} + \frac{1}{A_M}\frac{1}{\frac{1}{\tau_V+\tau_P+\tau_F} + q_2(\overline{s}_F) + \tau_\eta^{-1}})^{-1} > \mu, \tag{1.94}$$

$$\overline{a} = \overline{\omega} + (1-\overline{\omega})\frac{\tau_P}{\tau_V + \tau_P + \tau_F}(1 + q_2(\overline{s}_F)) > a, \tag{1.95}$$

$$\bar{b} = (1-\overline{\omega})q_1(\overline{s}_F) > b(\overline{s}_F - \bar{V}). \tag{1.96}$$

The equilibrium spot demands for $s_F \geq \overline{s}_F$ are

$$\overline{x}_H = \frac{\tau_\eta}{A_H}[V - \bar{V} - \overline{a}(s_P - \bar{V}) - \overline{b}] - (1 - \overline{\omega})\bar{X}, \tag{1.97}$$

$$\overline{x}_M = -\frac{\tau_\eta}{\alpha A_H}[V - \bar{V} - \overline{a}(s_P - \bar{V}) - \overline{b}] + \frac{1 - \overline{\omega}}{\alpha}\bar{X}. \tag{1.98}$$

In futures market, the equilibrium price for $s_F \geq \overline{s}_F$ is

$$\overline{F} = \overline{c} + \overline{c}_0, \tag{1.99}$$

where

$$\overline{c} = (1 - \frac{\overline{G}_{11}^H}{\overline{G}_{12}^H}\overline{\mu}_0)(\overline{a}q_1'(\overline{s}_F) + \overline{b}), \tag{1.100}$$

$$\overline{c}_0 = (1 - \frac{\overline{G}_{11}^H}{\overline{G}_{12}^H}\overline{\mu}_0)(\bar{V} - \overline{\mu}\bar{X}) - (1 + \frac{\overline{G}_{13}^H}{\overline{G}_{11}^H}\overline{\mu} + \frac{\overline{G}_{12}^M}{\overline{G}_{11}^M}\alpha\overline{\mu})\overline{\mu}_0\bar{X} + \overline{\mu}_0 L, \tag{1.101}$$

where $\overline{\mu}_0 = (\frac{1}{A_H\overline{G}_{11}^H} + \frac{\alpha}{A_M\overline{G}_{11}^M})^{-1}$ and all the $\overline{G}$s are functions of $\overline{s}_F$. Thus the equilibrium futures demands for $s_F \geq \overline{s}_F$ are

$$\overline{y}_H = \frac{\mathrm{E}_0^H[\overline{P}] - \overline{F}}{A_H\overline{G}_{11}^H} - \frac{\overline{G}_{12}^H}{\overline{G}_{11}^H}\mathrm{E}_0^H[\overline{P}] - \frac{\overline{G}_{13}^H}{\overline{G}_{11}^H}A_H\mathrm{Var}_1^H(\eta)\mathrm{E}_0^H[\overline{x}_H + X] - \mathrm{E}_0^H[X], \tag{1.102}$$

$$\overline{y}_M = \frac{\mathrm{E}_0^M[\overline{P}] - \overline{F}}{A_M\overline{G}_{11}^M} - \frac{\overline{G}_{12}^M}{\overline{G}_{11}^M}A_M\mathrm{Var}_1^M(V + \eta)\mathrm{E}_0^M[\overline{x}_M], \tag{1.103}$$

$$\overline{y}_S = \frac{\mathrm{E}_0^S[\overline{P}] - \overline{F} - A_S\mathrm{Cov}_0^S(\overline{P}, V)z}{N_S A_S\mathrm{Var}_0^S(\overline{P})} \geq L. \tag{1.104}$$

When speculators are stabilizing price, $d_S < d$ thus if speculators keep revealing their state, then position limit L binds for $s_F \leq \hat{s}_F$ where

$$\hat{s}_F = \bar{V} + \frac{N_S A_S\mathrm{Var}_0^S(P) + \lambda/N_S}{d_S - d}[L + \frac{h}{N_S A_S\mathrm{Var}_0^S(P) + \lambda/N_S} - h_S]. \tag{1.105}$$

Hence for $s_F \leq \hat{s}_F$, in equilibrium the market clearing condition is $N_H y_H + N_M y_M + N_S L = 0$. Keep in mind that s_F is still being fully revealed although L binds if speculators submit limit orders $\hat{y}_S$ for $s_F \leq \hat{s}_F$:

$$\hat{y}_F = \frac{\mathrm{E}_0^S[P] - \hat{F} - A_S \mathrm{Cov}_0^S(P,V)z}{N_S A_S \mathrm{Var}_0^S(P) + \hat{\lambda}/N_S}, \tag{1.106}$$

where $\hat{\lambda} > \lambda$, which means speculators provide less liquidity although they still reveal information.

On the right tail of s_F, position limit binds at $\overline{s}_F$ if speculators stop revealing their state, and $\overline{s}_F$ is determined by

$$y_S(\overline{s}_F) = \frac{\mathrm{E}_0^S[\overline{P}] - \overline{F} - A_S \mathrm{Cov}_0^S(\overline{P},V)z}{A_S \mathrm{Var}_0^S(\overline{P})} = N_S L. \tag{1.107}$$

As L becomes smaller, $\hat{s}_F$ increases and $\overline{s}_F$ decreases. Thus, when L is very tight, $\overline{s}_F$ can be smaller than $\hat{s}_F$. In such case, speculators' position is always at L: for $s_F \geq \overline{s}_F$ they submit market order without revealing their state and for $s_F < \overline{s}_F$ they submit limit order to reveal their state. Generally, when L is not too tight, $\hat{s}_F < \overline{s}_F$, and the equilibrium prices are

$$F_L = \begin{cases} \hat{F}, & \text{if } s_F \leq \hat{s}_F \\ F, & \text{if } \hat{s}_F < s_F < \overline{s}_F \\ \overline{F}, & \text{if } s_F \geq \overline{s}_F \end{cases} \qquad P_L = \begin{cases} P, & \text{if } s_F < \overline{s}_F \\ \overline{P}, & \text{if } s_F \geq \overline{s}_F \end{cases} \tag{1.108}$$

where $\hat{F}$ is a linear functions of s_F:

$$\hat{F} = \hat{d}(s_F - \bar{V}) + \hat{h}, \tag{1.109}$$

where

$$\hat{d} = (1 - \frac{G_{12}^H}{G_{11}^H}\mu_0)(\frac{\tau_F}{\tau_V + \tau_F}a + b) > d, \tag{1.110}$$

$$\hat{h} = (1 - \frac{G_{12}^H}{G_{11}^H}\mu_0)\bar{V} - \mu\bar{X} - \mu_0\bar{X} + \mu_0 L. \tag{1.111}$$

The equilibrium futures price is continuous at $\hat{s}_F$ but discontinuous at $\overline{s}_F$ with respect to state s_F. Then the equilibrium demands in futures and spot markets can be computed.

When speculators are destabilizing price and their population is large ($N_S >> 1$), i.e., $d_S > d$, thus if speculators keep revealing their state, then position limit L binds for $s_F \geq \hat{s}_F$. Hence for $s_F \geq \hat{s}_F$, in equilibrium the market clearing condition is $N_H y_H + N_M y_M + N_S L = 0$. Keep in mind that s_F is still being fully revealed although L binds if speculators submit limit orders. Also on the right tail of s_F, position limit binds at $\overline{s}_F$ if speculators stop revealing their state, and $\overline{s}_F$ is still determined by Eqn. (1.107). As L becomes smaller, both $\hat{s}_F$ and $\overline{s}_F$ decrease. Since speculators do not have much market power, $\hat{s}_F < \overline{s}_F$. As L increases, the gap between $\overline{s}_F$ and $\hat{s}_F$ becomes smaller. The equilibrium prices are

$$F_L = \begin{cases} F, & \text{if } s_F < \hat{s}_F \\ \hat{F}, & \text{if } \hat{s}_F \leq s_F < \overline{s}_F \\ \overline{F}, & \text{if } s_F \geq \overline{s}_F \end{cases} \qquad P_L = \begin{cases} P, & \text{if } s_F < \overline{s}_F \\ \overline{P}, & \text{if } s_F \geq \overline{s}_F \end{cases} \tag{1.112}$$

where $\hat{F}$ is a linear functions of s_F:

$$\hat{F} = \hat{d}(s_F - \bar{V}) + \hat{h}, \tag{1.113}$$

where

$$\hat{d} = (1 - \frac{G_{12}^H}{G_{11}^H}\mu_0)(\frac{\tau_F}{\tau_V + \tau_F}a + b) < d, \tag{1.114}$$

$$\hat{h} = (1 - \frac{G_{12}^H}{G_{11}^H}\mu_0)\bar{V} - \mu\bar{X} - \mu_0\bar{X} + \mu_0 L. \tag{1.115}$$

The equilibrium futures price is continuous at $\hat{s}_F$ but discontinuous at $\overline{s}_F$ with respect to state s_F. Then the equilibrium demands in futures and spot markets can be computed.

When speculators are destabilizing price and their population is small, $d_S > d$ thus if speculators choose to reveal their state then the position limit L binds for $s_F \geq \overline{s}_F$, where $\overline{s}_F$ is given by Eqn. (1.107); if speculators choose not to reveal their state then the position limit L binds for $s_F \geq \hat{s}_F$, where $\hat{s}_F$ is given by Eqn. (1.105). Speculators always prefer not to reveal their state. When there is only a few speculators, however, $\hat{s}_F$ can not bind without $\overline{s}_F$ binds i.e., $\hat{s}_F < \overline{s}_F$. As a result, in this situation speculators stop revealing their state for $s_F \geq \hat{s}_F$. Then the equilibrium demands in futures and spot markets can be computed.

Proposition 14 (**Equilibrium Demands in Futures Market with Position Limits**). *With the speculative position limit L, the equilibrium demands of futures are:*

1. *When speculators are stabilizing price, given any L, there exist $\overline{s}_F$ and $\hat{s}_F$ such that: in equilibrium the futures demand schedule of each speculator is y_S^* and the demand is executed at y_S^L,*

$$y_S^* = \begin{cases} \hat{y}_S \geq L, & \text{if } s_F \leq \hat{s}_F \\ y_S < L, & \text{if } \hat{s_F} < s_F < \overline{s}_F \\ L, & \text{if } s_F \geq \overline{s_F} \end{cases} \Rightarrow \quad y_S^L = \min\{y_S^*, L\} = \begin{cases} L, & \text{if } s_F \leq \hat{s}_F \\ y_S, & \text{if } \hat{s_F} < s_F < \overline{s}_F \\ L, & \text{if } s_F \geq \overline{s}_F \end{cases}$$

   The equilibrium futures demands of each market maker and each commercial hedger are y_M^L and y_H^L respectively:

$$y_M^L = \begin{cases} \hat{y}_M, & \text{if } s_F \leq \hat{s}_F \\ y_M, & \text{if } \hat{s}_F < s_F < \overline{s}_F \\ \overline{y}_M, & \text{if } s_F \geq \overline{s}_F \end{cases} \qquad y_H^L = \begin{cases} \hat{y}_H, & \text{if } s_F \leq \hat{s}_F \\ y_H, & \text{if } \hat{s}_F < s_F < \overline{s}_F \\ \overline{y}_H, & \text{if } s_F \geq \overline{s}_F \end{cases}$$

2. *When speculators are destabilizing price and their population is large ($N_S >> 1$), given any L, there exist $\hat{s}_F < \overline{s}_F$ such that: in equilibrium the futures demand schedule of each speculator is y_S^* and the demand is executed at y_S^L,*

$$y_S^* = \begin{cases} y_S < L, & \text{if } s_F < \hat{s}_F \\ \hat{y}_S \geq L, & \text{if } \hat{s}_F \leq s_F < \overline{s}_F \\ L, & \text{if } s_F \geq \overline{s}_F \end{cases} \quad \Rightarrow \quad y_S^L = \min\{y_S^*, L\} = \begin{cases} y_S, & \text{if } s_F < \hat{s}_F \\ L, & \text{if } \hat{s}_F \leq s_F < \overline{s}_F \\ L, & \text{if } s_F \geq \overline{s}_F \end{cases}$$

The equilibrium demands of each market maker and each commercial hedger are y_M^L and y_H^L respectively:

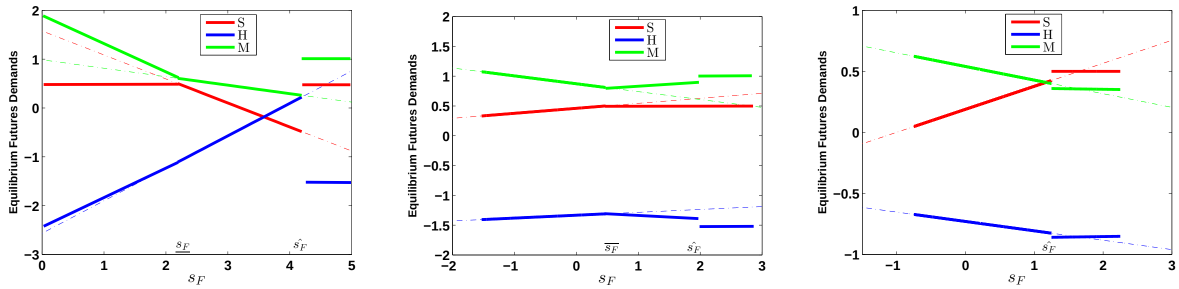
$$y_M^L = \begin{cases} \hat{y}_M, & \text{if } s_F < \hat{s}_F \\ y_M, & \text{if } \hat{s}_F \leq s_F < \overline{s}_F \\ \overline{y}_M, & \text{if } s_F \geq \overline{s}_F \end{cases} \qquad y_H^L = \begin{cases} \hat{y}_H, & \text{if } s_F < \hat{s}_F \\ y_H, & \text{if } \hat{s}_F \leq s_F < \overline{s}_F \\ \overline{y}_H, & \text{if } s_F \geq \overline{s}_F \end{cases}$$

3. *When speculators are destabilizing price and their population is small, given any L, there exists $\overline{s}_F$ such that: in equilibrium speculators collude and each of them submits demand y_S^L in futures market,*

$$y_S^L = \begin{cases} y_S < L, & \text{if } s_F < \overline{s}_F \\ L, & \text{if } s_F \geq \overline{s}_F \end{cases}$$

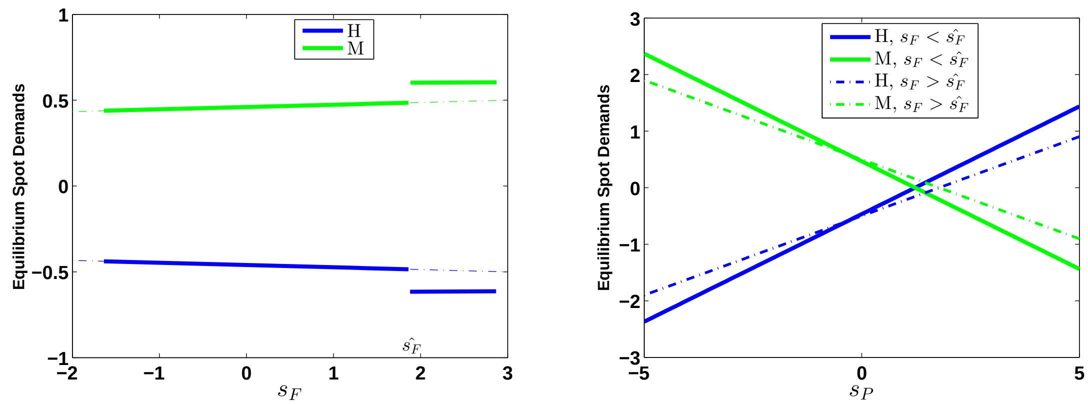
The equilibrium futures demands of each market maker and each commercial hedger are y_M^L and y_H^L respectively:

$$y_M^L = \begin{cases} y_M, & \text{if } s_F < \overline{s}_F \\ \overline{y}_M, & \text{if } s_F \geq \overline{s}_F \end{cases} \qquad y_H^L = \begin{cases} y_H, & \text{if } s_F < \overline{s}_F \\ \overline{y}_H, & \text{if } s_F \geq \overline{s}_F \end{cases}$$



(a) Stabilizing Price (b) Destabilizing Price (Competitive) (c) Destabilizing Price (Duopoly)

Figure 1.11: Equilibrium demands in futures market with position limit under different market conditions, where (a) $d_S = 0.1\bar{d}$ and $N_S = 2$, (b) $d_S = 1.5\bar{d}$ and $N_S = 15$, (c) $d_S = 2.5\bar{d}$ and $N_S = 2$. Other parameters are $\tau_\epsilon/\tau_V = 0.3$, $\tau_\eta = 1$, $N_H = N_M = 30$, $\bar{X} = 1$ and $\bar{V} = 0$.



(a) (b)

Figure 1.12: Equilibrium demands in the spot market with position limit with respect to s_F (a) and s_P (b) with parameters are $\tau_\epsilon/\tau_V = 0.3$, $\tau_\eta = 1$, $N_H = N_M = 30$, $N_S = 15$, $d_S = 1.5\bar{d}$, $\bar{X} = 1$ and $\bar{V} = 0$.

Proposition 15 (**Equilibrium Demands in Spot Market with Position Limits**). *With the speculative position limit, there always exists $\overline{s}_F$ such that the equilibrium commodity* *demands of each market maker and each commercial hedger are x_M^L and x_H^L respectively:*

$$x_M^L = \begin{cases} x_M, & \text{if } s_F < \overline{s}_F \\ \\ \overline{x}_M, & \text{if } s_F \geq \overline{s}_F \end{cases} \qquad x_H^L = \begin{cases} x_H, & \text{if } s_F < \overline{s}_F \\ \\ \overline{x}_H, & \text{if } s_F \geq \overline{s}_F \end{cases}$$

1.8.6 Proof of Proposition 7

Proof. Plug equilibrium prices in, then

$$P - F = a(s_P - \bar{V}) + (b - d)(s_F - \bar{V}) + \bar{V} - \mu\bar{X} - h,$$

$$F = d(s_F - \bar{V}) + h,$$

Since $\text{Cov}(s_F, s_P) = \tau_V^{-1}$ and $\text{Var}(s_F) = \tau_V^{-1} + \tau_F^{-1}$ and $\text{Var}(s_P) = \tau_V^{-1} + \tau_P^{-1}$, thus

$$\gamma = \frac{(\tau_V + \tau_F)/\tau_F}{\sqrt{\text{Var}(P - F)\text{Var}(F)}}(d_M - d) < 0. \tag{1.116}$$

For $s_F \geq \overline{s}_F$, the correlation between F and $P - F$ goes to zero. For $s_F \leq \hat{s}_F$, d gets greater, but so does $\text{Var}(F)$ and $\text{Var}(P)$. □

1.8.7 Welfare

Welfare without Speculative Position Limit

First, let's consider the welfare in the absence of position limits. Each speculator's utility at time $t = 0$ when he receives the private information and endowment is

$$J_0^S = -e^{-N_S A_S \left[\frac{1}{2} N_S A_S \text{Var}_0^S(P) y_S^2 + \frac{\lambda}{N_S} y_S^2 + \text{E}_0^S[V - \bar{V}]\frac{z}{N_S} - \frac{1}{2} A_S \text{Var}_0^S(V)\frac{z^2}{N_S}\right]}, \tag{1.117}$$

where

$$\lambda = \frac{\mu_0 d_S + A_S \mathrm{Var}_0^S(P)\bar{d}}{d_S - \bar{d} - \bar{d}/N_S^2},$$

$$y_S = \frac{\mathrm{E}_0^S[P] - F - A_S \mathrm{Cov}_0^S(P, V)z}{N_S A_S \mathrm{Var}_0^S(P) + \lambda/N_S}.$$

Firstly, in order to keep the welfare of the speculators finite, the sufficient condition is

$$A_S^2 \mathrm{Var}(V)\mathrm{Var}(z) = \frac{A_S^2}{\tau_V \tau_z} < 1.$$

Rewrite J_0^S as

$$J_0^S = -C_0 e^{-A_S C_1 C_2^2}, \tag{1.118}$$

where

$$C_0 = C_0(s, z) = e^{\mathrm{E}_0^S[V-\bar{V}]z - \frac{1}{2}A_S \mathrm{Var}_0^S(V)z^2}, \tag{1.119}$$

$$C_1 = \frac{\lambda/N_S^2 + \frac{1}{2}A_S \mathrm{Var}_0^S(P)}{[\lambda/N_S^2 + A_S \mathrm{Var}_0^S(P)]^2}, \tag{1.120}$$

$$C_2 = C_2(s_F) = (d_S - d)(s_F - \bar{V}) + \lambda_F(\bar{X} - \theta\bar{V}). \tag{1.121}$$

As N_S increases, λ decreases and C_1 increases monotonically. Note that

$$C_2 \sim \mathcal{N}\Big(\lambda_F(\bar{X} - \theta\bar{V}), (1-\omega_S)^2(d_S - \bar{d})^2(\tau_V^{-1} + \tau_F^{-1})\Big), \tag{1.122}$$

and it is easy to show that as N_S increases, the mean of C_2 decreases and the variance of C_2 increases. Therefore, $C_2(N_S)$ second order stochastic dominates $C_2(N_S + 1)$. In sum, the competitiveness among speculators has two effects. On one hand, less competitiveness (small N_S) means more profit ($C_2(N_S)$ SOSD $C_2(N_S + 1)$). On the other hand, it means higher cost from more severe adverse selection ($C_1(N_S) < C_2(N_S + 1)$).

