Universitat Autònoma de Barcelona

Essays in Rational Inattention and Market Microstructure

Alfredo Contreras

Essays in Rational Inattention and Market Microstructure

Alfredo Contreras[1]

Advisor: Jordi Caballé

July, 2021

[1]A Thesis submitted to Departament d'Economia i d'Història Econòmica of The Universitat Autònoma de Barcelona in fulfillment of the requirements for the degree of Doctor of Philosophy by the International Doctorate in Economic Analysis (IDEA) Program.

To my parents, Alfredo and Elizabeth

To my wife, Lilian

Acknowledgments

First, I would like to thank my advisor, Jordi Caballé, for his invaluable guidance, support, and dedication. I am also grateful to Johannes Gierlinger, Inés Macho, Pau Olivella, Xavier Martínez, Chaoran Sun, Gaurav Mehta, Jesús Otero, Fernando Jaramillo, and Gabriel Penagos for their helpful comments.

I am deeply thankful to my family for their emotional support during these past years. Most importantly, I thank my wife, Lilian, for her patience and support during the most challenging times.

Additionally, I thank Lourdes Escolar and Jose A. Sánchez-Contador for making my stay in Spain remarkably pleasant. I am most indebted to Àngels López for her timely help with certificates and other invaluable administrative tasks.

I am also grateful to my professors in Colombia for encouraging me to pursue a life in academia. In particular, I am thankful to Alvaro Montenegro, who introduced me, among many other things, the fundamentals of Information Theory.

Finally, I acknowledge COLCIENCIAS for financial support throughout my doctoral studies via the "Segunda convocatoria para estudios de doctorado en el exterior" program and the Banco de la República via the "Beca para estudios de doctorado en el exterior."

Contents

Introduction

In my dissertation, I explain different asset price anomalies in the financial market by characterizing the information sets of those who interact within it. In particular, I study the effects of facing a restriction on the volume of information that an actor can process. For this purpose, I construct the analysis from a financial market microstructure perspective, where a group of investors has access to inside information about the traded assets. I find that these anomalies can originate when a rational agent faces a constraint on the information volume they can process. The dissertation contains results when investors and price-setting agents face this restriction.

The dissertation is composed of three chapters. In the first chapter, I try to explain how a shock to the fundamental value of a specific asset spreads to various asset prices through an informational channel. To show how the asset price covariance arises, I use a model for insider trading where a partially-informed monopolistic investor selects a portfolio. Moreover, the analysis considers that the asset payoffs are statistically independent to eliminate any other potential source of asset price covariance. Thus, the only possible source of correlation appears when the investor chooses the information structure that maximizes his expected profits given a limited information-processing capacity. As a result, the imperfectly informed investor observes correlated information and exhibits hedging behavior, which results in a negative covariance of asset prices. Finally, I find that risk aversion scales down the market orders, and this translates into a reduction in the magnitude of the covariance.

In the second chapter, I address the correlation of the Bid-Ask spread across financial markets. Specifically, I study how investors' information structures can impact different asset prices simultaneously. To this end, I use an analytical framework in which two kinds of investors, insiders and noise traders, submit bid/ask orders to a market maker. I identify that a first source of correlation in bid and ask prices across markets occurs in investors' information structures when they have limited access to information. Since information is costly to process, insiders may choose to observe a more precise correlated information structure over other alternatives. This structure induces correlation in the investor's market orders, and makes the order flow correlated across markets. Thus, market makers' information sets are structurally correlated as well, and the bid/ask prices inherit the correlation given the connection between their information structures. Alternatively, I find that when market makers can trade in all markets, they compensate for the adverse selection problem that insider trading poses and adjust the asset prices to cross-subsidize across markets for the potential losses. The market makers' conduct induces a correlated pricing behavior, even if the investors do not supply correlated information to the market.