The ex-ante welfare of each speculator with no position limit is

$$\mathcal{W}^S = \mathrm{E}[J_0^S] = -\frac{1}{\sqrt{I + A_S C^S \Sigma^S}} e^{-\frac{1}{2}\frac{h^2}{\mathrm{Var}_0^S(P)} + \frac{1}{2}[\frac{h(d_S - d)}{\mathrm{Var}_0^S(P)}]^2 G_{11}^S}, \tag{1.123}$$

where

$$u = \begin{pmatrix} s_F - \bar{V} \\ z \\ s - \bar{V} \end{pmatrix}, \quad B = \begin{pmatrix} [h_S - \frac{h}{A_S \mathrm{Var}_0^S(P)}](d_S - d) \\ 0 \\ 0 \end{pmatrix},$$

$$C = \begin{pmatrix} \frac{(d_S - d)^2}{A_S \mathrm{Var}_0^S(P)} & 0 & 0 \\ 0 & -\frac{A_S \mathrm{Var}_0^S(V)}{\beta^2} & \frac{1}{\beta}\frac{\tau_\epsilon}{\tau_V + \tau_\epsilon} \\ 0 & \frac{1}{\beta}\frac{\tau_\epsilon}{\tau_V + \tau_\epsilon} & 0 \end{pmatrix},$$

and $A = \frac{1}{2} A_S \mathrm{Var}_0^S(P)[h_S - \frac{h}{A_S \mathrm{Var}_0^S(P)}]^2$, $G^S = \Sigma^S(I + A_S C^S \Sigma^S)^{-1}$. As discussed above, the competitiveness among speculators (or N_S) has two offsetting effects on speculators' welfare. Therefore, whether a oligopoly market is better for speculators depends on the market condition, i.e., how large $d_S - \bar{d}$ is. Less competitive market is strictly better for speculators only when $|d_S - \bar{d}|$ is large enough to give sufficient rent that can be extracted from information advantage and offset the cost of adverse selection.

For the market makers, their utility at $t = 0$ is

$$J_0^M = -\frac{1}{\sqrt{|I + A_M C^M \Sigma^M|}} e^{-\frac{1}{2} A_M [A_M G_{11}^M y_M^2 + (\frac{1}{A_M \mathrm{Var}_1^M(V+\eta)} - A_M G_{22}^M)(\mu \bar{X})^2]}, \tag{1.124}$$

where

$$|I + A_M C^M \Sigma^M| = A_M^2 \mathrm{Var}_1^M(V + \eta)\Sigma_{22}^M + 1,$$

$$\frac{1}{A_M \mathrm{Var}_1^M(V + \eta)} - A_M G_{22}^M = \frac{1}{A_M \mathrm{Var}_1^M(V + \eta)[A_M^2 \mathrm{Var}_1^M(V + \eta)\Sigma_{22}^M + 1]}.$$

Therefore, the ex-ante welfare of the market makers with no position limit is

$$\mathcal{W}^M = \mathrm{E}[J_0^M] = -\frac{e^{-\frac{1}{2}\frac{(\mu\bar{X})^2}{\mathrm{Var}_1^M(V+\eta)[1+A_M^2\mathrm{Var}_1^M(V+\eta)\Sigma_{22}^M]}}}{\sqrt{1+A_M^2\mathrm{Var}_1^M(V+\eta)\Sigma_{22}^M}}\frac{e^{-\frac{1}{2}\frac{A_M^2 G_{11}^M(\mathrm{E}[y_M])^2}{1+A_M^2 G_{11}^M\mathrm{Var}(y_M)}}}{\sqrt{1+A_M^2 G_{11}^M\mathrm{Var}(y_M)}}, \tag{1.125}$$

where

$$\mathrm{E}[y_M] = h_M - \frac{1}{N_M A_M G_{11}^M}h,$$
$$\mathrm{Var}(y_M) = (\frac{d_M - d}{N_M A_M G_{11}^M})^2(\frac{1}{\tau_V} + \frac{1}{\tau_\epsilon} + \frac{A_S^2}{\tau_\epsilon^2}\frac{1}{\tau_z}).$$

As N_S increases, the variance of y_M stays the same but its mean decreases, thus $y_M(N_S)$ SOSD $y_M(N_S+1)$. Therefore, decreasing competitiveness will benefit market makers.

For commercial hedgers, their utility at $t=0$ is

$$\begin{aligned} J_0^H &= -\frac{e^{-\frac{1}{2}A_H\left[A_H G_{11}^H(y_H+\bar{X})^2+(\frac{1}{A_H\mathrm{Var}_1^H(\eta)}-A_H G_{33}^H)(\mu\bar{X})^2-A_H G_{22}^H(E_0^H[P])^2-2A_H G_{23}^H E_0^H[P]\mu\bar{X}+2F\bar{X}\right]}}{\sqrt{|I+A_H C^H\Sigma^H|}}, \\ &= -\Gamma_0 e^{-A_H\left[\frac{1}{2}\Gamma_2(s_F-\bar{V})^2+\Gamma_1(s_F-\bar{V})\right]}, \end{aligned} \tag{1.126}$$

where

$$\Gamma_2 = \frac{(d_H-d)^2}{N_H A_H G_{11}^H} - A_H G_{22}^H(\frac{\tau_F}{\tau_V+\tau_F})^2,$$
$$\Gamma_1 = (h_H + \bar{X} - \frac{h}{N_H A_H G_{11}^H})(d_H - d) - A_H G_{22}^H(\bar{V}-\mu\bar{X})\frac{\tau_F}{\tau_V+\tau_F} - A_H G_{23}^H\mu\bar{X}\frac{\tau_F}{\tau_V+\tau_F} + d\bar{X},$$
$$\Gamma_0 = \frac{e^{-\frac{1}{2}A_H\left[(\frac{1}{A_H\mathrm{Var}_1^H(\eta)}-A_H G_{33}^H)(\mu\bar{X})^2+A_H G_{11}^H(h_H+\bar{X}-\frac{h}{N_H A_H G_{11}^H})^2-A_H G_{22}^H(\bar{V}-\mu\bar{X})^2-2A_H G_{23}^H(\bar{V}-\mu\bar{X})\mu\bar{X}+2h\bar{X}\right]}}{\sqrt{|I+A_H C^H\Sigma^H|}}.$$

Therefore, the ex-ante welfare of the commercial hedgers with no position limit is

$$\mathcal{W}^H = \mathrm{E}[J_0^H] = -\frac{\Gamma_0}{\sqrt{1+A_H\Gamma_2\mathrm{Var}(s_F)}}e^{\frac{1}{2}\frac{A_H^2\Gamma_1^2\mathrm{Var}(s_F)}{1+A_H\Gamma_2\mathrm{Var}(s_F)}}, \tag{1.127}$$

where

$$\mathrm{Var}(s_F) = \frac{1}{\tau_V} + \frac{1}{\tau_\epsilon} + \frac{A_S^2}{\tau_\epsilon^2}\frac{1}{\tau_z}.$$

The change of $\mathcal{W}^H$ with the change of N_S is ambiguous.

Welfare with Speculative Position Limit

In this section, I compute the welfare of different market participants with speculative position limit L.

First, consider the case in which $d_S < \bar{d}$, speculators demands are downward sloping with respect to s_F, (i.e., τ_X is small enough), then there exist $\overline{s}_F(L) > \hat{s}_F(L)$ such that for $\hat{s}_F < s_F < \overline{s}_F$ they trade like without position limits because the position limit does not bind, while for $s_F \geq \overline{s}_F$ the speculators place market orders at L, and for $s_F \leq \hat{s}_F$ they place limit orders but get executed at L. As $L \to \infty$, $\hat{s}_F \to \overline{s}_F$. There is an order gap at $\overline{s}_F$, but not at $\hat{s}_F$.

The speculators' utility at $t = 0$ in the presence of position limit L is

$$J_L^S = \begin{cases} \hat{J}_0^S(s, z), & \text{if } s_F \leq \hat{s}_F, \\ J_0^S(s, z), & \text{if } \hat{s}_F < s_F < \overline{s}_F, \\ \overline{J}_0^S(s, z), & \text{if } s_F \geq \overline{s}_F, \end{cases}$$

where

$$\hat{J}_0^S(s, z) = -C_0 e^{-N_S A_S \left[\left[\mathrm{E}_0^S[P] - \hat{F} - A_S \mathrm{Cov}_0^S(P, V) z \right] L - \frac{1}{2} N_S A_S \mathrm{Var}_0^S(P) L^2 \right]},$$

$$\overline{J}_0^S(s, z) = -C_0 e^{-N_S A_S \left[\left[\mathrm{E}_0^S[\overline{P}] - \overline{F} - A_S \mathrm{Cov}_0^S(\overline{P}, V) z \right] L - \frac{1}{2} N_S A_S \mathrm{Var}_0^S(\overline{P}) L^2 \right]},$$

where $\overline{s}_F$ and $\hat{s}_F$ are determined by

$$\hat{y}_S(s_F, L) = \frac{\mathrm{E}_0^S[P] - \hat{F} - A_S \mathrm{Cov}_0^S(P, V) z}{N_S A_S \mathrm{Var}_0^S(P) + \lambda / N_S} \geq L,$$

$$\overline{y}_S(s_F, L) = \frac{\mathrm{E}_0^S[\overline{P}] - \overline{F} - A_S \mathrm{Cov}_0^S(\overline{P}, V) z}{N_S A_S \mathrm{Var}_0^S(\overline{P})} \geq L,$$

where

$$\mathrm{E}_0^S[P] - \hat{F} - A_S \mathrm{Cov}_0^S(P, V)z = (d_S - \underline{d})(s_F - \bar{V}) + \lambda_F(\bar{X} - \theta\bar{X}),$$

$$\mathrm{E}_0^S[\overline{P}] - \overline{F} - A_S \mathrm{Cov}_0^S(\overline{P}, V)z = \mathrm{E}[s_F - \bar{V}]1_{s_F \geq \overline{s}_F} + \lambda_F(\bar{X} - \theta\bar{X}).$$

Therefore, the welfare with the position limit is

$$\mathcal{W}_L^S = \int_{-\infty}^{\hat{s}_F} \hat{J}_0^S dp(s_F) + \int_{\hat{s}_F}^{\overline{s}_F} J_0^S dp(s_F) + \int_{\overline{s}_F}^{\infty} \overline{J}_0^S dp(s_F), \tag{1.128}$$

where $p(s_F)$ is the CDF of s_F, which is normally distributed.

Similarly, commercial hedgers' utility at $t = 0$ is

$$J^H = \begin{cases} \hat{J}_0^H(s_F), & s_F \leq \hat{s}_F, \\ J_0^H(s_F), & \text{if } \hat{s}_F < s_F < \overline{s}_F, \\ \overline{J}_0^H(\overline{s}_F), & \text{if } s_F \geq \overline{s}_F, \end{cases}$$

where

$$\hat{J}_0^H(s_F) = -\frac{e^{-\frac{1}{2}A_H\left[A_H G_{11}^H(\hat{y}_H + \bar{X})^2 + (\frac{1}{A_H \mathrm{Var}_1^H(\eta)} - A_H G_{33}^H)(\mu\bar{X})^2 - A_H G_{22}^H(\mathrm{E}_0^H[P])^2 - 2A_H G_{23}^H \mathrm{E}_0^H[P]\mu\bar{X} + 2\hat{F}\bar{X}\right]}}{\sqrt{|I + A_H C^H \Sigma^H|}},$$

$$\overline{J}_0^H(\overline{s}_F) = -\frac{e^{-\frac{1}{2}A_H\left[A_H \overline{G}_{11}^H(\overline{y}_H + \bar{X})^2 + (\frac{1}{A_H \mathrm{Var}_1^H(\eta)} - A_H \overline{G}_{33}^H)(\overline{\mu}\bar{X})^2 - A_H \overline{G}_{22}^H(\mathrm{E}_0^H[\overline{P}])^2 - 2A_H \overline{G}_{23}^H \mathrm{E}_0^H[\overline{P}]\overline{\mu}\bar{X} + 2\overline{F}\bar{X}\right]}}{\sqrt{|I + A_H \overline{C}^H \overline{\Sigma}^H|}},$$

and where

$$\hat{y}_H(s_F) = \frac{\mathrm{E}_0^H[P] - \hat{F}}{A_H G_{11}^H} - \frac{G_{12}^H}{G_{11}^H}\mathrm{E}_0^H[P] - \frac{G_{13}^H}{G_{11}^H}\mu\bar{X} - \bar{X},$$

$$\overline{y}_H(\overline{s}_F) = \frac{\mathrm{E}_0^H[\overline{P}] - \overline{F}}{A_H \overline{G}_{11}^H} - \frac{\overline{G}_{12}^H}{\overline{G}_{11}^H}\mathrm{E}_0^H[\overline{P}] - \frac{\overline{G}_{13}^H}{\overline{G}_{11}^H}\overline{\mu}\bar{X} - \bar{X}.$$

Note that $\overline{J}_0^H$ is a constant which depends on L through $\overline{s}_F$. Therefore, the welfare with position limit is

$$\mathcal{W}_L^H = \int_{-\infty}^{\hat{s}_F} \hat{J}_0^H dp(s_F) + \int_{\hat{s}_F}^{\overline{s}_F} J_0^H dp(s_F) + \overline{J}_0^H \int_{\overline{s}_F}^{\infty} dp(s_F). \tag{1.129}$$

Similarly, market makers' utility at $t = 0$ is

$$J_L^M = \begin{cases} \hat{J}_0^M(s_F), & \text{if } s_F \leq \hat{s}_F, \\ \\ J_0^M(s_F), & \text{if } \hat{s}_F < s_F < \overline{s}_F, \\ \\ \overline{J}_0^M(\overline{s}_F), & \text{if } s_F \geq \overline{s}_F, \end{cases}$$

where

$$\hat{J}_0^M(s_F) = -\frac{1}{\sqrt{|I + A_M C^M \Sigma^M|}} e^{-\frac{1}{2}A_M[A_M G_{11}^M \hat{y}_M^2 + (\frac{1}{A_M \mathrm{Var}_1^M(V+\eta)} - A_M G_{22}^M)(\mu\bar{X})^2]},$$

$$\overline{J}_0^M(\overline{s}_F) = -\frac{1}{\sqrt{|I + A_M \overline{C}^M \overline{\Sigma}^M|}} e^{-\frac{1}{2}A_M[A_M \overline{G}_{11}^M \overline{y}_M^2 + (\frac{1}{A_M \mathrm{Var}_1^M(V+\eta)} - A_M \overline{G}_{22}^M)(\overline{\mu}\bar{X})^2]},$$

and

$$\hat{y}_M = \frac{\mathrm{E}_0^M[P] - \hat{F}}{A_M G_{11}^M} - \frac{G_{12}^M}{G_{11}^M}\mu\bar{X}.$$

$$\overline{y}_M = \frac{\mathrm{E}_0^M[\overline{P}] - \overline{F}}{A_M \overline{G}_{11}^M} - \frac{\overline{G}_{12}^M}{\overline{G}_{11}^M}\overline{\mu}\bar{X}.$$

Note that $\overline{J}_0^M$ is a constant which depends on L through $\overline{s}_F$. Therefore, the welfare with position limit is

$$\mathcal{W}_L^M = \int_{-\infty}^{\hat{s}_F} \hat{J}_0^M dp(s_F) + \int_{\hat{s}_F}^{\overline{s}_F} J_0^M dp(s_F) + \overline{J}_0^M \int_{\overline{s}_F}^{\infty} dp(s_F). \tag{1.130}$$

Second consider the second case in which $d_S > \bar{d}$ and with large number of speculators, speculators' demands are upward sloping with respect to s_F, (i.e., τ_X is large enough), then there exist $\overline{s}_F > \hat{s}_F$ such that for $s_F < \hat{s}_F$ they trade like without position limit because the limit doesn't bind, while for $\hat{s}_F \leq s_F < \overline{s}_F$ the limit binds but the speculators still submit limit orders so the mixed information s_F still be fully revealed and the futures price is set to clear the market given that the speculators' position is L, and for $s_F \geq \overline{s}_F$ the speculators begin to submit market order of L instead of limit order. As $L \to \infty$, $\hat{s}_F \to \overline{s}_F$.

The speculators' utility at $t = 0$ is

$$J_L^S = \begin{cases} J_0^S(s_F), & \text{if } s_F < \hat{s}_F, \\ \hat{J}_0^S(s_F), & \text{if } \hat{s}_F \leq s_F < \overline{s}_F, \\ \overline{J}_0^S(\overline{s}_F), & \text{if } s_F \geq \overline{s}_F, \end{cases}$$

where

$$\hat{J}_0^S(s, z) = -e^{-A_S\left[A_S \mathrm{Var}_0^S(P) L \hat{y}_S - \frac{1}{2} A_S \mathrm{Var}_0^S(P) L^2 + E_0^S[V - \bar{V}]\frac{z}{\beta} - \frac{1}{2} A_S \mathrm{Var}_0^S(V)\frac{z^2}{N_S}\right]},$$

where

$$\hat{y}_S(s_F, L) = \frac{\mathrm{E}_0^S[P] - \hat{F} - A_S \mathrm{Cov}_0^S(P, V)\frac{z}{\beta}}{A_S \mathrm{Var}_0^S(P)} \geq L, \text{ for } s_F \geq \hat{s}_F.$$

Therefore, the welfare with the position limit is

$$\mathcal{W}_L^S = \int_{-\infty}^{\hat{s}_F} J_0^S dp(s_F) + \int_{\hat{s}_F}^{\overline{s}_F} \hat{J}_0^S dp(s_F) + \int_{\overline{s}_F}^{\infty} \overline{J}_0^S dp(s_F). \tag{1.131}$$

Similarly, the commercial hedgers' utility at $t = 0$ is

$$J_L^H = \begin{cases} J_0^H(s_F), & \text{if } s_F < \hat{s}_F, \\ \hat{J}_0^H(s_F), & \text{if } \hat{s}_F \leq s_F < \overline{s}_F, \\ \overline{J}_0^H(\overline{s}_F), & \text{if } s_F \geq \overline{s}_F, \end{cases}$$

Note that $\overline{J}_0^H$ is a constant which depends on L through $\overline{s}_F$. Therefore, the welfare of the commercial trader is

$$\mathcal{W}_L^H = \int_{-\infty}^{\hat{s}_F} J_0^H dp(s_F) + \int_{\hat{s}_F}^{\overline{s}_F} \hat{J}_0^H dp(s_F) + \overline{J}_0^H \int_{\overline{s}_F}^{\infty} dp(s_F). \tag{1.132}$$

Similarly, market makers' utility at $t = 0$ is

$$J_L^M = \begin{cases} J_0^M(s_F), & \text{if } s_F < \hat{s}_F, \\ \hat{J}_0^M(s_F), & \text{if } \hat{s}_F \leq s_F < \overline{s}_F, \\ \overline{J}_0^M(\overline{s}_F), & \text{if } s_F \geq \overline{s}_F, \end{cases}$$

Note that $\overline{J}_0^M$ is a constant which depends on L through $\overline{s}_F$. Therefore, the welfare with position limit is

$$\mathcal{W}_L^M = \int_{-\infty}^{\hat{s}_F} J_0^M dp(s_F) + \int_{\hat{s}_F}^{\overline{s}_F} \hat{J}_0^M dp(s_F) + \overline{J}_0^M \int_{\overline{s}_F}^{\infty} dp(s_F). \tag{1.133}$$

Lastly, consider the third case in which $d_S > \bar{d}$ and with small number of speculators, thus $\hat{s}_F > \overline{s}_F$. Hence, their utility at $t = 0$ is

$$J_L^S = \begin{cases} J_0^S(s_F), & \text{if } s_F < \overline{s}_F, \\ \overline{J}_0^S(\overline{s}_F), & \text{if } s_F \geq \overline{s}_F. \end{cases}$$

Then the welfare with SPL is

$$\mathcal{W}_L^S = \int_{-\infty}^{\overline{s}_F} J_0^S dp(s_F) + \int_{\overline{s}_F}^{\infty} \overline{J}_0^S dp(s_F). \tag{1.134}$$

Similarly, the commercial hedgers' utility at $t = 0$ is

$$J_L^H = \begin{cases} J_0^H(s_F), & \text{if } s_F < \overline{s}_F, \\ \overline{J}_0^H(\overline{s}_F), & \text{if } s_F \geq \overline{s}_F. \end{cases}$$

Then their welfare with SPL is

$$\mathcal{W}_L^H = \int_{-\infty}^{\overline{s}_F} J_0^H dp(s_F) + \overline{J}_0^H \int_{\overline{s}_F}^{\infty} dp(s_F). \tag{1.135}$$

Similarly, market makers' utility at $t = 0$ is

$$J_L^M = \begin{cases} J_0^M(s_F), & \text{if } s_F < \bar{s}_F, \\ \\ \bar{J}_0^M(\bar{s}_F), & \text{if } s_F \geq \bar{s}_F. \end{cases}$$

Then their welfare with SPL is

$$\mathcal{W}_L^M = \int_{-\infty}^{\bar{s}_F} J_0^M dp(s_F) + \bar{J}_0^M \int_{\bar{s}_F}^{\infty} dp(s_F). \tag{1.136}$$

Therefore, the total ex-ante welfare with SPL is

$$\mathcal{W}_L = N_H \mathcal{W}_L^H + N_M \mathcal{W}_L^M + N_S \mathcal{W}_L^S. \tag{1.137}$$

Unsurprisingly, $\mathcal{W}_L < \mathcal{W}$ in any market condition. However, for each individual group of traders, this is not the case. Unfortunately, for commercial hedgers, in any market condition,

$$\mathcal{W}_L^H < \mathcal{W}^H, \quad \frac{d\mathcal{W}_L^H}{dL} > 0,$$

which implies that the tighter L is, the more welfare loss commercial hedgers suffers. For market makers, in the first case and the second case,

$$\mathcal{W}_L^M > \mathcal{W}^H, \quad \frac{d\mathcal{W}_L^M}{dL} < 0,$$

which implies that the tighter L is, the more welfare gain market makers have in the first case and the second case. However, in Market Condition the second case, $\mathcal{W}_L^M$ is not a monotone function of L, and

$$\lim_{L \to 0} \mathcal{W}_L^M > \mathcal{W}^M, \quad \min_L \mathcal{W}_L^M < \mathcal{W}^M,$$

which implies that in the second case market makers can also suffer a welfare loss with certain range of L.

For speculators, $\mathcal{W}_L^S$ is also not a monotone function of L in any market condition,

$$\lim_{L\to 0} \mathcal{W}_L^S < \mathcal{W}^S, \quad \max_L \mathcal{W}_L^S > \mathcal{W}^S,$$

which implies that speculators can be always better off with certain range of L.

1.8.8 Proof of Proposition 11

Proof. Suppose there is a one period Grossman-Stiglitz economy (24) without noisy traders. The liquidation value of the stock is $V+\eta$ (i.e., the commodity), and informed traders know V privately and also have a liquidity shock X (i.e., commercial hedgers). Uninformed traders have no private information or liquidity shock (i.e., market makers). There is a public information about V (i.e., futures price F): $s_F \sim \mathcal{N}(\bar{V}, \tau_F^{-1})$, which is announced before the realization of liquidity shock X. Thus, it is equivalent to show that as τ_F decreases, both informed traders and uninformed traders are worse off. The welfare of informed and uninformed traders are

$$\mathcal{W}_I = -\frac{e^{-\frac{A_I}{2}\bar{u}^\top [C - A_I C^\top \Sigma (I + A_I C\Sigma)^{-1} C]\bar{u}}}{\sqrt{|I + A_I C\Sigma|}},$$

$$\mathcal{W}_U = -\frac{e^{-\frac{1}{2[1+A_U^2 \mathrm{Var}^U(V+\eta)\mathrm{Var}(x_U)]}\frac{(\mu\bar{X})^2}{\mathrm{Var}^U(V+\eta)}}}{\sqrt{1 + A_U^2 \mathrm{Var}^U(V+\eta)\mathrm{Var}(x_U)}},$$

where $A_{I,U}$ are risk aversions for informed and uninformed traders respectively, and $\bar{u}$, Σ and C can be found in 1.8.1. Both $\mathcal{W}_I$ and $\mathcal{W}_U$ are monotone increasing functions of τ_F as following figure shows.

The pattern in Figure 1.13 still holds for non-zero $\bar{V}$ and $\bar{X}$. For non-zero $\bar{V}$ and $\bar{X}$, the change of $\mathcal{W}_I$ is more dramatic with respect to the change of τ_F. Welfare is more sensitive

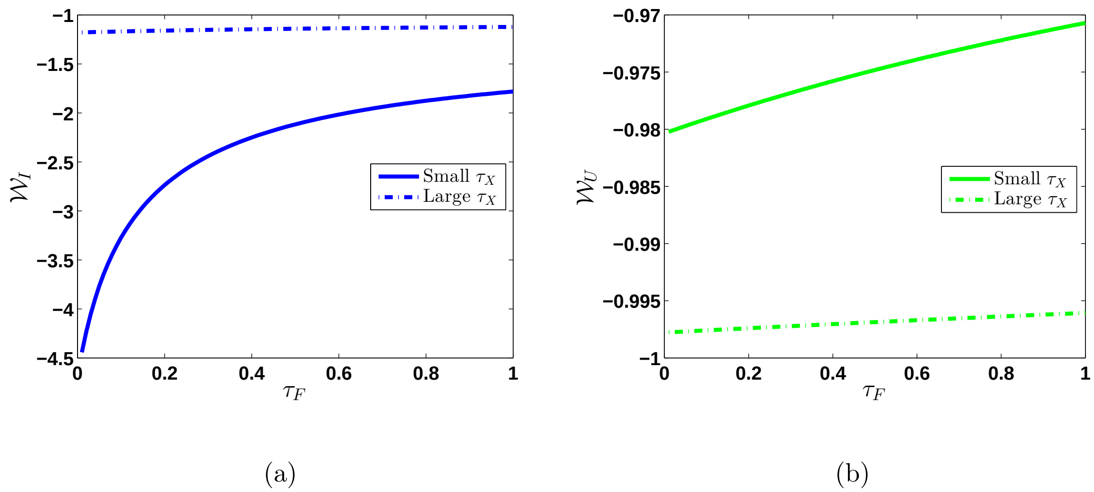


Figure 1.13: Welfare of informed trader (a) and uninformed trader (b) with respect to the accuracy of public information for $\bar{V} = \bar{X} = 0$.

to the accuracy of public information for small τ_X. As τ_X increases, informed traders are better off and uninformed traders are worse off. This is consistent with risk sharing trading, not speculating in spot market.

Although both noise trading and liquidity shock prevent private information from fully revealing, they are not always equivalent. In standard Grossman-Stiglitz model with noise trading, public information always hurts informed traders but benefits uninformed traders, because it reduces information asymmetry between informed and uninformed traders, no matter the public information is revealed before the realization of noise trading or after. In liquidity shock setup public information still always benefits uninformed traders. However, when the public information is revealed makes huge difference for informed traders. If the public information is announced before the realization of the liquidity shock, then it also benefits informed traders by reducing their hedging cost. If the public information is announced after the realization of the liquidity shock, then it hurts informed traders. This is why speculators prefer not to reveal their state with position limit. Although they decide the critical point $\hat{s}_F$ before the realization of X, the information asymmetry is increased after the realization of X. Thus, speculators benefit from reducing their information revealing. $\square$

1.8.9 Proof of Proposition 12

Proof. The existence condition $\lambda > 0$ requires that $d_S - 2\bar{d} > 0$. By plugging the formulas of ds,

$$(\frac{\tau_\epsilon}{\tau_V})^2 + \frac{\tau_\epsilon}{\tau_V} - 1 < 0, \tag{1.138}$$

which implies $0 < \frac{\tau_\epsilon}{\tau_V} < \frac{\sqrt{5}-1}{2} \approx 0.618$. This gives a necessary condition $\tau_z < \frac{A_S^2}{\beta^2 \tau_\epsilon} \frac{\tau_V}{\tau_V + \tau_\epsilon}$. This is the existence condition for collusive equilibrium. However, this can not guarantee speculators to collude. λ measures the cost of implementing market power, thus λ can not be too big, otherwise the cost can not be compensated by the monopoly informational advantage. Therefore, there exist a upper bound $\hat{\lambda}$ for λ, i.e., $\lambda \le \hat{\lambda}$. Equivalently there exists a lower bound $\kappa > 0$ such that for $d_S \ge 2\bar{d} + \kappa\bar{d}$ speculators choose to collude and exert market power. κ is numerically calculated. Thus, there exists a boundary function $\tau_z = g(\tau_X)$ in the space (τ_X, τ_z) such that:

$$\tau_z = g(\tau_X) \Longleftrightarrow d_S(\tau_z, \tau_X) = (2+\kappa)\bar{d}(\tau_z, \tau_X). \tag{1.139}$$

Thus, the asymptotic equation of (τ_z, τ_X) on boundary g is

$$\frac{A_M + A_H}{A_M}\tau_X + \frac{A_H^3}{A_M \tau_\eta} \frac{\tau_X}{\tau_X + \frac{A_H^2}{\tau_\eta^2}(\tau_V + \tau_F)} = A_H^2 \frac{(\frac{1}{\tau_\epsilon} + \frac{1}{\tau_\eta} + \frac{\tau_V}{\tau_\epsilon \tau_\eta})\tau_z + \frac{A_S^2}{\tau_\epsilon^2}(1 + \frac{\tau_V}{\tau_\eta})}{\kappa\bar{d} + \frac{A_S^2 \tau_V}{\tau_\epsilon^2} - (1 + \frac{\tau_V}{\tau_\epsilon})\tau_z}. \tag{1.140}$$

Therefore, $g(\tau_X)$ is a monotone increasing concave function. Thus, g is invertible, and $g^{-1}(\tau_z)$ is monotone increasing convex. The asymptotic formula of g^{-1} for large τ_X is given by

$$g^{-1}(\tau_z) \approx \frac{A_M}{A_M + A_H}[A_H^2 \frac{(\frac{1}{\tau_\epsilon} + \frac{1}{\tau_\eta} + \frac{\tau_V}{\tau_\epsilon \tau_\eta})\tau_z + \frac{A_S^2}{\tau_\epsilon^2}(1 + \frac{\tau_V}{\tau_\eta})}{\kappa \bar{d} + \frac{A_S^2 \tau_V}{\tau_\epsilon^2} - (1 + \frac{\tau_V}{\tau_\epsilon})\tau_z} - \frac{A_H^3}{A_M \tau_\eta}],$$

and for small τ_X is given by

$$g^{-1}(\tau_z) \approx [\frac{A_M + A_H}{A_M} + \frac{A_H}{A_M} \frac{\frac{\tau_\eta}{\tau_\epsilon} + \frac{A_S^2}{\tau_\epsilon^2}\frac{\tau_\eta}{\tau_z}}{1 + \frac{\tau_V}{\tau_\epsilon} + \frac{A_S^2}{\tau_\epsilon^2}\frac{\tau_V}{\tau_z}}]^{-1} A_H^2 \frac{(\frac{1}{\tau_\epsilon} + \frac{1}{\tau_\eta} + \frac{\tau_V}{\tau_\epsilon \tau_\eta})\tau_z + \frac{A_S^2}{\tau_\epsilon^2}(1 + \frac{\tau_V}{\tau_\eta})}{\kappa \bar{d} + \frac{A_S^2 \tau_V}{\tau_\epsilon^2} - (1 + \frac{\tau_V}{\tau_\epsilon})\tau_z}.$$

$\square$

Chapter 2

Dynamic of Market Making and Asset Pricing

2.1 Introduction

Previous literature focuses on the informational oligopoly in the financial markets and assumes that market makers are risk neutral and perfect competitive. In this paper, we focus on the imperfect competition among market makers who are so risk averse that they won't take inventory over time. We want to study the dynamics of bid-ask spread and trading volume to understand how these may interact with each other in shaping asset prices and market liquidity. We develop a dynamic model of market making with asymmetric information where imperfectly competitive market makers match offsetting trades. More specifically, this model is finite time horizon with multiple trading periods. Hence, the equilibrium solution for this model is not steady-state, but time-dependent. This model helps to better understand how equilibrium prices evolve to steady state. There are three types of traders:

informed traders with an initial liquidity shock, uninformed traders, and market makers. We assume that there are a large number of informed and uninformed traders so that their market power can be negligible. We also assume that informed and uninformed traders cannot trade directly but have to trade through market makers. This assumption captures the feature of over-the-counter markets, where 98% trades are facilitated by dealers. Since there is no need of searching for informed or uninformed traders, this model is more suitable to study the mature OTC markets with stable dealer networks. Unlike informed or uninformed traders, market makers are imperfectly competitive and hold zero inventory over each time period, and they optimally facilitate trading in both bid and ask markets by adjusting their trading quantities. The market power of market makers may comes the limited number of them, or comes from their monopoly quoting speed. Therefore, this model can also be applied to the financial markets with high frequent traders, who quote faster than other traders in the market so that most of the trades have to go through them and they hold close-to-zero inventory at the end of the day. This is a very general framework which can be applied to many financial markets.

This is a very challenge work for two main reasons. First, it is a dynamic model with multiple rounds of trading and all traders in the model are rational non-myopic and choose optimal trading quantity. Second, there are basically two markets: bid market for sellers and ask market for buyers, and market makers need to clear both markets simultaneously. Still, we are able to solve for equilibrium bid and ask prices and market depths in closed-form, and examine how informed traders dynamically hedge their liquidity shock and reveal private information. Because of the challenges, this model give some results that are distinct from a one-period model with perfectly competitive traders. Firstly, unlike a static model in which traders always hedge their liquidity shock no matter they are informed or uninformed, in this dynamic model informed traders may speculate on their liquidity shock when their private information is very accurate. In other words, when informed traders expect to receive accurate private information, they will be over-hedge before the accurate private information coming in. Because the hedging cost for them will be higher after they receive accurate private information due to a higher adverse selection. After they receive the accurate private information, they will speculate on their previous liquidity shock instead of hedge. Secondly, unlike the model with perfectly competitive market makers, the average trading volume in this model is not monotonic decreasing with time. In a competitive model, average trading volume always decreases with time, because more information is revealed with time. However, this model can generate a reverse-J shape with time. Because rational traders take into account the impact of trading volume on bid-ask spread, the trading volume becomes part-dependent, which means high trading volume in the past can lead to a higher trading volume in the current time. We also plan to examine what is correlation among bid-ask spread, trading volume and return.

Most existing literature about the dynamic of bid and ask prices stems from the paper by Glosten and Milgram 1985. Our paper differs from that stream of literature in two ways. First, there is no quantity effect in those papers, where each trade only has one unit. Second, market makers in those paper are perfectly competitive. The bid-ask spread comes from adverse selection, while the bid-ask spread comes from imperfect competition among market makers in our model.

2.2 The Model

We consider a multiperiod model of trading in a market where all the transactions have to go through market makers. There are three types of traders: informed traders, uninformed traders and market makers. Informed traders receive a common private information about the fundamental value of the asset. They also receive a liquidity shock at the beginning of the trading, and only themselves know the size of the shock. Other traders, including uninformed traders and market makers, have no private information but can infer a mixed signal from the price, which is a linear combination of informed traders' private information and the size of their liquidity shock. Since neither the private information nor the size of liquidity shock can be observed by uninformed traders or market makers, in equilibrium the private information of informed traders will not be fully revealed. The model is further defined as follows.

2.2.1 The Assets

There is a riskless asset and a risky asset (stock) available for trading at dates $1, ..., T-1$. There also has an illiquid asset, which can not be liquidated until the final date T. The riskless asset is of perfectly elastic supply with the rate of return r being a nonnegative constant. For simplicity, we assume $r = 0$. There is zero supply of the riskless asset. Each share of the stock pays a liquidation value of V at the final date T. Shares of the stock are infinitely divisible and are traded through market makers. The total supply of of the risky asset is fixed at Θ. Without loss of generality, we assume that each share of the illiquid asset pays a liquidation value of V at the final date T.

2.2.2 Investors

There are N_I of identical informed investors, N_U of identical uninformed investors, and N_M designated market makers who are also uninformed. We assume $N_{I,U} >> N_M$ so that informed and uninformed traders are perfectly competitive and market makers are imperfectly competitive. All investors are endowed with zero share of riskless asset. Every type $i \in \{I, U, M\}$ investor is endowed with θ_0^i shares of risky assets before trading starts. So the total supply of the stock is fixed at $\Theta = N_I\theta_0^I + N_U\theta_0^U + N_M\theta_0^M$. Moreover, informed traders are also endowed with X shares of illiquid asset in total (i.e., each informed trader receives X/N_I units of illiquid assets). We also assume that investors only consume at the final date, and W_T^i is type i investor final wealth. Each trader of type i maximizes expected utility by choosing their positions in the stock market

$$\max_{\theta^i} \mathrm{E}[-e^{-N_i\lambda^i W_T^i} \mid \mathcal{F}_T^i],$$

where $\mathcal{F}_t^i$ is type i investor's information set at date t, and λ^i is the total risk aversion coefficient of type i traders. Since all the orders have to go through market makers, they can set ask prices A_t, at which they sell, and bid price B_t, at which they buy, by choosing their positions as a cournot game.

$$\begin{aligned}
W_T^I &= W_0^I + \sum_{t=1}^{T-1}[(\theta_t^I - \theta_{t-1}^I)^- B_t - (\theta_t^I - \theta_{t-1}^I)^+ A_t] + \theta_{T-1}^I V + VX/N_I \\
W_T^U &= W_0^U + \sum_{t=1}^{T-1}[(\theta_t^U - \theta_{t-1}^U)^- B_t - (\theta_t^U - \theta_{t-1}^U)^+ A_t] + \theta_{T-1}^U V \\
W_T^{M,j} &= W_0^M + \sum_{t=1}^{T-1}(\alpha_t^{M,j} A_t - \beta_t^{M,j} B_t) + \theta_{T-1}^M V
\end{aligned}$$

where $\alpha_t^{M,j}$ and $\beta_t^{M,j}$ are the quantities that market maker $j \in \{1, 2, ..., N_M\}$ buys and sells at time t. $x^+ = x$ if $x > 0$ otherwise 0, and similarly $x^- = -x$ if $x < 0$ otherwise 0. Since their utility functions are CARA, their initial wealth should not matter.

2.2.3 Information Structure

All investors have the same prior information about V and X. Assume that the prior distribution are $V \sim \mathcal{N}(\bar{V}, \tau_V^{-1})$ and $X \sim \mathcal{N}(0, \tau_X^{-1})$, and V and X are independent.

At $t = 0$, i.e. before the trading, each type I investor receives X/N_I shares of illiquid assets, and X is known only by themselves. At each date t, only type I informed investors receive a common private signal v_t about stock's liquidation part V:

$$v_t = V + \varepsilon_t,$$

where $\varepsilon_t \sim \mathcal{N}(0, \tau_{\varepsilon,t}^{-1})$ is the *i.i.d.* noise. The bid and ask prices of the stock A_t and B_t are observable to all investors. Thus, we can write the investors's information set as follows:

$$\mathcal{F}_t^I = \{\mathcal{F}_0, X, \underline{v}_t, \underline{A}_t, \underline{B}_t\} \tag{2.1}$$

$$\mathcal{F}_t^U = \mathcal{F}_t^M = \{\mathcal{F}_0, \underline{A}_t, \underline{B}_t\} \tag{2.2}$$

where $\mathcal{F}_0$ represents prior information as given by the prior distributions. We introduce the notation $\underline{Z}_t \equiv (Z_1, ..., Z_t)$ for any stochastic process $\{Z_t\}$, i.e., $\underline{Z}_t$ represents the history of Z_t up to and including t. Also, for notation simplicity, we define, respectively, expectations conditional on $\mathcal{F}_t^I$ and $\mathcal{F}_t^U$:

$$\hat{Z}_t^I = \mathrm{E}_t^I[Z_t], \quad \hat{Z}_t^U = \mathrm{E}_t^U[Z_t],$$

$$o_{Z,t}^I = \mathrm{Var}_t^I(Z), \quad o_{Z,t}^U = \mathrm{Var}_t^U(Z),$$

where

$$\mathrm{E}_t^I[.] \equiv \mathrm{E}[. \mid \mathcal{F}_t^I], \quad \mathrm{E}_t^U[.] \equiv \mathrm{E}[.|\mathcal{F}_t^U] \equiv \mathrm{E}[.|\mathcal{F}_t^M],$$

$$\mathrm{Var}_t^I[.] \equiv \mathrm{Var}[. \mid \mathcal{F}_t^I], \quad \mathrm{Var}_t^U[.] \equiv \mathrm{Var}[. \mid \mathcal{F}_t^U] \equiv \mathrm{Var}[. \mid \mathcal{F}_t^M].$$

For simplicity, we assume that all noise $\{\varepsilon_t\}$ are independent with V and X. As the number of trading dates increases, more private information will be revealed to informed traders and true value of V will eventually be revealed to informed traders, so to uninformed traders and market makers by the equilibrium prices. For simplicity, we solve the equilibrium without public information. The extension to the case with public information is straightforward.

2.3 Equilibrium with Competitive Market Makers

In this section, we solve for the equilibrium of the economy with large number of market makers so that they drive their profits from market making to zero. In this case, there is no bid-ask spread, so $A_t = B_t = P_t$. Markets makers are the same as uninformed investors, just like [24], in which uninformed investors are market making.

Given the well-known properties of CARA preference under normal distributions of payoffs and signals, we only consider the linear equilibria of the economy. In a linear equilibrium, the equilibrium stock price can be expressed as a linear function of the state variables of the economy. In other words, we have

$$P_t = \mathbf{L}[\Phi_t], \tag{2.3}$$

where Φ_t represents the vector of state variables of the economy at date t. The general history dependence of the equilibrium under differential information leads to difficulties in solving the equilibrium since the dimensionality of the state variables increases over time without bound. Generally speaking, the state vector Φ_t can contain all the information up to time t: $\Phi_t = (X, \underline{v}_t)$. We use $\mathbf{L}[.]$ to denote a general linear relation. Since we often do not care about the actual functional form within the linear class, the same symbol is used for different functions.

In the current setting, however, the general history dependence can be simplified by properly choosing the state space. The equilibrium only include first-order expectation due to hierarchy information structure. This allows us to solve for the equilibrium prices and trading strategies just like [28].

Lemma 2. *In a linear equilibrium of the economy, the price function can be expressed as follows:*

$$P_t = \mathbf{L}[X, \underline{v}_t]. \tag{2.4}$$

Furthermore, we can rewrite P_t *as*

$$\begin{aligned} P_t &= \mathbf{L}[X, v_t, \underline{P}_{t-1}] && (2.5) \\ &= a_t(v_t - h_t X) + \mathbf{L}[\underline{P}_{t-1}]. && (2.6) \end{aligned}$$

Define

$$s_t = v_t - h_t X. \tag{2.7}$$

Thus, $\{\underline{s}_t\}$ *is informationally equivalent to* $\{\underline{P}_t\}$*.*

This implies that, given past prices, observing the current price is equivalent to observe s_t, which is a linear combination of the two unknowns for uninformed investors. Consequently, in a linear equilibrium the uninformed investors' information set $\mathcal{F}_t^U = \{\mathcal{F}_0, \underline{P}_t\} = \{\mathcal{F}_0, \underline{s}_t\}$, that is $\{\underline{P}_t\} \Leftrightarrow \{\underline{s}_t\}$.

In order to derive each investor's optimal stock holding, we have to solve the conditional expectations, given his information set. In the linear equilibrium, calculating the conditional expectations of the state variables is a linear filtering problem, since all the signals are linear in the state variables, including endogenous signals such as prices. Utilizing the equivalence between $\{\mathcal{F}_0, \underline{P}_t\} = \{\mathcal{F}_0, \underline{s}_t\}$, we can solve for uninformed investors' conditional expectation of X and v_t. Informed investors know X and v_t, so their expectations of X and v_t are the true values. Therefore, in this hierarchy structure of information, the higher-order expectations are reduced to first-order expectations.

Lemma 3. *Given the linear price function, $\hat{V}_t^I$ and $(\hat{V}_t^U, \hat{X}_t^U)$ are determined by the following stochastic differential equations:*

$$\hat{V}_t^I = \hat{V}_{t-1}^I + K_t^I(v_t - \mathrm{E}_{t-1}^I[v_t]), \tag{2.8}$$

and

$$\begin{pmatrix} \hat{V}_t^U \\ \\ \hat{X}_t^U \end{pmatrix} = \begin{pmatrix} \hat{V}_{t-1}^U \\ \\ \hat{X}_{t-1}^U \end{pmatrix} + K_t^U(s_t - \mathrm{E}_{t-1}^U[s_t]), \tag{2.9}$$

and $\hat{X}_0^U = 0$, $\hat{V}_0^U = \bar{V}$, K_t^I is a scalar, and K_t^U is a 2×1 vector.

Proof. See Appendix 2.8.1. □

For uninformed traders, $\{\hat{X}_t^U, \hat{V}_t^U\}$ follows a Gaussian Markov process under the information process generated by $\mathcal{F}_t^U$, since $(\hat{V}_t^U, \hat{X}_t^U)$ can be expressed as a recursive equation of $(\hat{V}_{t-1}^U, \hat{X}_{t-1}^U)$ with the surprise in s_t as innovations. Similarly, for informed traders, $\{\hat{V}_t^I\}$

follows a Gaussian Markov process under the information process generated by $\mathcal{F}_t^I$, and $\{\hat{X}_t^U, \hat{V}_t^U\}$ is measurable with respect to $\mathcal{F}_t^I$. Furthermore, $s_t \equiv v_t - h_t X \subseteq \mathcal{F}_t^U \subseteq \mathcal{F}_t^I$, thus $v_t - h_t X = \hat{v}_t^U - h_t \hat{X}_t^U$, where $\hat{v}_t^U = \mathrm{E}_t^U[v_t]$.

Now we start to solve investors' optimal stock holdings. Let $P_T = V$. An investor's optimal stock holding at each date is given by the solution to the following optimization problem:

$$\max_{\theta_t^U} \mathrm{E}[-e^{-N_U \lambda^U W_T^U} \mid \mathcal{F}_t^U] \tag{2.10}$$

$$s.t. \quad W_{t+1}^U = W_t^U + \theta_t^U (P_{t+1} - P_t)$$

$$\max_{\theta_t^I} \mathrm{E}[-e^{-N_I \lambda^I W_T^I} \mid \mathcal{F}_t^I] \tag{2.11}$$

$$s.t. \quad W_{t+1}^I = W_t^I + \theta_t^I (P_{t+1} - P_t), \quad t \le T-2$$

$$and \quad W_T^I = W_{T-1}^I + \theta_{T-1}^I (V - P_{T-1}) + VX/N_I.$$

Conjecture 1. We conjuecture that the equilibrium price is

$$P_t = \omega_t(\hat{V}_t^I - \mu_t X) + (1-\omega_t)\hat{V}_t^U - \mu_t \omega_t \Theta, \tag{2.12}$$

$$= \hat{V}_t^I - \mu_t (X - \hat{X}_t^U) - \mu_t \omega_t (\hat{X}_t^U + \Theta), \tag{2.13}$$

$$= \hat{V}_t^U - \mu_t \omega_t (\hat{X}_t^U + \Theta), \tag{2.14}$$

where $0 < \omega_t < 1$. Define

$$S_t = \hat{V}_t^I - \mu_t X, \tag{2.15}$$

which is informationally equivalent to P_t, and $\hat{V}_t^I - \mu_t X = \hat{V}_t^U - \mu_t \hat{X}_t^U$.

Under this conjecture, this is easy to find that $\{\underline{s}_t\} \Leftrightarrow \{\underline{S}_t\}$. Since $\hat{V}_t^I = \mathbf{L}[\underline{v}_t] = \mathbf{L}[\underline{s}_t] + \mu_t X$, $\hat{V}_t^I - \mu_t X = \mathbf{L}[\underline{s}_t] = S_t$. Therefore, $\{\underline{\mu}_t\}$ and $\{\underline{h}_t\}$ have a one-to-one mapping as follows:

$$\mu_t = \frac{\sum_{s=1}^t \tau_{\varepsilon,s} h_s}{\tau_V + \sum_{s=1}^t \tau_{\varepsilon,s}}, \quad \text{or} \tag{2.16}$$

$$\tau_{\varepsilon,t} h_t = (\tau_V + \sum_{s=1}^t \tau_{\varepsilon,s})\mu_t - (\tau_V + \sum_{s=1}^{t-1} \tau_{\varepsilon,s})\mu_{t-1}, \tag{2.17}$$

where $\tau_{\varepsilon,1} h_1 = (\tau_V + \tau_{\varepsilon,1})\mu_1$. Once we know $\{h_t\}$, then we know $\{\mu_t\}$, and vice versa. These coefficients determine the information revealing through prices. So in this section, we will compute the sets of these information coefficients backward.

Proposition 16. *There is only one equilibrium solution for* $\{\mu_t\}$ *as*

$$\mu_t = \lambda^I o^I_{V,t} \equiv \frac{\lambda^I}{\tau_V + \sum_{s=1}^t \tau_{\varepsilon,s}}. \tag{2.18}$$

Equivalently,

$$h_1 = \frac{\lambda^I}{\tau_{\varepsilon,1}}, \tag{2.19}$$

$$h_t = 0, \ \textit{for } t = 2, 3, ..., T-1. \tag{2.20}$$

Proof. See Appendix 2.8.2. □

Note that from the price, uninformed traders can infer $\hat{V}^I_t - \mu_t X$, which serves as a signal for informed traders' estimation about stock value $\hat{V}^I_t$ and informed traders' total liquidity shock X. According to this proposition, the noise-signal ratio for uninformed traders is μ_t, which decreases over time when there is new private information for informed traders every period. Thus, the current price becomes more and more informative about information traders' estimation of the stock value. More specifically, this proposition tells us that uninformed traders know $s_1 = v_1 - h_1 X$, which is a mixed signal about informed traders' private information v_1 and total liquidity shock X in the first period, and know $s_t = v_t$

for $t = 2, 3, ..., T-1$, which is the exact informed traders' private information in the later periods. Thus, uninformed traders' estimation about stock value $\{\hat{V}_t^U\}$ and liquidity shock $\{\hat{X}_t^U\}$ and variance $\{O_t^U\}$ can be expressed explicitly.

Proposition 17. *The value functions for informed and uninformed traders have the forms as below*

$$J_t^I = -\rho_t^I e^{-\lambda^I N_I W_t^I - \lambda^I X \hat{V}_t^I - \frac{1}{2}\Phi_t^\top H_t^I \Phi_t}, \tag{2.21}$$

$$J_t^U = -\rho_t^U e^{-\lambda^U N_U W_t^U - \frac{1}{2} H_t^U (\hat{X}_t^U + \Theta)^2}, \tag{2.22}$$

where $\Phi_t = \begin{pmatrix} \hat{X}_t^U \\ 1 \end{pmatrix}$ *is the state vector for informed traders, and* $H_t^I = \begin{pmatrix} k_t & -k_t^\Theta \Theta \\ -k_t^\Theta \Theta & 0 \end{pmatrix}$, *which means there is no cross term of* $X\hat{X}_t^U$ *in the value function.*

In the last period,

$$\omega_{T-1} = \frac{(\lambda^I o_{V,T-1}^I)^{-1}}{(\lambda^I o_{V,T-1}^I)^{-1} + (\lambda^U o_{V,T-1}^U)^{-1}}. \tag{2.23}$$

By Proposition 16, ω_t *can be computed recursively as*

$$\begin{aligned} &\mu_t \omega_t - \mu_{t+1}\omega_{t+1} \\ &= \left(1 + \frac{H_{t+1}^U K_{X,t+1}^U}{\lambda^U b_{Q,t+1}} - \frac{k_{t+1}^\Theta K_{X,t+1}^U}{\lambda^I a_{Q,t+1}}\right)\left(\frac{1}{\lambda^I a_{Q,t+1}^2 \Xi_t^I} + \frac{1}{\lambda^U b_{Q,t+1}^2 \Xi_t^U}\right)^{-1}, \end{aligned} \tag{2.24}$$

where H_{t+1}^U, k_{t+1} *and* k_{t+1}^Θ *(i.e.,* H_{t+1}^I*) can be computed backward, and* a_{t+1}, b_{t+1} *and* Ξ_t *can be calculated by Proposition 16.*

The equilibrium stock holdings are

$$N_I \theta_t^I = \gamma_t(\hat{X}_t^U + \Theta) - \hat{X}_t^U, \tag{2.25}$$

$$N_U \theta_t^U = (1-\gamma_t)(\hat{X}_t^U + \Theta), \tag{2.26}$$

where

$$\gamma_t = \left(1 + \frac{H^U_{t+1} b_{X,t+1}}{\lambda^U b_{Q,t+1}}\right) c_t + \frac{k^{\Theta}_{t+1} K^U_{X,t+1}}{\lambda^I a_{Q,t+1}} (1 - c_t), \tag{2.27}$$

and

$$c_t = \frac{(c^I_t)^{-1}}{(c^I_t)^{-1} + (c^U_t)^{-1}} \in (0, 1), \tag{2.28}$$

where

$$c^I_t = \lambda^I a^2_{Q,t+1} \Xi^I_t, \tag{2.29}$$

$$c^U_t = \lambda^U b^2_{Q,t+1} \Xi^U_t. \tag{2.30}$$

Proof. See Appendix 2.8.2. □

This proposition verifies the Conjecture 1 that the equilibrium price is linear in state variables.

2.4 Equilibrium with Oligopolistic Market Makers

In this section, we study the case where the population of market makers is not large enough to neglect market makers' market power. For example, in [46], there is a monopolistic market maker who can indirectly set both bid and ask prices through the quantities they buy and sell. It is well known that price has a dual role: information revealing and market clearing. In above section, we separate price's information role from its market clearing role. Since market makers have no private information, their trading brings no new information into prices. Therefore, the information role stays the same as in competitive case.

Lemma 4. *In a linear equilibrium, the ask price A_t and bid price B_t are informational equivalent to $s_t = v_t - h_t X$ or $S_t = \hat{V}_t^I - \mu_t X$: $\{\underline{A}_t, \underline{B}_t\} \Leftrightarrow \{\underline{s}_t\} \Leftrightarrow \{\underline{S}_t\}$.*

This lemma implies that although there are two prices (ask and bid), both of the prices contain the same linear combination of the two unknowns (informed traders' private information and liquidity shock), i.e. both ask and bid prices reveal the same information. Therefore, state variables $\{\Phi_t\}$ also follows a Gaussian Markov process under information process generated by $\mathcal{F}_t^I$ and $\mathcal{F}_t^U$ respectively.

Now we start to solve traders' optimal stock holding problem. Define P_t^i as the price at which type $i \in \{I, U\}$ traders are trading at time t. For example, for informed and uninformed traders i.e. $i \in \{I, U\}$, $P_t^i = A_t$ if they buy at t, and $P_t^i = B_t$ if they sell at t. We can always write informed and uninformed traders' trading strategies in terms of the difference between their private value P_t^{iR} and their trading price P_t^i as in [46]:

$$N_i \theta_t^i = N_i \theta_{t-1}^i + \frac{P_t^{iR} - P_t^i}{c_t^i} \tag{2.31}$$

for $i \in \{I, U\}$, and where P_t^{iR} is the reservation price of type i traders at time t, i.e. type i traders don't trade at P_t^{iR}. Define

$$\Delta_t = P_t^{IR} - P_t^{UR} \tag{2.32}$$

as the reservation price difference at time t. Reservation prices are revealed to market makers. Since they are infinite risk averse, they hold no inventory and bear no risk. They always set ask and bid prices through their trading quantities without change informed and uninformed traders' trading direction in competitive case. We denote $\alpha_t^{M,j}$ and $\beta_t^{M,j}$ as the quantities that market maker $j \in \{1, 2, ..., N_M\}$ chooses to buy and sell at time t. We also assume that at initial time $t = 0$ market makers hold on inventory, i.e., $\theta_0^M = 0$, due to their infinite risk aversion.

Lemma 5. *The optimal quantities that market maker j buys and sells at time $t = 1, 2, 3, ..., T-1$ are*

$$\alpha_t^{M,j} = \beta_t^{M,j} = \frac{1}{N_M + 1} \frac{|\Delta_t|}{c_t^I + c_t^U}. \tag{2.33}$$

The equilibrium bid and ask prices are respectively

$$A_t = P_t^{UR} + \frac{N_M}{N_M + 1} c_t \Delta_t + \frac{\Delta_t^+}{N_M + 1} = P_t^{IR} - \frac{N_M}{N_M + 1}(1 - c_t)\Delta_t + \frac{\Delta_t^-}{N_M + 1}, \tag{2.34}$$

$$B_t = P_t^{UR} + \frac{N_M}{N_M + 1} c_t \Delta_t - \frac{\Delta_t^-}{N_M + 1} = P_t^{IR} - \frac{N_M}{N_M + 1}(1 - c_t)\Delta_t - \frac{\Delta_t^+}{N_M + 1}. \tag{2.35}$$

The bid-ask spread is $A_t - B_t = \frac{|\Delta_t|}{N_M + 1}$.

Proof. See Appendix 2.8.3. □

Now we start to solve informed and uninformed traders' optimal stock holding. Let $Q_{t+1}^i = P_{t+1}^i - P_t^i$ $(t = 1, 2, ..., T-2)$ and $Q_T^i = V - P_{T-1}^i$ be the excess return on one share of the stock for type $i \in \{I, U\}$ traders. An investor's optimal stock holding at each date is given by the solution to the following optimization problem:

$$\max_{\theta_t^U} E[-e^{-\lambda^U N_U W_T^U} | \mathcal{F}_t^U] \tag{2.36}$$

$$s.t. \quad W_{t+1}^U = W_t^U + \theta_t^U Q_{t+1}^U$$

$$\max_{\theta_t^I} E[-e^{-\lambda^I N_I W_T^I} | \mathcal{F}_t^I] \tag{2.37}$$

$$s.t. \quad W_{t+1}^I = W_t^I + \theta_t^I Q_{t+1}^I, \quad t \leq T-2$$

$$and \quad W_T^I = W_{T-1}^I + \theta_{T-1}^I Q_T^I + XN/N_I.$$

Their excess return depends on whether they buy or sell, i.e. depends on the signs of Δ_{t+1} and Δ_t. For example, if $\Delta_t > 0$ and $\Delta_{t+1} < 0$, then informed traders buy at t and sell at $t+1$, hence $P_t^I = A_t$ and $P_{t+1}^I = B_{t+1}$. Therefore, $Q_{t+1}^I = B_{t+1} - A_t$ in this example. By examining the four cases, it interestingly turns out that

$$Q_{t+1}^I = (P_{t+1}^{IR} - P_t^{IR}) - \frac{N_M}{N_M+1}[(1-c_{t+1})\Delta_{t+1} - (1-c_t)\Delta_t], \tag{2.38}$$

$$Q_{t+1}^U = (P_{t+1}^{UR} - P_t^{UR}) + \frac{N_M}{N_M+1}(c_{t+1}\Delta_{t+1} - c_t\Delta_t), \tag{2.39}$$

$$Q_{t+1}^I - Q_{t+1}^U = \frac{1}{N_M+1}(\Delta_{t+1} - \Delta_t), \tag{2.40}$$

which means the form of the excess return Q_t^i nicely does not depend on the sign of Δ_t. As $N_M \to \infty$, $Q_t^I = Q_t^U = Q_t$, where Q_t is the excess return in competitive case. The reservations prices are revealed to all traders. We conjecture that for $t = 1, 2, 3, ..., T-1$

$$P_t^{IR} = \mathbf{L}[\Phi_t] - c_t^I N_I \theta_{t-1}^I, \tag{2.41}$$

$$P_t^{UR} = \mathbf{L}[(\hat{X}_t^U + \Theta)] - c_t^U N_U \theta_{t-1}^U. \tag{2.42}$$

This conjecture implies the following one.

Conjecture 2. Q_t^I and Q_t^U, measurable with respect to $\mathcal{F}_t^I$ and $\mathcal{F}_t^U$ respectively, are Gaussian process under information $\{\mathcal{F}_t^I\}$ and $\{\mathcal{F}_t^U\}$ respectively:

$$Q_{t+1}^I = \mathbf{L}[\Phi_t, e_{t+1}^I] - \frac{c_{t+1}^I}{N_M+1} N_I \theta_t^I + \frac{c_t^I}{N_M+1} N_I \theta_{t-1}^I, \tag{2.43}$$

$$Q_{t+1}^U = \mathbf{L}[(\hat{X}_t^U + \Theta), e_{t+1}^U] - \frac{c_{t+1}^U}{N_M+1} N_U \theta_t^U + \frac{c_t^U}{N_M+1} N_U \theta_{t-1}^U. \tag{2.44}$$

As $N_M \to \infty$, $Q_t^I = Q_t^U = Q_t$, and they don't depend on previous inventories.

With the above lemma and conjecture in this section, we can obtain our main result which is presented in the following proposition:

Proposition 18. *In a linear equilibrium of the economy, the price function has the following form: for* $t = 1, 2, 3, ..., T-1$,

$$A_t = \max\left\{P_t + \frac{c_t^I}{N_M+1}\left[(1-\gamma_t)(\hat{X}_t^U + \Theta) - N_U\theta_{t-1}^U\right], P_t - \frac{c_t^U}{N_M+1}\left[(1-\gamma_t)(\hat{X}_t^U + \Theta) - N_U\theta_{t-1}^U\right]\right\}$$

$$B_t = \min\left\{P_t + \frac{c_t^I}{N_M+1}\left[(1-\gamma_t)(\hat{X}_t^U + \Theta) - N_U\theta_{t-1}^U\right], P_t - \frac{c_t^U}{N_M+1}\left[(1-\gamma_t)(\hat{X}_t^U + \Theta) - N_U\theta_{t-1}^U\right]\right\}$$

The bid-ask spread is $A_t - B_t = \frac{|\Delta_t|}{N_M+1}$, *where*

$$\Delta_t = (c_t^I + c_t^U)\left[(1-\gamma_t)(\hat{X}_t^U + \Theta) - N_U\theta_{t-1}^U\right]. \tag{2.45}$$

The equilibrium stock holdings are

$$N_I\theta_t^I = \frac{N_M}{N_M+1}[\gamma_t(\hat{X}_t^U + \Theta) - \hat{X}_t^U] + \frac{1}{N_M+1}N_I\theta_{t-1}^I, \tag{2.46}$$

$$N_U\theta_t^U = \frac{N_M}{N_M+1}(1-\gamma_t)(\hat{X}_t^U + \Theta) + \frac{1}{N_M+1}N_U\theta_{t-1}^U. \tag{2.47}$$

Proof. See Appendix 2.8.4. □

This proposition verifies the Conjecture 2 that the excess returns are also linear in state variables and previous inventories.

2.5 Dynamics of Trading Volume and Bid-Ask Spread

We now examine in more detail the behavior of trading volume and its relation to bid-ask spread and price volatility under different specifications of the information flow in the market.

2.5.1 Trading Volume

In our model, there are three types of traders. Since market makers are assumed to hold zero inventory over time, all the trades are actually exchanged between informed and uninformed traders. The trading volume at time t, denoted by Vol_t, can be calculated by Eq. (2.33)

$$\mathrm{Vol}_t = \frac{N_M}{N_M+1}\frac{|\Delta_t|}{c_t^I + c_t^U}, \tag{2.48}$$

where we can rewrite Δ_t as

$$\Delta_t = (c_t^I + c_t^U)\left[(1-\gamma_t)(\hat{X}_t^U + \Theta) - N_M \sum_{n=1}^{t-1} \frac{(1-\gamma_{t-n})(\hat{X}_{t-n}^U + \Theta)}{(N_M+1)^n} - \frac{N_U \theta_0^U}{(N_M+1)^{t-1}}\right]. \tag{2.49}$$

Obviously, $\Delta_t \sim \mathcal{N}\left((c_t^I + c_t^U)\mu_t^\Delta, (c_t^I + c_t^U)^2(\sigma_t^\Delta)^2\right)$, where

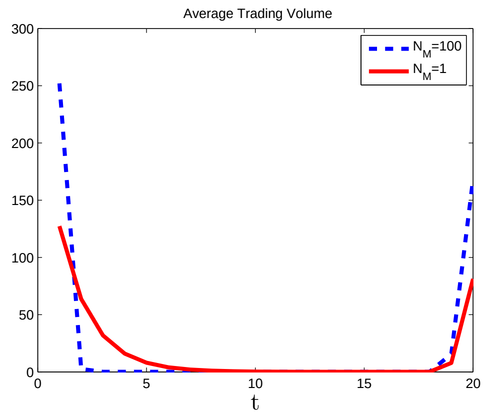
$$\mu_t^\Delta = \Theta(1-\gamma_t) - \Theta N_M \sum_{n=1}^{t-1} \frac{1-\gamma_{t-n}}{(N_M+1)^n} - \frac{N_U \theta_0^U}{(N_M+1)^{t-1}}, \tag{2.50}$$

and σ_t^Δ is given in Appendix 2.8.5. Thus, the average trading volume is

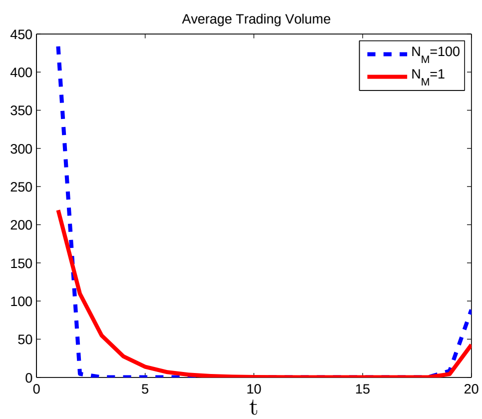
$$\overline{\mathrm{Vol}}_t = \frac{N_M}{N_M+1}\frac{\mathrm{E}\left[|\Delta_t|\right]}{c_t^I + c_t^U} = \frac{N_M}{N_M+1}\left[\sqrt{\frac{2}{\pi}}\sigma_t^\Delta e^{-\frac{1}{2}(\frac{\mu_t^\Delta}{\sigma_t^\Delta})^2} + \mu_t^\Delta\left(1 - 2\mathbf{\Phi}\left(-\frac{\mu_t^\Delta}{\sigma_t^\Delta}\right)\right)\right], \tag{2.51}$$

where $\mathbf{\Phi}$ is normal cumulative distribution function. Therefore, we can compute the average trading volume for each period.

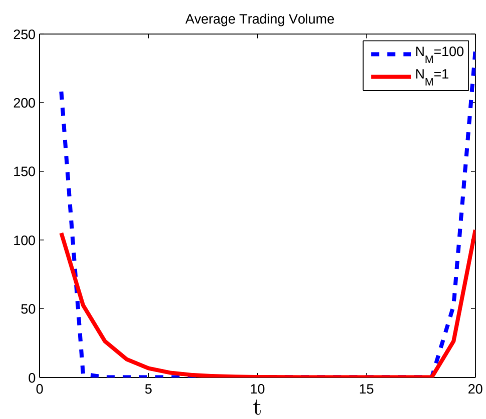
When the information only comes to the market at the opening, the average trading volume exhibits a U-shape pattern throughout the trading duration, i.e., high trading volume happens at the beginning and at the end. The high trading volume at the beginning comes from the information arrival. As Figure 2.1(b) shows, when there is more information (high $\tau_{\varepsilon,1}/\tau_V$), the trading volume is higher. The high trading volume at the end comes from the realization of the liquidity shock for informed traders. The more volatile the shock is, the higher the trading volume at the market closure as Figure 2.1(c) shows. The red lines in Figure 1. show that when market makers have significant market power, traders will smooth out their trading even though they have no market power. Hence, there are cases the average trading volume can be higher in a market with oligopolistic market makers than in a market with competitive market makers.



(a) $\tau_V = 1$, $\tau_X = 1$

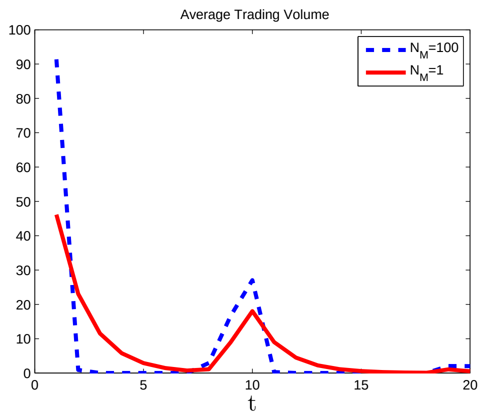


(b) $\tau_V = 0.1$, $\tau_X = 1$

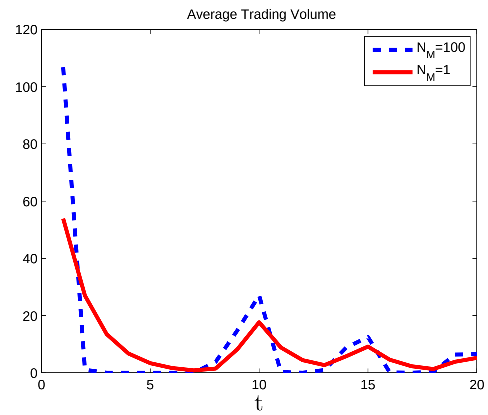


(c) $\tau_V = 1$, $\tau_X = 0.1$

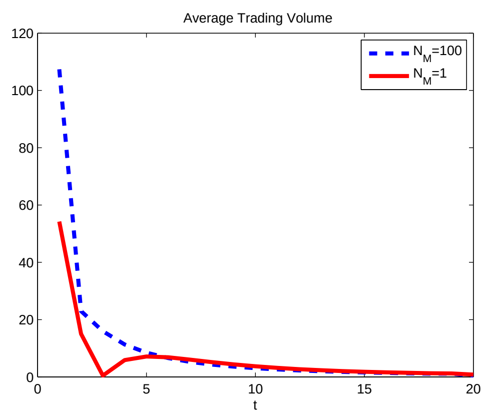
Figure 2.1: Parameters are $\lambda^I = \lambda^U = 1$, $\Theta = 1000$, $N_I\theta_0^I = N_U\theta_0^U = \Theta/2$, $\bar{V} = 100$, $\bar{X} = 0$, and $\tau_{\varepsilon,t} = 0$ except for $\tau_{\varepsilon,1} = 1$.



(a) $\tau_{\varepsilon,10} = 1$



(b) $\tau_{\varepsilon,10} = \tau_{\varepsilon,15} = \tau_{\varepsilon,20} = 1$



(c) $\tau_{\varepsilon,t} = 10\tau_V/(T-1)$

Figure 2.2: Parameters are $\lambda^I = \lambda^U = 1$, $\Theta = 1000$, $N_I\theta_0^I = N_U\theta_0^U = \Theta/2$, $\bar{V} = 100$, $\bar{X} = 0$, $\tau_V = 1$, $\tau_X = 1$, and $\tau_{\varepsilon,t} = 0$ except for $\tau_{\varepsilon,1} = 1$ and other certain dates as indicated.

As Figure 2.2 shows, the average trading volume is very sensitive to information flow. We focus on two common types of information flows. First, as in Figure 2.2(a) and (b), the information only comes at certain points of time. Once there is information coming to the market, the trading volume will spike. Thus, the information has a ripple effect on average trading volume. This information ripple effect is more strong in the competitive case. However, when there are only a few market makers in the market, the market power of market makers lead traders to smooth their orders, which weakens the information ripple effect on trading volume. Second, the information comes to the market at a constant rate as Figure 2.2(c) shows. In this case, the trading volume goes to zero after several rounds of trading, and then renounces to the level which is slightly higher than the one in a competitive market. In competitive markets the trading volume only depends on uninformed traders' expected value of X, thus the trading volume decreases monotonically with the number of trading rounds as information gradually comes to the market. However, in monopoly markets trading volume also depends on the inventory level, thus after several rounds of trading the expected liquidity shock reaches the inventory level, both informed and uninformed have the same reservation value for the stock, resulting in no trading volume.

2.5.2 Dynamics of Bid-Ask Spread

This model allows us to examine the dynamic of both average trading volume and bid-ask spread. By Eq. (2.45) we can compute the average spread as

$$\overline{A_t - B_t} = \frac{1}{N_M + 1} \mathrm{E}\left[|\Delta_t|\right], \tag{2.52}$$

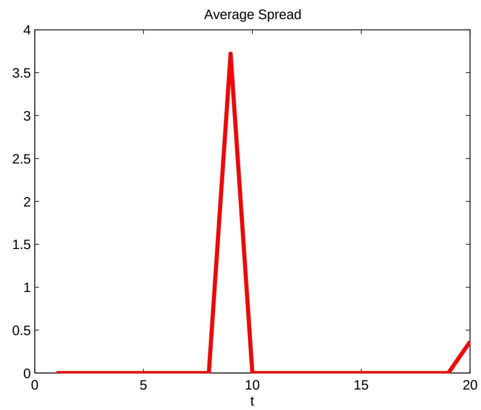
Thus, the average bid-ask spread is perfectly correlated with the average trading volume. The average bid-ask spread, hence, also exhibits a U-shape pattern throughout the trading duration with large spread at the market opening and closure. However, when there is information coming in throughout the trading duration, the average spread exhibits different patterns from average trading volume. This is because average spread not only depends on the state variable $\hat{X}_t^U$ and inventory, but also depends on $c_t^I + c_t^U$ which serves a multiplier. $c_t^I + c_t^U$ depends on how much weight traders put on the next period's information. Unsurprisingly, $c_t^I + c_t^U$ dramatically increases when the next period's information is very accurate. Therefore, when information comes at certain points of the time, the average spread spikes one period before the information arrival as Figure 2.3(a) and (b) show. When information comes at a constant rate, the average spread exhibits a similar pattern as the average trading volume, as Figure 2.3(c) shows. This is because in this type of information environment, $c_t^I + c_t^U$ changes gradually, so the variation of the average spread mostly comes from the change of state variable and inventory.

Another often used quantification of spread is the relative spread with respect to the price level $\frac{\overline{A_t - B_t}}{\mathrm{E}[|P_t|]}$, which we discuss in the Appendix 2.8.5.

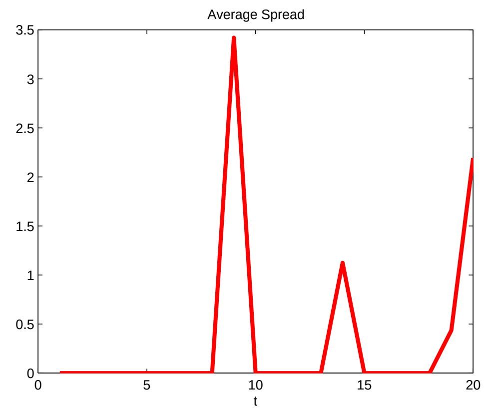
2.6 Market Makers' Profits

Since we shut down the inventory risk of market makers, market makers' profits purely come from making the market instead of speculating by carrying inventory. The total profit for each market maker in a N_M-market-maker market with T trading periods is

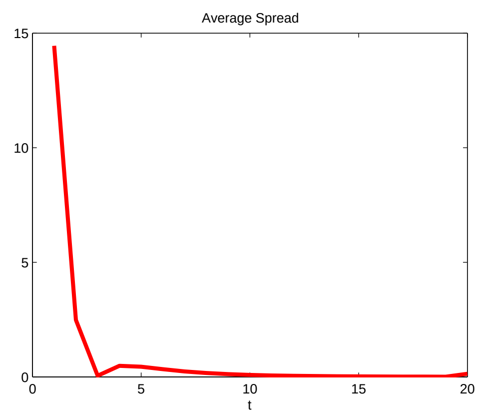
$$\mathit{Profit}(N_M, T) = \frac{1}{(N_M + 1)^2} \sum_{t=1}^{T-1} \frac{|\Delta_t|^2}{c_t^I + c_t^U}. \tag{2.53}$$



(a) $\tau_{\varepsilon,10} = 1$

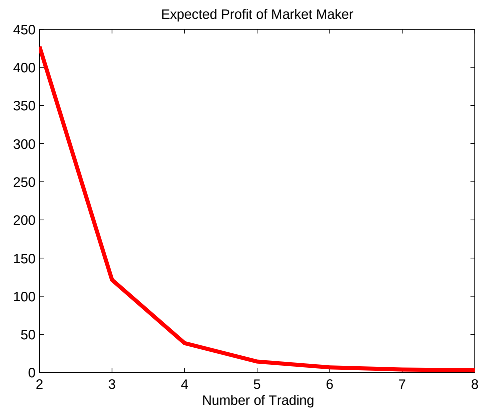


(b) $\tau_{\varepsilon,10} = \tau_{\varepsilon,15} = \tau_{\varepsilon,20} = 1$

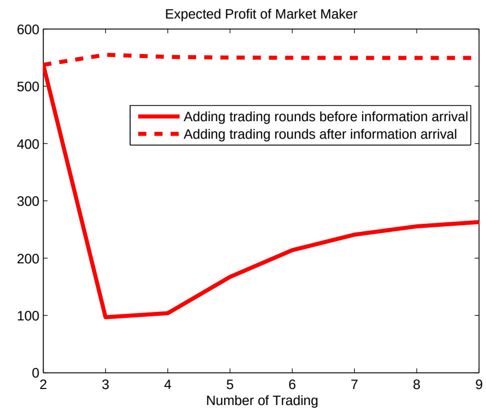


(c) $\tau_{\varepsilon,t} = 10\tau_V/(T-1)$

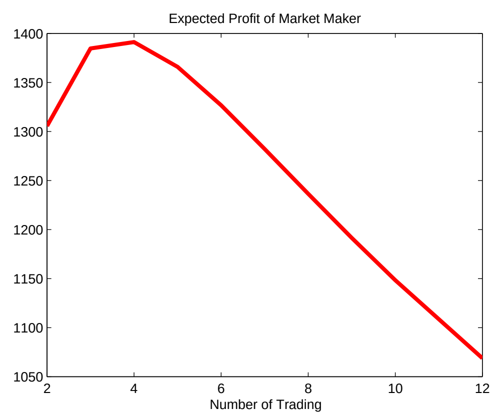
Figure 2.3: Parameters are $\lambda^I = \lambda^U = 1$, $\Theta = 1000$, $N_I\theta_0^I = N_U\theta_0^U = \Theta/2$, $\bar{V} = 100$, $\bar{X} = 0$, $\tau_V = 1$, $\tau_X = 1$, and $\tau_{\varepsilon,t} = 0$ except for $\tau_{\varepsilon,1} = 1$ and other certain dates as indicated.



(a) No information after opening



(b) Information arrives at certain time



(c) Information arrives at a constant rate

Figure 2.4: Parameters are $\lambda^I = \lambda^U = 1$, $\Theta = 1000$, $N_I \theta_0^I = N_U \theta_0^U = \Theta/2$, $\bar{V} = 100$, $\bar{X} = 0$, $\tau_V = 1$, $\tau_X = 1$, and $\tau_{\varepsilon,t} = 0$ except for $\tau_{\varepsilon,1} = 1$ and other certain dates as indicated.

Thus, the expected profit is

$$\mathrm{E}[Profit(N_M, T)] = \frac{1}{(N_M+1)^2} \sum_{t=1}^{T-1} (c_t^I + c_t^U) \left[\left(\mu_t^\Delta\right)^2 + \left(\sigma_t^\Delta\right)^2 \right]. \tag{2.54}$$

Obviously, reducing the number of market makers always benefit market makers. Thus one way to increase market makers' profit is to increase their trading speed to gain market power. There is existing literature studying the trading speed racing among market makers. In this paper, we focus on homogeneous speed to examine whether increasing trading frequency can benefit market makers without giving them more market power. For simplicity, we study the monopoly market with only one market maker. Now we change the number of trading rounds T but fixing the information environment in three common ways as follows. In Figure 2.4(a), the information only arrives at the beginning, i.e., $\tau_{\varepsilon,t} = 0$ except for $t = 1$. In this case, adding trading rounds make market makers less profitable. This is because informed and uninformed traders tend to shred their orders in monopoly market to reduce the transaction cost. More trading rounds, narrower the average spread, thus the market maker make less profit. In Figure 2.4(b), there is information arrives at certain point of time. For example, the whole trading duration is one day, and in the middle of the day there is an accoutrement, then whether adding trading rounds before the information or after the information can benefit the market maker? As Figure 2.4(b) shows, adding trading rounds before the information arrival can make the market maker more profitable. In Figure 2.4(c), information comes at a constant rate. In this information environment, as the number of trading rounds increases, the market maker's profit increases at first, and then falls down.

As Figure 2.4 shows, a monopolistic market maker does not necessarily prefer make the market more frequently. This may explain why in some OTC markets, for example, in municipal bond markets where the market makers have significant market power, the trading frequency is very low.

So far we study the trading frequency from the market maker's standing point. In the future work, we will study how trading frequency affects informed and uninformed traders' welfare or a social plan's objective function to try to understand whether continuous trading can make all market participants better off.

2.7 Concluding Remarks

In this paper, we develop a multi-period model of stock trading in which market markers have significant market power to set bid and ask prices in the presence of information asymmetry and heterogeneous hedging demands. We show that the volume pattern over time is closely related to the flow and the nature of the information.

If private information mainly arrives at the opening, both trading volume and bid-ask spread exhibit U-shape patterns, consistent with empirical findings. Traders tend to trade more aggressively with more information arrivals. As a result, market makers charge high bid-ask spreads and earn higher profits. In addition, we find that the market power of market makers tends to dampen trading spikes due to new information arrivals. Our model generates several new testable empirical predictions. For example, our model predicts that the variation of trading volume tends to be smaller for stocks with less competition during periods of news events. Our model is very flexible and can be extended to study the equilibrium when there is residual noise in the fundamental value or when there is new liquidity shock every period.

2.8 Appendix

2.8.1 Proof of Lemma 3

Proof. To derive the filtering equations, we use the results in the following lemma, the proof of which can be found in [45].

Lemma 6. *Let*

$$x_t = A_t x_{t-1} + B_t \varepsilon_{x,t}, \quad y_t = H_t x_t + \varepsilon_{y,t}, \quad t = 1, 2, ...$$

x_t is the n-vector of state variables at t, y_t is the m-vector of observations at t. A_t, B_t and H_t are , respectively, $(n \times n)$, $(n \times k)$, $(m \times n)$ constant matrices. $\{\varepsilon_{x,t}, t = 1, 2, ...\}$ and $\{\varepsilon_{y,t}, t = 1, 2, ...\}$ are respectively a k-vector and an m-vector white Gaussian sequence. $\varepsilon_{x,t} \sim \mathcal{N}(0, Q_t)$, $\varepsilon_{y,t} \sim \mathcal{N}(0, R_t)$, and $x_0 \sim \mathcal{N}(\bar{x}_0, \Sigma_{x,0})$. x_0, $\{\varepsilon_{x,t}\}$ and $\{\varepsilon_{y,t}\}$ are independent. Let

$$\begin{aligned} \hat{x}_t &= \hat{x}_{t|t} = \mathrm{E}[x_t | y_\tau : 1 \leq \tau \leq t], \\ O_t &= O_{t,t} = \mathrm{E}[(x_t - \hat{x}_t)(x_t - \hat{x}_t)^\top | y_\tau : 1 \leq \tau \leq t]. \end{aligned}$$

Then,

$$\begin{aligned} \hat{x}_t &= A_t \hat{x}_{t-1} + K_t (y_t - H_t A_t \hat{x}_{t-1}), \\ O_t &= (I_n - K_t H_t)(A_t O_{t-1} A_t^\top + B_t Q_t B_t^\top), \\ K_t &= (A_t O_{t-1} A_t^\top + B_t Q_t B_t^\top) H_t^\top [H_t (A_t O_{t-1} A_t^\top + B_t Q_t B_t^\top) H_t^\top + R_t]^{-1}. \end{aligned}$$

where I_n is the $(n \times n)$ identity matrix.

We can now solve for the informed filters $\hat{V}_t^I$ by applying this lemma. Make the following substitution:

$$x_t = x_0 = V, \quad \varepsilon_{x,t} = 0; \quad y_t = v_t, \quad \varepsilon_{y,t} = \varepsilon_t. \tag{2.55}$$

so the constant matrices are

$$A_t = 1, \quad H_t = 1, \quad Q_t = 0, \quad R_t = \tau_{\varepsilon,t}^{-1}. \tag{2.56}$$

By definition

$$\hat{x}_t = \hat{V}_t^I, \quad O_t = o_{V,t}^I,$$

so by the above lemma, we have $o_{V,t}^I = (1 - K_t^I) o_{V,t-1}^I$ and

$$\frac{1}{o_{V,t}^I} = \frac{1}{o_{V,t-1}^I} + \tau_{\varepsilon,t}, \tag{2.57}$$

$$K_t^I = \frac{o_{V,t-1}^I}{o_{V,t-1}^I + \tau_{\varepsilon,t}^{-1}} = o_{V,t}^I \tau_{\varepsilon,t}, \tag{2.58}$$

i.e., K_t^I can be expressed in terms of $o_{V,t-1}^I$ and $\{o_{V,t}^I\}$ can be expressed recursively with the initial value $o_{V,0}^I = \tau_V^{-1}$. Now we can express informed investors' expectation as follows:

$$\hat{V}_t^I = \hat{V}_{t-1}^I + K_t^I (v_t - \hat{V}_{t-1}^I). \tag{2.59}$$

We can now solve for the uninformed filters $\hat{V}_t^U$ and $\hat{X}_t^U$ by applying this lemma. Make the following substitution:

$$x_t = x_0 = \begin{pmatrix} V \\ X \end{pmatrix}, \quad \varepsilon_{x,t} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}; \quad y_t = s_t, \quad \varepsilon_{y,t} = \varepsilon_t. \tag{2.60}$$

The constant matrices are

$$A_t = I_2 \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad H_t = \begin{pmatrix} 1 & -h_t \end{pmatrix}, \quad Q_t = 0, \quad R_t = \tau_{\varepsilon,t}^{-1}. \tag{2.61}$$

Since $s_t = v_t - h_t X = V - h_t X + \varepsilon_t$, by definition

$$\hat{x}_t = \begin{pmatrix} \hat{V}_t^U \\ \hat{X}_t^U \end{pmatrix}, \quad O_t = \begin{pmatrix} o_{V,t}^U & o_{VX,t}^U \\ o_{VX,t}^U & o_{X,t}^U \end{pmatrix},$$

where $o_{VX,t}^U = \mathrm{Cov}_t^U(V, X)$. Therefore, by the above lemma,

$$K_t^U = \frac{1}{o_{t-1}^U} O_{t-1} \begin{pmatrix} 1 \\ -h_t \end{pmatrix} = \frac{1}{o_{t-1}^U} \begin{pmatrix} o_{V,t-1}^U - h_t o_{VX,t-1}^U \\ o_{VX,t-1}^U - h_t o_{X,t-1}^U \end{pmatrix}, \tag{2.62}$$

$$O_t = O_{t-1} - \frac{1}{o_{t-1}^U} O_{t-1} \begin{pmatrix} 1 & -h_t \\ -h_t & h_t^2 \end{pmatrix} O_{t-1}, \tag{2.63}$$

where o_{t-1}^U is a scalar determined by O_{t-1} and h_t

$$o_{t-1}^U = H_t O_{t-1} H_t^\top + R_t = o_{V,t-1}^U - 2h_t o_{VX,t-1}^U + h_t^2 o_{X,t-1}^U + \tau_{\varepsilon,t}^{-1} \equiv \mathrm{Var}_{t-1}^U(s_t). \tag{2.64}$$

Thus, the elements of O_t can be determined by the elements of O_{t-1}:

$$\begin{aligned} o_{V,t}^U &= o_{V,t-1}^U - \frac{1}{o_{t-1}^U}(o_{V,t-1}^U - h_t o_{VX,t-1}^U)^2, \\ o_{X,t}^U &= o_{X,t-1}^U - \frac{1}{o_{t-1}^U}(o_{VX,t-1}^U - h_t o_{X,t-1}^U)^2, \\ o_{VX,t}^U &= o_{VX,t-1}^U - \frac{1}{o_{t-1}^U}(o_{V,t-1}^U - h_t o_{VX,t-1}^U)(o_{VX,t-1}^U - h_t o_{X,t-1}^U). \end{aligned}$$

Reversely, O_{t-1} can be determined by O_t as

$$o_{V,t-1}^U = \frac{[o_{V,t}^U o_{X,t}^U - (o_{VX,t}^U)^2] h_t^2 - o_{V,t}^U \tau_{\varepsilon,t}^{-1}}{o_{V,t}^U - 2o_{VX,t}^U h_t + o_{X,t}^U h_t^2 - \tau_{\varepsilon,t}^{-1}}, \tag{2.65}$$

$$o_{X,t-1}^U = \frac{o_{V,t}^U o_{X,t}^U - (o_{VX,t}^U)^2 - o_{X,t}^U \tau_{\varepsilon,t}^{-1}}{o_{V,t}^U - 2o_{VX,t}^U h_t + o_{X,t}^U h_t^2 - \tau_{\varepsilon,t}^{-1}}, \tag{2.66}$$

$$o_{VX,t-1}^U = \frac{[o_{V,t}^U o_{X,t}^U - (o_{VX,t}^U)^2] h_t - o_{VX,t}^U \tau_{\varepsilon,t}^{-1}}{o_{V,t}^U - 2o_{VX,t}^U h_t + o_{X,t}^U h_t^2 - \tau_{\varepsilon,t}^{-1}}. \tag{2.67}$$

Thus, if we take a guess of O^U_{T-1}, then all the $\{O^U_t\}$ for $t < T-1$ can be computed recursively, and $\{O^U_{T-1}\}$ can be pinned down recursively with initial value $O_0 = \begin{pmatrix} \tau_V^{-1} & 0 \\ 0 & \tau_X^{-1} \end{pmatrix}$.

Moreover,

$$o^U_{t-1} \equiv \mathrm{Var}^U_{t-1}(s_t) = \frac{\tau_{\varepsilon,t}^{-2}}{\tau_{\varepsilon,t}^{-1} - (o^U_{V,t} - 2o^U_{VX,t}h_t + o^U_{X,t}h_t^2)}, \tag{2.68}$$

which can be proved positive.

Similarly, $K^U_t \equiv \begin{pmatrix} K^U_{V,t} \\ K^U_{X,t} \end{pmatrix}$ can be expressed by O^U_t:

$$\begin{aligned} K^U_{V,t} &= \frac{o^U_{V,t} - h_t o^U_{VX,t}}{\tau_{\varepsilon,t}^{-1}} = o^U_{V,t}\tau_{\varepsilon,t} - (\frac{\mu_t}{o^I_{V,t}} - \frac{\mu_{t-1}}{o^I_{V,t-1}})o^U_{VX,t}, && (2.69) \\ K^U_{X,t} &= \frac{o^U_{VX,t} - h_t o^U_{X,t}}{\tau_{\varepsilon,t}^{-1}} = o^U_{VX,t}\tau_{\varepsilon,t} - (\frac{\mu_t}{o^I_{V,t}} - \frac{\mu_{t-1}}{o^I_{V,t-1}})o^U_{X,t}, && (2.70) \\ K^I_t &= o^I_{V,t}\tau_{\varepsilon,t}. && (2.71) \end{aligned}$$

Thus, $\{O^U_t\}$ and $\{K^U_t\}$ for $t < T-1$ can be determined by the initial guess of O^U_{T-1} and $\{h_t\}$, and so does

$$\begin{pmatrix} \hat{V}^U_t \\ \hat{X}^U_t \end{pmatrix} = \begin{pmatrix} \hat{V}^U_{t-1} \\ \hat{X}^U_{t-1} \end{pmatrix} + \begin{pmatrix} K^U_{V,t} \\ K^U_{X,t} \end{pmatrix} [s_t - (\hat{V}^U_{t-1} - h_t \hat{X}^U_{t-1})]. \tag{2.72}$$

□

2.8.2 Proof of Proposition 16 and 17

Proof. To derive the optimal holdings and price, we use the following lemma.

Lemma 7. *Let u be an $n \times 1$ normal vector with mean $\bar{u}$ and covariance matrix Σ, A a scalar, B an $n \times 1$ vector, C an $n \times n$ symmetric matrix, I the $n \times n$ identity matrix, and*

$|M|$ the determinant of a matrix M. Then,

$$\mathrm{E}_u \exp\{-\rho[A + B^\top u + \frac{1}{2}u^\top C u]\} = \frac{1}{\sqrt{|I + \rho C \Sigma|}} \exp\big\{ -\rho[A + B^\top \bar{u} + \frac{1}{2}\bar{u}^\top C \bar{u} \\ -\frac{1}{2}\rho(B + C\bar{u})^\top (\Sigma^{-1} + \rho C)^{-1}(B + C\bar{u})]\big\}.$$

Now we compute the optimal holdings and price for the last period. For every informed trader, his terminal wealth at $t = T$ is

$$W_T^I = W_{T-1}^I + \theta_{T-1}^I (V - P_{T-1}) + VX/N_I, \tag{2.73}$$

which is normally distributed. So his expected utility at $t = T - 1$ is

$$\mathrm{E}_{T-1}^I[W_T^I] = -e^{-\lambda^I N_I W_{T-1}^I + \lambda^I N_I \theta_{T-1}^I P_{T-1} - \lambda^I (N_I \theta_{T-1}^I + X)\hat{V}_{T-1}^I + \frac{1}{2}(\lambda^I)^2 (N_I \theta_{T-1}^I + X)^2 o_{V,T-1}^I}. \tag{2.74}$$

By the first order condition, we have

$$N_I \theta_{T-1}^I = \frac{\hat{V}_{T-1}^I - P_{T-1} - \lambda^I o_{V,T-1}^I X}{\lambda^I o_{V,T-1}^I}, \tag{2.75}$$

and the second order condition $\lambda^I o_{V,T-1}^I > 0$ is automatically satisfied. Thus, informed traders' trading reveals a mixed signal $S_{T-1} = \hat{V}_{T-1}^I - \lambda^I o_{V,T-1}^I X$, i.e., $\mu_{T-1} = \lambda^I o_{V,T-1}^I$.

Similarly, we can compute uninformed traders' optimal holding. Every uninformed trader's terminal wealth at $t = T$ is

$$W_T^U = W_{T-1}^U + \theta_{T-1}^U (V - P_{T-1}), \tag{2.76}$$

which is also normally distributed. So his expected utility at $t = T - 1$ is

$$\mathrm{E}_{T-1}^U[W_T^U] = -e^{-\lambda^U N_U W_{T-1}^U + \lambda^U N_U \theta_{T-1}^U P_{T-1} - \lambda^U N_U \theta_{T-1}^U \hat{V}_{T-1}^U + \frac{1}{2}(\lambda^U N_U \theta_{T-1}^U)^2 o_{V,T-1}^U}. \tag{2.77}$$

By the first order condition, we have

$$N_U\theta^U_{T-1} = \frac{\hat{V}^U_{T-1} - P_{T-1}}{\lambda^U o^U_{V,T-1}}, \tag{2.78}$$

and the second order condition $\lambda^U o^U_{V,T-1} >$ is automatically satisfied.

In equilibrium, by market clearing condition $N_I\theta^I_{T-1} + N_U\theta^U_{T-1} = \Theta$,

$$P_{T-1} = \omega_{T-1}(\hat{V}^I_{T-1} - \mu_{T-1}X) + (1-\omega_{T-1})\hat{V}^U_{T-1} - \mu_{T-1}\omega_{T-1}\Theta, \tag{2.79}$$

where

$$\mu_{T-1} = \lambda^I o^I_{V,T-1}, \tag{2.80}$$

$$\omega_{T-1} = \frac{(\lambda^I o^I_{V,T-1})^{-1}}{(\lambda^I o^I_{V,T-1})^{-1} + (\lambda^U o^U_{V,T-1})^{-1}}, \tag{2.81}$$

and equilibrium holdings can be computed

$$N_I\theta^I_{T-1} = \gamma_{T-1}(\hat{X}^U_{T-1} + \Theta) - \hat{X}^U_{T-1}, \tag{2.82}$$

$$N_U\theta^U_{T-1} = (1-\gamma_{T-1})(\hat{X}^U_{T-1} + \Theta), \tag{2.83}$$

where

$$\gamma_{T-1} = \omega_{T-1}. \tag{2.84}$$

Now let's compute backward. Define state vector $\Phi_t = \begin{pmatrix} \hat{X}^U_t \\ 1 \end{pmatrix}$, innovation term for informed traders $e^I_t = V - \hat{V}^I_{t-1} + \varepsilon_t$, and innovation term for uniformed traders $e^U_t = V - \hat{V}^U_{t-1} - h_t(X - \hat{X}^U_{t-1}) + \varepsilon_t$. Thus,

$$e^I_t \mid \mathcal{F}^I_{t-1} \sim \mathcal{N}(0, \Sigma^I_{t-1}),$$

$$e^U_t \mid \mathcal{F}^U_{t-1} \sim \mathcal{N}(0, \Sigma^U_{t-1}),$$

where

$$\Sigma_{t-1}^{I} = \mathrm{Var}_{t-1}^{I}(v_t) = o_{V,t-1}^{I} + \tau_{\varepsilon,t}^{-1},$$

$$\Sigma_{t-1}^{U} = \mathrm{Var}_{t-1}^{U}(s_t) = o_{V,t-1}^{U} - 2h_t o_{VX,t-1}^{U} + h_t^2 o_{X,t-1}^{U} + \tau_{\varepsilon,t}^{-1} = o_{t-1}^{U}.$$

Under the price conjuncture, define $Q_t = P_t - P_{t-1}$. Then different filtration, Φ_t and Q_t have different expressions.

Under informed traders' filtration, we have

$$\Phi_t = f_{\Phi,t}\Phi_{t-1} + a_{\Phi,t}e_t^I, \tag{2.85}$$

$$Q_t = f_{Q,t}\Phi_{t-1} + a_{Q,t}e_t^I, \tag{2.86}$$

$$\hat{V}_t^I = \hat{V}_{t-1}^I + a_{V,t}e_t^I, \tag{2.87}$$

where

$$f_{\Phi,t} = \begin{pmatrix} 1 - K_{X,t}^U(\mu_{t-1} - h_t) & K_{X,t}^U(\mu_{t-1} - h_t)X \\ 0 & 1 \end{pmatrix} \equiv \begin{pmatrix} 1 - f_t^{\Phi} & f_t^{\Phi}X \\ 0 & 1 \end{pmatrix}, \tag{2.88}$$

$$\begin{aligned} f_{Q,t} &= \begin{pmatrix} \mu_t(1-\omega_t)[1 - K_{X,t}^U(\mu_{t-1} - h_t)] - \mu_{t-1}(1-\omega_{t-1}) \\ [\mu_{t-1} - \mu_t + \mu_t(1-\omega_t)K_{X,t}^U(\mu_{t-1} - h_t)]X + [\mu_{t-1}\omega_{t-1} - \mu_t\omega_t]\Theta \end{pmatrix}^{\top} \\ &\equiv \begin{pmatrix} f_{\Theta,t}^Q - f_{X,t}^Q & f_{X,t}^Q X + f_{\Theta,t}^Q\Theta \end{pmatrix}, \end{aligned} \tag{2.89}$$

and

$$a_{\Phi,t} = \begin{pmatrix} K_{X,t}^U \\ 0 \end{pmatrix}, \tag{2.90}$$

$$a_{Q,t} = K_t^I + \mu_t(1-\omega_t)K_{X,t}^U, \tag{2.91}$$

$$a_{V,t} = K_t^I, \tag{2.92}$$

since $\hat{V}_t^I - \hat{V}_t^U = \mu_t(X - \hat{X}_t^U)$.

Under uninformed traders' filtration, we have

$$\hat{X}_t^U = \hat{X}_{t-1}^U + b_{X,t} e_t^U, \tag{2.93}$$

$$Q_t = g_{Q,t}(\hat{X}_{t-1}^U + \Theta) + b_{Q,t} e_t^U, \tag{2.94}$$

where

$$g_{Q,t} = \mu_{t-1}\omega_{t-1} - \mu_t\omega_t, \tag{2.95}$$

and

$$b_{X,t} = K_{X,t}^U, \tag{2.96}$$

$$b_{Q,t} = K_{V,t}^U - \mu_t\omega_t K_{X,t}^U. \tag{2.97}$$

Now we compute backward to $t = T-2$. For informed traders, we plug Eq. 2.82 into Eq. 2.74 and obtain

$$J_{T-1}^I = -\rho_{T-1}^I e^{-\lambda^I N_I W_{T-1}^I - \lambda^I X \hat{V}_{T-1}^I - \frac{1}{2}\Phi_{T-1}^\top H_{T-1}^I \Phi_{T-1}} \tag{2.98}$$

$$= -\rho_{T-1}^I e^{-\lambda^I N_I W_{T-2}^I - \lambda^I N_I \theta_{T-2}^I Q_{T-1} - \lambda^I X \hat{V}_{T-1}^I - \frac{1}{2}\Phi_{T-1}^\top H_{T-1}^I \Phi_{T-1}} \tag{2.99}$$

where

$$H_{T-1}^I = \frac{\mu_{T-1}^2}{o_{V,T-1}^I} \begin{pmatrix} (1-\omega_{T-1})^2 & -(1-\omega_{T-1})\omega_{T-1}\Theta \\ -(1-\omega_{T-1})\omega_{T-1}\Theta & 0 \end{pmatrix} \equiv \begin{pmatrix} k_{T-1} & -k_{T-1}^\Theta \Theta \\ -k_{T-1}^\Theta \Theta & 0 \end{pmatrix} \tag{2.100}$$

and ρ_{T-1}^I is a constant that absorbs the quadratic terms of X and Θ. The value function at

$T-1$ of informed traders can be expressed as

$$\begin{aligned} J^I_{T-1} = \quad & -\rho^I_{T-1} e^{-\lambda^I N_I W^I_{T-2} - \lambda^I N_I \theta^I_{T-2}(f_{Q,T-1}\Phi_{T-2} + a_{Q,T-1}e^I_{T-1}) - \lambda^I X(\hat{V}^I_{T-2} + a_{V,T-1}e^I_{T-1})} \\ & \times e^{-\frac{1}{2}(f_{\Phi,T-1}\Phi_{T-2} + a_{\Phi,T-1}e^I_{T-1})^\top H^I_{T-1}(f_{\Phi,T-1}\Phi_{T-2} + a_{\Phi,T-1}e^I_{T-1})} \quad (2.101) \\ = \quad & -\rho^I_{T-1} e^{-\lambda^I N_I W^I_{T-2} - \lambda^I X \hat{V}^I_{T-2} - \frac{1}{2}\Phi^\top_{T-2} f^\top_{\Phi,T-1} H^I_{T-1} f_{\Phi,T-1}\Phi_{T-2}} \\ & \times e^{-\lambda^I N_I \theta^I_{T-2} f_{Q,T-1}\Phi_{T-2} - B^I_{T-1}e^I_{T-1} - \frac{1}{2}(e^I_{T-1})^\top C^I_{T-1} e^I_{T-1}}, \quad (2.102) \end{aligned}$$

where

$$A^I_{T-1} = \Phi^\top_{T-2} f^\top_{\Phi,T-1} H^I_{T-1} a_{\Phi,T-1} + \lambda^I X a_{V,T-1}, \quad (2.103)$$

$$B^I_{T-1} = \lambda^I N_I \theta^I_{T-2} a_{Q,T-1} + A^I_{T-1}, \quad (2.104)$$

$$C^I_{T-1} = a^\top_{\Phi,T-1} H^I_{T-1} a_{\Phi,T-1} = k_{T-1}(K^U_{T-1})^2. \quad (2.105)$$

Recall that $e^I_{T-1} \mid \mathcal{F}^I_{T-2} \sim \mathcal{N}(0, \Sigma^I_{T-2})$. Thus,

$$\begin{aligned} \mathrm{E}^I_{T-2}[J^I_{T-1}] = \quad & -(\rho^I_{T-1})' e^{-\lambda^I N_I W^I_{T-2} - \lambda^I X \hat{V}^I_{T-2} - \frac{1}{2}\Phi^\top_{T-2} f^\top_{\Phi,T-1} H^I_{T-1} f_{\Phi,T-1}\Phi_{T-2}} \\ & \times e^{-\lambda^I N_I \theta^I_{T-2} f_{Q,T-1}\Phi_{T-2} + \frac{1}{2}\Xi^I_{T-2}(\lambda^I N_I \theta^I_{T-2} a_{Q,T-1} + A^I_{T-1})^2}, \quad (2.106) \end{aligned}$$

where $\Xi^I_{T-2} = ((\Sigma^I_{T-2})^{-1} + C^I_{T-1})^{-1}$, and $(\rho^I_{T-1})'$ absorbs the constants that contain the quadratic terms of X and Θ. The first order condition with respect to θ^I_{T-2} gives

$$f_{Q,T-1}\Phi_{T-2} = a_{Q,T-1}\Xi^I_{T-2} B^I_{T-1}, \quad (2.107)$$

and

$$\begin{aligned} N_I \theta^I_{T-2} \quad &= \quad [\frac{f_{Q,T-1}}{\lambda^I a^2_{Q,T-1}\Xi^I_{T-2}} - \frac{a^\top_{\Phi,T-1} H^I_{T-1} f_{\Phi,T-1}}{\lambda^I a_{Q,T-1}}]\Phi_{T-2} - \frac{a_{V,T-1}}{a_{Q,T-1}} X, \quad (2.108) \\ &\equiv \quad \gamma^I_{T-2}\hat{X}^U_{T-2} + \gamma^I_{X,T-2} X + \gamma^I_{\Theta,T-2}\Theta, \quad (2.109) \end{aligned}$$

which contains $\hat{X}^U_{T-2}$, Θ and X, and where

$$\gamma^I_{T-2} = \frac{f^Q_{\Theta,T-1} - f^Q_{X,T-1}}{\lambda^I a^2_{Q,T-1}\Xi^I_{T-2}} - \frac{(1-f^\Phi_{T-1})k_{T-1}K^U_{X,T-1}}{\lambda^I a_{Q,T-1}}, \tag{2.110}$$

$$\gamma^I_{X,T-2} = \frac{f^Q_{X,T-1}}{\lambda^I a^2_{Q,T-1}\Xi^I_{T-2}} - \frac{f^\Phi_{T-1}k_{T-1}K^U_{X,T-1}}{\lambda^I a_{Q,T-1}} - \frac{a_{V,T-1}}{a_{Q,T-1}}, \tag{2.111}$$

$$\gamma^I_{\Theta,T-2} = \frac{f^Q_{\Theta,T-1}}{\lambda^I a^2_{Q,T-1}\Xi^I_{T-2}} + \frac{k^\Theta_{T-1}K^U_{X,T-1}}{\lambda^I a_{Q,T-1}}. \tag{2.112}$$

The second order condition is $\Xi^I_{T-2} > 0$, which is automatically satisfied.

For uninformed traders, we plug Eq. 2.83 into Eq. 2.77 and obtain

$$J^U_{T-1} = -\rho^U_{T-1}e^{-\lambda^U N_U W^U_{T-1} - \frac{1}{2}H^U_{T-1}(\hat{X}^U_{T-1}+\Theta)^2} \tag{2.113}$$

$$= -\rho^U_{T-1}e^{-\lambda^U N_U W^U_{T-2} - \lambda^U N_U \theta^U_{T-2}Q_{T-1} - \frac{1}{2}H^U_{T-1}(\hat{X}^U_{T-1}+\Theta)^2}, \tag{2.114}$$

where

$$H^U_{T-1} = \frac{\mu^2_{T-1}\omega^2_{T-1}}{o^U_{V,T-1}}. \tag{2.115}$$

The value function at $T-1$ of uninformed traders can be expressed as

$$J^U_{T-1} = -\rho^U_{T-1}e^{-\lambda^U N_U W^U_{T-2} - \frac{1}{2}H^U_{T-1}(\hat{X}^U_{T-2}+\Theta)^2} \times e^{-\lambda^U N_U \theta^U_{T-2} g_{Q,T-1}(\hat{X}^U_{T-2}+\Theta) - B^U_{T-1}e^U_{T-1} - \frac{1}{2}C^U_{T-1}(e^U_{T-1})^2}, \tag{2.116}$$

where

$$A^U_{T-1} = H^U_{T-1}(\hat{X}^U_{T-2}+\Theta)b_{X,T-1}, \tag{2.117}$$

$$B^U_{T-1} = \lambda^U N_U \theta^U_{T-2} b_{Q,T-1} + A^U_{T-1}, \tag{2.118}$$

$$C^U_{T-1} = H^U_{T-1}b^2_{X,T-1}. \tag{2.119}$$

Recall that $e^U_{T-1} \mid \mathcal{F}^U_{T-2} \sim \mathcal{N}(0, \Sigma^U_{T-2})$. Thus,

$$\begin{aligned}\mathrm{E}^U_{T-2}[J^U_{T-1}] = \quad & -\rho^U_{T-2} e^{-\lambda^U N_U W^U_{T-2} - \frac{1}{2} H^U_{T-1} (\hat{X}^U_{T-2} + \Theta)^2} \\ & \times e^{-\lambda^U N_U \theta^U_{T-2} g_{Q,T-1} (\hat{X}^U_{T-2} + \Theta) + \frac{1}{2} \Xi^U_{T-2} (\lambda^U N_U \theta^U_{T-2} b_{Q,T-1} + A^U_{T-1})^2}, \end{aligned} \tag{2.120}$$

where $\Xi^U_{T-2} = ((\Sigma^U_{T-2})^{-1} + C^U_{T-1})^{-1}$. The first order condition with respect to θ^U_{T-2} gives

$$g_{Q,T-1}(\hat{X}^U_{T-2} + \Theta) = b_{Q,T-1} \Xi^U_{T-2} B^U_{T-1}, \tag{2.121}$$

and

$$N_U \theta^U_{T-2} = \gamma^U_{T-2} (\hat{X}^U_{T-2} + \Theta), \tag{2.122}$$

where

$$\gamma^U_{T-2} = \frac{g_{Q,T-1}}{\lambda^U b^2_{Q,T-1} \Xi^U_{T-2}} - \frac{H^U_{T-1} b_{X,T-1}}{\lambda^U b_{Q,T-1}}. \tag{2.123}$$

The second order condition is $\Xi^U_{T-2} > 0$, which is automatically satisfied.

Now we have best responses of informed traders Eq. (2.108) and uninformed traders Eq. (2.122), we can compute the equilibrium at $T-2$ by imposing the market clearing condition $N_I \theta^I_{T-2} + N_U \theta^U_{T-2} = \Theta$. Therefore,

$$\gamma^I_{X,T-2} = 0, \tag{2.124}$$

$$\gamma^I_{\Theta,T-2} + \gamma^U_{T-2} = 1, \tag{2.125}$$

$$\gamma^I_{T-2} + \gamma^U_{T-2} = 0. \tag{2.126}$$

The above three equations are not independent. If two of them are satisfied, then the third will be automatically satisfied. Thus, we only need to solve Eq. (2.124) and Eq. (2.125).

Eq. (2.124) leads to the solution of μ_{T-2} and h_{T-1} as follows

$$\mu_{T-2} = \lambda^I o^I_{V,T-2}, \tag{2.127}$$

$$h_{T-1} = 0. \tag{2.128}$$

Proposition 16 can be proved by following the same logic, and thus

$$\mu_{t-1} - \mu_t = K^I_t(\mu_{T-1} - h_t). \tag{2.129}$$

Eq. (2.125) leads to the solution of ω_{T-2} as follows

$$\begin{aligned} &\mu_{T-2}\omega_{T-2} - \mu_{T-1}\omega_{T-1} = g_{Q,T-1} \\ &= \left(1 + \frac{H^U_{T-1}K^U_{X,T-1}}{\lambda^U b_{Q,T-1}} - \frac{k^\Theta_{T-1}K^U_{X,T-1}}{\lambda^I a_{Q,T-1}}\right)\left(\frac{1}{\lambda^I a^2_{Q,T-1}\Xi^I_{T-2}} + \frac{1}{\lambda^U b^2_{Q,T-1}\Xi^U_{T-2}}\right)^{-1} \end{aligned} \tag{2.130}$$

where

$$\frac{1}{\Xi^I_{T-2}} = (\Sigma^I_{T-2})^{-1} + k_{T-1}(K^U_{X,T-1})^2, \tag{2.131}$$

$$\frac{1}{\Xi^I_{T-2}} = (\Sigma^U_{T-2})^{-1} + H^U_{T-1}(K^U_{X,T-1})^2, \tag{2.132}$$

and $H^I_{T-1} = \begin{pmatrix} k_{T-1} & -k^\Theta_{T-1}\Theta \\ -k^\Theta_{T-1}\Theta & 0 \end{pmatrix}$, $a_{Q,T-1} = K^I_{T-1} + \mu_{T-1}(1-\omega_{T-1})K^U_{X,T-1}$ and $b_{Q,T-1} = K^U_{V,T-1} - \mu_{T-1}\omega_{T-1}K^U_{X,T-1}$. Thus, ω_{T-2} can be calculated backward.

Define $\gamma_{T-2} = 1 - \gamma^U_{T-2} = 1 + \gamma^I_{T-2}$, then in equilibrium

$$N_I\theta^I_{T-2} = \gamma_{T-2}(\hat{X}^U_{T-2} + \Theta) - \hat{X}^U_{T-2}, \tag{2.133}$$

$$N_U\theta^U_{T-2} = (1 - \gamma_{T-2})(\hat{X}^U_{T-2} + \Theta), \tag{2.134}$$

and

$$\begin{aligned} \gamma_{T-2} &= 1 + \frac{H^U_{T-1}b_{X,T-1}}{\lambda^U b_{Q,T-1}} - \frac{\mu_{T-2}\omega_{T-2} - \mu_{T-1}\omega_{T-1}}{\lambda^U b^2_{Q,T-1}\Xi^U_{T-2}} && (2.135) \\ &= \left(1 + \frac{H^U_{T-1}b_{X,T-1}}{\lambda^U b_{Q,T-1}}\right)c_{T-2} + \frac{k^\Theta_{T-1}K^U_{X,T-1}}{\lambda^I a_{Q,T-1}}(1 - c_{T-2}), && (2.136) \end{aligned}$$

where

$$c_{T-2} = \frac{(\lambda^I a_{Q,T-1}^2 \Xi_{T-2}^I)^{-1}}{(\lambda^I a_{Q,T-1}^2 \Xi_{T-2}^I)^{-1} + (\lambda^U b_{Q,T-1}^2 \Xi_{T-2}^U)^{-1}}. \tag{2.137}$$

Now let's plug back the equilibrium holdings into the expected value functions to compute $J_{T-2}^{I,U}$. It is easy to show that for uninformed traders, their value function still has the form as

$$J_{T-2}^U = -\rho_{T-2}^U e^{-\lambda^U N_U W_{T-2}^U - \frac{1}{2} H_{T-2}^U (\hat{X}_{T-2}^U + \Theta)^2}, \tag{2.138}$$

where

$$H_{T-2}^U = H_{T-1}^U - 2\frac{H_{T-1}^U K_{X,T-1}^U g_{Q,T-1}}{b_{Q,T-1}} + \frac{g_{Q,T-1}^2}{b_{Q,T-1}^2 \Xi_{T-2}^U}. \tag{2.139}$$

Similarly, for informed traders, their value functions also has the form as

$$J_{T-2}^I = -\rho_{T-2}^I e^{-\lambda^I N_I W_{T-2}^I - \lambda^I X \hat{V}_{T-2}^I - \frac{1}{2} \Phi_{T-2}^\top H_{T-2}^I \Phi_{T-2}}, \tag{2.140}$$

where ρ_{T-2}^I absorbs all the quadratic terms of X and Θ, and H_{T-2}^I also has the form of

$$H_{T-2}^I = \begin{pmatrix} k_{T-2} & -k_{T-2}^\Theta \Theta \\ -k_{T-2}^\Theta \Theta & 0 \end{pmatrix}, \tag{2.141}$$

which has no cross term of $X\hat{X}_{T-2}^U$ because of Eq. (2.129). The elements of H_{T-2}^I can be calculated recursively by the elements of H_{T-1}^I, i.e., by k_{T-1} and k_{T-1}^Θ as follows

$$\begin{aligned} k_{T-2} &= k_{T-1}(1 - f_{T-1}^\Phi)^2 + \frac{(f_{\Theta,T-1}^Q - f_{X,T-1}^Q)^2}{a_{Q,T-1}^2 \Xi_{T-2}^I} - 2\frac{k_{T-1} K_{X,T-1}^U}{a_{Q,T-1}}(1 - f_{T-1}^\Phi)(f_{\Theta,T-1}^Q - f_{X,T-1}^Q) \\ &= k_{T-1}\left[(1 - f_{T-1}^\Phi) - \frac{K_{X,T-1}^U}{a_{Q,T-1}}(f_{\Theta,T-1}^Q - f_{X,T-1}^Q)\right]^2 + \frac{(f_{\Theta,T-1}^Q - f_{X,T-1}^Q)^2}{a_{Q,T-1}^2 \Sigma_{T-2}^I}, \end{aligned} \tag{2.142}$$

$$\begin{aligned} k_{T-2}^\Theta = \quad & k_{T-1}^\Theta (1 - f_{T-1}^\Phi) + \lambda^I (1 - \gamma_{T-2}) f_{\Theta,T-1}^Q - \lambda^I \gamma_{T-2}(f_{\Theta,T-1}^Q - f_{X,T-1}^Q) \\ & + \frac{(f_{\Theta,T-1}^Q - f_{X,T-1}^Q)^2}{a_{Q,T-1}^2 \Xi_{T-2}^I}. \end{aligned} \tag{2.143}$$

Therefore, we can compute ω_t recursively by applying Eq. (2.130), and then use computed ω_t to compute $J_t^{I,U}$. Now Proposition 17 has been fully proved.

□

2.8.3 Proof of Lemma 5

Proof. Firstly, we study the case $\Delta_t < 0$. In this case, we conjecture that I investors sell at the bid at time t, and U investors buy at the ask at time t, i.e. $P_t^{UR} > A_t > B_t > P_t^{IR}$,

$$\begin{aligned} A_t &= P_t^{UR} - c_t^U \sum_{j=1}^{N_M} \alpha_t^{M,j} \\ B_t &= P_t^{IR} + c_t^I \sum_{j=1}^{N_M} \beta_t^{M,j} \end{aligned}$$

Since market makers are infinitely risk averse, they are zero tolerance of uncertainty. The net positions of market maker should be zero, i.e. $\theta_t^{M,j} = \theta_{t-1}^{M,j} + \beta_t^{M,j} - \alpha_t^{M,j} = 0$ for $t = 1, 2, 3, ..., T-1$, where $\theta_{t-1}^{M,j} = \theta_{t-2}^{M,j} + \beta_{t-1}^{M,j} - \alpha_{t-1}^{M,j} = 0$. They maximize their profits from ask-bid by choosing $\alpha_t^{M,j}$ and $\beta_t^{M,j}$ at time t. Market maker j's problem is to maximize profit

$$\max_{\substack{\alpha_t^{M,j} \\ \alpha_t^{M,j} = \beta_t^{M,j} + \theta_1^{M,j}}} \quad \alpha_t^{M,j} A_t - \beta_t^{M,j} B_t \tag{2.144}$$

The optimal quantities each market maker chooses to buy and sell at time t are

$$\alpha_t^{M,j} = \beta_t^{M,j} = -\frac{1}{N_M + 1} \frac{\Delta_t}{c_t^I + c_t^U} \tag{2.145}$$

Therefore,

$$\begin{aligned} A_t &= P_t^{UR} + c_t \frac{N_M}{N_M + 1} \Delta_t \\ B_t &= P_t^{IR} - (1 - c_t) \frac{N_M}{N_M + 1} \Delta_t \\ A_t - B_t &= -\frac{\Delta_t}{N_M + 1} \end{aligned} \tag{2.146}$$

It can be verified that $P_t^{UR} > A_t > B_t > P_t^{IR}$ if $\Delta_t < 0$.

Secondly, if $\Delta_t > 0$, we conjecture that U investors sell at the bid at date t, and I investors buy at the ask at date t, i.e., $P_t^{IR} > A_t > B_t > P_t^{UR}$. The optimal quantities each market maker chooses to buy and sell at date t are

$$\alpha_t^{M,j} = \beta_t^{M,j} = \frac{1}{N_M+1}\frac{\Delta_t}{c_t^I + c_t^U} \tag{2.147}$$

and the ask and bid prices are

$$\begin{aligned} A_t &= P_t^{IR} - (1-c_t)\frac{N_M}{N_M+1}\Delta_t \\ B_t &= P_t^{UR} + c_t\frac{N_M}{N_M+1}\Delta_t \\ A_t - B_t &= \frac{\Delta_t}{N_M+1}. \end{aligned} \tag{2.148}$$

$\square$

2.8.4 Proof of Proposition 18

Coming soon...

2.8.5 Computation of Average Trading Volume and Bid-ask Spread

σ_t^{Δ} depends on the variance of $\hat{X}_t^U$. Since

$$\hat{X}_t^U = \mu_t^{-1}(\hat{V}_t^U - \hat{V}_t^I) + X, \tag{2.149}$$

and

$$\hat{V}_t^I = (\tau_V + \sum_{n=1}^{t}\tau_{\varepsilon,n})^{-1}\left(\bar{V}\tau_V + V\sum_{n=1}^{t}\tau_{\varepsilon,n} + \sum_{n=1}^{t}\tau_{\varepsilon,n}\varepsilon_n\right) \tag{2.150}$$

$$\hat{V}_t^U = (\tau_V + \tau_1 + \sum_{n=2}^{t}\tau_{\varepsilon,n})^{-1}\left(\bar{V}\tau_V + V(\tau_1 + \sum_{n=2}^{t}\tau_{\varepsilon,n}) + \tau_1(\varepsilon_1 - h_1 X) + \sum_{n=2}^{t}\tau_{\varepsilon,n}\varepsilon_n\right) \tag{2.151}$$

where $\frac{1}{\tau_1} = \frac{1}{\tau_{\varepsilon,1}} + \frac{h_1^2}{\tau_X}$. We can rewrite

$$\begin{aligned}
\hat{X}_t^U = \quad & \mu_t^{-1}\left((o_{V,t}^U - o_{V,t}^I)\tau_V\bar{V} + \left[o_{V,t}^U(\tau_1 + \sum_{n=2}^t \tau_{\varepsilon,n}) - o_{V,t}^I\sum_{n=1}^t \tau_{\varepsilon,n}\right]V + (o_{V,t}^U - o_{V,t}^I)\sum_{n=2}^t \tau_{\varepsilon,n}\varepsilon_n\right) \\
& + \mu_t^{-1}(o_{V,t}^U\tau_1 - o_{V,t}^I\tau_{\varepsilon,1})\varepsilon_1 + (1 - \mu_t^{-1}h_1\tau_1 o_{V,t}^U)X \qquad (2.152) \\
:= \quad & d_t\bar{V} + d_t^V V + d_t^\varepsilon\varepsilon_1 + \mu_t^{-1}(o_{V,t}^U - o_{V,t}^I)\sum_{n=2}^t \tau_{\varepsilon,n}\varepsilon_n + d_t^X X, \qquad (2.153)
\end{aligned}$$

where

$$d_t^V = \mu_t^{-1}[o_{V,t}^U(\tau_1 + \sum_{n=2}^t \tau_{\varepsilon,n}) - o_{V,t}^I\sum_{n=1}^t \tau_{\varepsilon,n}], \qquad (2.154)$$

$$d_t^\varepsilon = \mu_t^{-1}(o_{V,t}^U\tau_1 - o_{V,t}^I\tau_{\varepsilon,1}), \qquad (2.155)$$

$$d_t^X = 1 - \mu_t^{-1}h_1\tau_1 o_{V,t}^U. \qquad (2.156)$$

Since V, X and $\{\varepsilon_t\}$ are independent,

$$\begin{aligned}
(\sigma_t^\Delta)^2 &= \mathrm{Var}\left((1-\gamma_t)\hat{X}_t^U - N_M\sum_{n=1}^{t-1}\frac{(1-\gamma_{t-n})\hat{X}_{t-n}^U}{(N_M+1)^n}\right) \\
&= \left((1-\gamma_t)d_t^V - N_M\sum_{n=1}^{t-1}\frac{(1-\gamma_{t-n})d_{t-n}^V}{(N_M+1)^n}\right)^2\tau_V^{-1} + \left((1-\gamma_t)d_t^X - N_M\sum_{n=1}^{t-1}\frac{(1-\gamma_{t-n})d_{t-n}^X}{(N_M+1)^n}\right)^2\tau_X^{-1} \\
&+ \left((1-\gamma_t)\mu_t^{-1}(o_{V,t}^U - o_{V,t}^I) - N_M\sum_{n=1}^{t-2}\frac{(1-\gamma_{t-n})\mu_{t-n}^{-1}(o_{V,t-n}^U - o_{V,t-n}^I)}{(N_M+1)^n}\right)^2\tau_{\varepsilon,2} + \cdots \\
&+ \left((1-\gamma_t)\mu_t^{-1}(o_{V,t}^U - o_{V,t}^I) - \frac{N_M}{N_M+1}(1-\gamma_{t-1})\mu_{t-1}^{-1}(o_{V,t-1}^U - o_{V,t-1}^I)\right)^2\tau_{\varepsilon,t-1} \\
&+ \left((1-\gamma_t)\mu_t^{-1}(o_{V,t}^U - o_{V,t}^I)\right)^2\tau_{\varepsilon,t} + \left((1-\gamma_t)d_t^\varepsilon - N_M\sum_{n=1}^{t-1}\frac{(1-\gamma_{t-n})d_{t-n}^\varepsilon}{(N_M+1)^n}\right)^2\tau_{\varepsilon,1}^{-1} \\
&= \left((1-\gamma_t)d_t^V - N_M\sum_{n=1}^{t-1}\frac{(1-\gamma_{t-n})d_{t-n}^V}{(N_M+1)^n}\right)^2\tau_V^{-1} + \left((1-\gamma_t)d_t^X - N_M\sum_{n=1}^{t-1}\frac{(1-\gamma_{t-n})d_{t-n}^X}{(N_M+1)^n}\right)^2\tau_X^{-1} \\
&+ \sum_{s=2}^{t-1}\left((1-\gamma_t)\mu_t^{-1}(o_{V,t}^U - o_{V,t}^I) - N_M\sum_{n=1}^{t-s}\frac{(1-\gamma_{t-n})\mu_{t-n}^{-1}(o_{V,t-n}^U - o_{V,t-n}^I)}{(N_M+1)^n}\right)^2\tau_{\varepsilon,s} \\
&+ \left((1-\gamma_t)\mu_t^{-1}(o_{V,t}^U - o_{V,t}^I)\right)^2\tau_{\varepsilon,t} + \left((1-\gamma_t)d_t^\varepsilon - N_M\sum_{n=1}^{t-1}\frac{(1-\gamma_{t-n})d_{t-n}^\varepsilon}{(N_M+1)^n}\right)^2\tau_{\varepsilon,1}^{-1}. \qquad (2.157)
\end{aligned}$$

The competitive price can be rewritten as

$$P_t = p_t\bar{V} + p_t^V V + p_t^X X + p_t^\varepsilon \varepsilon_1 + \left[(1-\omega_t)o_{V,t}^U + \omega_t o_{V,t}^I\right]\sum_{n=2}^{t}\tau_{\varepsilon,n}\varepsilon_n - \mu_t\omega_t\Theta, \tag{2.158}$$

where

$$p_t^V = (1-\omega_t)o_{V,t}^U(\tau_1 + \sum_{n=2}^{t}\tau_{\varepsilon,n}) + \omega_t o_{V,t}^I\sum_{n=1}^{t}\tau_{\varepsilon,n}, \tag{2.159}$$

$$p_t^X = -(1-\omega_t)o_{V,t}^U\tau_1 h_1 - \mu_t\omega_t, \tag{2.160}$$

$$p_t^\varepsilon = (1-\omega_t)o_{V,t}^U\tau_1 + \omega_t o_{V,t}^I\tau_{\varepsilon,1}. \tag{2.161}$$

Therefore,

$$\sigma_t^P = \sqrt{\mathrm{Var}(P_t)} = \sqrt{\left\{\frac{(p_t^V)^2}{\tau_V} + \frac{(p_t^X)^2}{\tau_X} + \frac{(p_t^\varepsilon)^2}{\tau_{\varepsilon,1}} + \left[(1-\omega_t)o_{V,t}^U + \omega_t o_{V,t}^I\right]^2\sum_{n=2}^{t}\tau_{\varepsilon,n}\right\}}. \tag{2.162}$$

Obviously, $P_t \sim \mathcal{N}\left(\mu_t^P, (\sigma_t^P)^2\right)$, where

$$\mu_t^P = \bar{V} - \mu_t\omega_t\Theta, \tag{2.163}$$

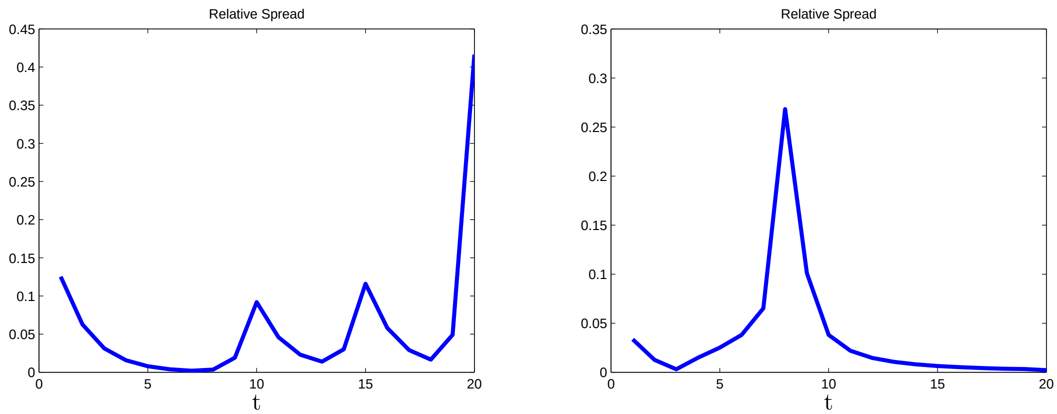
and σ_t^P is given in Appendix 2.8.5. These lead to the average absolute value of price level

$$\mathrm{E}[|P_t|] = \sqrt{\frac{2}{\pi}}\sigma_t^P e^{-\frac{1}{2}\left(\frac{\mu_t^P}{\sigma_t^P}\right)^2} + \mu_t^P\left(1 - 2\boldsymbol{\Phi}\left(-\frac{\mu_t^P}{\sigma_t^P}\right)\right). \tag{2.164}$$

Thus, the relative spread can be computed as

$$\overline{\frac{A_t - B_t}{\mathrm{E}[|P_t|]}} = \frac{c_t^I + c_t^U}{N_M}\frac{\sqrt{\frac{2}{\pi}}\sigma_t^\Delta e^{-\frac{1}{2}(\frac{\mu_t^\Delta}{\sigma_t^\Delta})^2} + \mu_t^\Delta\left(1 - 2\boldsymbol{\Phi}\left(-\frac{\mu_t^\Delta}{\sigma_t^\Delta}\right)\right)}{\sqrt{\frac{2}{\pi}}\sigma_t^P e^{-\frac{1}{2}\left(\frac{\mu_t^P}{\sigma_t^P}\right)^2} + \mu_t^P\left(1 - 2\boldsymbol{\Phi}\left(-\frac{\mu_t^P}{\sigma_t^P}\right)\right)}. \tag{2.165}$$

Therefore, we can plot the relative spread $\overline{\frac{A_t - B_t}{\mathrm{E}[|P_t|]}}$ for each period under different configuration of information flows.



(a) $\tau_{\varepsilon,1} = \tau_{\varepsilon,10} = \tau_{\varepsilon,15} = \tau_{\varepsilon,20} = 1$

(b) $\tau_{\varepsilon,t} = \tau_\varepsilon/T$, $\tau_\varepsilon = 10\tau_V$

Figure 2.5: Parameters are $N_M = 1$, $\lambda^I = \lambda^U = 1$, $\Theta = 1000$, $N_I\theta_0^I = N_U\theta_0^U = \Theta/2$, $\bar{V} = 100$, $\bar{X} = 0$, $\tau_V = 1$, $\tau_X = 1$ and $\tau_{\varepsilon,t} = 0$ without indication.

When there is no information, the relative spread exhibits similar pattern as absolute spread. However, when there is information coming in, the relative spread takes price takes price volatility into account, so it exhibits different pattern from absolute spread or average trading volume. As Figure 2.5(a) shows, the spikes of relative spread are higher near the market closure compared to Figure 2.2(b). In Figure 2.5(b) there is a constant flow of information, i.e., each period there is information with accuracy τ_ε/T coming in. In such information environment, there is no spike in average trading volume or absolute spread. However, there is a spike in relative spread as Figure 2.5(b) shows. When the spike happens is determined by the total information accuracy τ_ε flowing into the market. As τ_ε is higher, the sooner the spike will happen. How large is the spike is determined by the number of trading periods T. As T gets bigger, the smaller the spike will be. Since the information is more spread out, the spike of relative spread becomes lower. However, since this is CARA-normal setup, no matter how large $\bar{V}$ is, P_t can still be close to zero. Therefore, we do not include the relative spread analysis in our main paper.

Bibliography

[1] Franklin Allen, Stephen Morris, and Andrew Postlewaite. Finite bubbles with short sale constraints and asymmetric information. *Journal of Economic Theory*, 61(2):206–229, 1993.

[2] Yang Bai, Eric Chang, and Jiang Wang. Asset prices under short-sale constraints. 2006.

[3] Steven D Baker. The financialization of storable commodities. *Finance Down Under 2014 Building on the Best from the Cellars of Finance*, 2014.

[4] Gurdip Bakshi, Xiaohui Gao Bakshi, and Alberto G Rossi. Understanding the sources of risk underlying the cross-section of commodity returns. *Robert H. Smith School Research Paper No. RHS*, 2589057, 2015.

[5] Snehal Banerjee and Jeremy J Graveline. Trading in derivatives when the underlying is scarce. *Journal of Financial Economics*, 111(3):589–608, 2014.

[6] Suleyman Basak and Anna Pavlova. A model of financialization of commodities. *Journal of Finance, forthcoming*, 2015.

[7] Martin T Bohl and Patrick M Stephan. Does futures speculation destabilize spot prices?

new evidence for commodity markets. *Journal of Agricultural and Applied Economics*, 45(04):595–616, 2013.

[8] Michael J Brennan. A theory of price limits in futures markets. *Journal of Financial Economics*, 16(2):213–233, 1986.

[9] Jonathan Brogaard, Matthew Ringgenberg, and David Sovich. The real impact of speculative trading in financial markets. *Available at SSRN*, 2015.

[10] Celso Brunetti and Bahattin Buyuksahin. Is speculation destabilizing? *Available at SSRN 1393524*, 2009.

[11] Giovanni Cespa and Thierry Foucault. Illiquidity contagion and liquidity crashes. *Review of Financial Studies*, 27(6):1615–1660, 2014.

[12] Eric C Chang. Returns to speculators and the theory of normal backwardation. *The Journal of Finance*, 40(1):193–208, 1985.

[13] Wen Chen and Albert S Kyle. Futures manipulation with "cash settlement" revisited. *Working paper*, 2016.

[14] Haw Cheng, Andrei Kirilenko, and Wei Xiong. Convective risk flows in commodity futures markets. *Review of Finance*, 19:1733—-1781, 2015.

[15] Ing-Haw Cheng and Wei Xiong. The financialization of commodity markets. Technical report, National Bureau of Economic Research, 2013.

[16] Commodity Futures Trading Commission. Aggregate of positions; proposed rule. *Federal Register*, 78(221), 2013.

[17] Commodity Futures Trading Commission. Position limits for derivatives; proposed rule. *Federal Register*, 78(239), 2013.

[18] Peter DeMarzo and Costis Skiadas. Aggregation, determinacy, and informational efficiency for a class of economies with asymmetric information. *journal of economic theory*, 80(1):123–152, 1998.

[19] Olivier Dessaint, Thierry Foucault, Laurent Frésard, and Adrien Matray. Ripple effects of noise on corporate investment. *Available at SSRN*, 2015.

[20] Ivar Ekeland, Delphine Lautier, and Bertrand Villeneuve. Speculation in commodity futures markets: A simple equilibrium model. *Available at SSRN 2323560*, 2015.

[21] Itay Goldstein and Liyan Yang. Commodity financialization: Risk sharing and price discovery in commodity futures markets. *Available at SSRN 2555996*, 2015.

[22] Gary B Gorton, Fumio Hayashi, and K Geert Rouwenhorst. The fundamentals of commodity futures returns. *Review of Finance*, 17(1):35–105, 2013.

[23] Sanford J Grossman and Merton H Miller. Liquidity and market structure. *the Journal of Finance*, 43(3):617–633, 1988.

[24] Sanford J Grossman and Joseph E Stiglitz. On the impossibility of informationally efficient markets. *The American economic review*, 70(3):393–408, 1980.

[25] James D Hamilton and Jing Cynthia Wu. Risk premia in crude oil futures prices. *Journal of International Money and Finance*, 42:9–37, 2014.

[26] Jeffrey H Harris and Bahattin Buyuksahin. The role of speculators in the crude oil futures market. *Available at SSRN 1435042*, 2009.

[27] Oliver D Hart and David M Kreps. Price destabilizing speculation. *The Journal of Political Economy*, pages 927–952, 1986.

[28] Hua He and Jiang Wang. Differential informational and dynamic behavior of stock trading volume. *Review of Financial Studies*, 8(4):919–972, 1995.

[29] Brian J Henderson, Neil D Pearson, and Li Wang. The price impact of large hedging trades. *Available at SSRN 1572000*, 2012.

[30] Brian J Henderson, Neil D Pearson, and Li Wang. New evidence on the financialization of commodity markets. *Review of Financial Studies*, page hhu091, 2014.

[31] John Richard Hicks et al. Value and capital: An inquiry into some fundamental principles of economic theory. *OUP Catalogue*, 1975.

[32] Jack Hirshleifer. Speculation and equilibrium: Information, risk, and markets. *The quarterly journal of Economics*, pages 519–542, 1975.

[33] Conghui Hu and Wei Xiong. Are commodity futures prices barometers of the global economy? Technical report, National Bureau of Economic Research, 2013.

[34] Robert Jarrow. Heterogeneous expectations, restrictions on short sales, and equilibrium asset prices. *The Journal of Finance*, 35(5):1105–1113, 1980.

[35] Norman L Johnson, Samuel Kotz, and N Balakrishnan. *Continuous multivariate distri-*

butions, volume 1, models and applications, volume 59. New York: John Wiley & Sons, 2002.

[36] John Maynard Keynes. A treatise on money, 1930.

[37] April Knill, Kristina Minnick, and Ali Nejadmalayeri. Selective hedging, information asymmetry, and futures prices. *The Journal of Business*, 79(3):1475–1501, 2006.

[38] Albert K. Kyle and Anna A. Obizhaeva. Market microstructure invariance: empirical hypotheses. *Econometrica*, 84(4):1345–1404, 2016.

[39] Albert S Kyle. Continuous auctions and insider trading. *Econometrica: Journal of the Econometric Society*, pages 1315–1335, 1985.

[40] Albert S Kyle. Trading halts and price limits. *Review of Futures Markets*, 7(3):426–434, 1988.

[41] Albert S Kyle. Informed speculation with imperfect competition. *The Review of Economic Studies*, 56(3):317–355, 1989.

[42] Albert S Kyle. Cash settlement, price manipulation, and the modigliani-miller theorem. *Working paper*, 2008.

[43] Emmanuel Leclercq and Rémy Praz. Equilibrium commodity trading. *Available at SSRN 2464400*, 2014.

[44] Jeongmin Lee and Albert S Kyle. Information aggregation with symmetry. *Working paper*, 2016.

[45] Robert S Liptser, Al'bert Nikolaevich Shiriaev, and Albert N Shiryaev. *Statistics of Random Processes II: II. Applications*, volume 2. Springer Science & Business Media, 2001.

[46] Hong Liu and Yajun Wang. Market making with asymmetric information and inventory risk. *Journal of Economic Theory*, 163:73–109, 2016.

[47] Mark Loewenstein and Gregory A Willard. The limits of investor behavior. *The Journal of Finance*, 61(1):231–258, 2006.

[48] Andreas Neuhierl and Andrew J Thompson. Liquidity timing in commodity markets and the impact of financialization. *Available at SSRN 2682698*, 2015.

[49] David MG Newbery and Joseph E Stiglitz. The theory of commodity price stabilisation rules: welfare impacts and supply responses. *The Economic Journal*, 89(356):799–817, 1979.

[50] Richard Roll. A simple implicit measure of the effective bid-ask spread in an efficient market. *The Journal of Finance*, 39(4):1127–1139, 1984.

[51] Michael Sockin and Wei Xiong. Informational frictions and commodity markets. *The Journal of Finance*, 70(5):2063–2098, 2015.

[52] Jeremy C Stein. Informational externalities and welfare-reducing speculation. *Journal of political economy*, 95(6):1123–1145, 1987.

[53] Ke Tang and Wei Xiong. Index investment and the financialization of commodities. *Financial Analysts Journal*, 68(5):54–74, 2012.

[54] Ke Tang and Haoxiang Zhu. Commodities as collateral. *Review of Financial Studies*, page hhw029, 2016.

[55] Dimitri Vayanos and Jiang Wang. Liquidity and asset returns under asymmetric information and imperfect competition. *Review of Financial Studies*, 25(5):1339–1365, 2012.

[56] Yajun Wang. Why can margin requirements increase volatility and benefit margin constrained investors? *Review of Finance*, page rfv041, 2016.

[57] Holbrook Working. The theory of price of storage. *The American Economic Review*, pages 1254–1262, 1949.

[58] Holbrook Working. Futures trading and hedging. In *The Economics of Futures Trading*, pages 68–82. Springer, 1953.