Finally, in the third chapter of this dissertation, I address the stochastic behavior of asset prices set by an imperfectly informed specialist who uses learning technology to refine her knowledge of the order flow. The specialist's endogenous choice of an information structure is analyzed given a learning technology and thus fully characterizing her responses to large orders in the market. Specifically, large orders can either be of structural origin, i.e., a disturbance in the asset's payoff, or an exogenous one associated with noise trading. A specialist with a large learning capacity optimally chooses a pricing function where structural shocks display high persistence, whereas exogenous shocks disappear rapidly. This market structure provides a natural setup to address market resilience, in the sense of the recovery speed of prices, and its relation with the specialist's information structure. Thus, market resilience is illustrated through various impulse-response functions for some known stationary stochastic processes, where the effect of the two possible shocks on the order flow is decomposed.

Chapter 1

Attention allocation and price comovement under imperfect competition

1.1 Introduction

Asset price behavior is one of the most extensively analyzed topics in financial economics, from both the theoretical and empirical perspectives. Nevertheless, the explanation for the covariance between some asset prices has always been a grey area. In this paper, I provide a potential explanation for the appearance of negative asset price covariance through a purely informational channel. Informed investors (insiders) are the only agents who supply useful information to the market through their orders. The market prices are usually noisy spreads of the assets' fundamental values. If these values are statistically independent, the appearance of asset price covariance is considered a market anomaly (from any rational expectations perspective). However, when insiders have limited information-processing capacity, they may face a trade-off between the reduction of aggregate external noise and the allocation of all their resources to learn about each asset independently. This learning constraint may lead investors to choose correlated information in order to select their portfolios. Hence, the selected portfolio spreads correlated information across markets, which results in correlated prices.

Mondria (2010) is a related analysis, which introduces an informational channel to explain the presence of market price covariance for assets whose fundamental values are statistically independent. In his model, Mondria considers a perfectly competitive financial market. To center the analysis on the informational channel, I consider a financial market where market orders are traded, and informed investors know the effect their order will have on the asset price. A market order is a choice that the insider submits to the market, which depends on the investor's information rather than the price. So, the framework proposed to analyze this phenomenon is an insider trading model with an imperfectly competitive financial market for two securities, where the insider must deal with a limited capacity to process information. In the model, the information is available to the insider in the form of signals. The insider decides how to allocate his information-processing capacity to a vector of informative signals that deliver noisy information about the asset payoffs. Intuitively, the transmission channel means that whenever the insider faces an information-processing capacity constraint, a volatility shock to either market will affect both prices since it is unfeasible to trace the origin back to the fundamental values from the information structure. Since such a shock impacts both market orders, the information that the market maker observes in the order flow to set the asset prices is correlated. This covariance between asset prices can arise in equilibrium due to a low information-processing capacity, which disappears as this capacity increases.

A common framework used to explain market anomalies is to attribute them to behavioral agents rather than rational ones. Instead, this paper considers a purely rational agent who gets to choose his information structure in the form of a signal and then selects a portfolio. This agent determines the information structure based on an exogenous information-processing capacity. The choice of an information structure is what the economics literature calls 'rational inattention'. A significant advantage of models with rational inattention problems is their convergence to a rational expectations result when the constraint is slacked. Thus, rational inattention offers an intermediate situation between rational and (disciplined-) behavioral agents.

The paper is structured as follows: Section 2 introduces the related literature for both insider trading and rational inattention models and shows how this paper relates to each.

Section 3 describes the market and informational structure. Then, Section 4 shows how covariance behaves if there is a risk-neutral insider, and Section 5 shows how it is affected by the introduction of different levels of risk aversion. Finally, Section 6 provides some concluding remarks.

1.2 Related Literature

The analytical framework considered here for insider trading was first proposed by Kyle (1985). This initial approach considered a model for insider trading with one security and provided analysis for both the competitive and non-competitive cases. Three types of agent interact in this family of models - informed investors, noise traders, and a market maker. Informed agents have access to privileged information, in the form of a signal, about the fundamentals of the securities traded in the market. Uninformed (or noise) traders place random liquidity demands in the market. Finally, there is a market maker who sets the price after she observes the collection of anonymous orders placed in the market.[1] For his analysis, Kyle considered risk-averse insiders for the perfectly competitive setup, and risk-neutral insiders under imperfect competition.[2] Admati (1985) developed a multi-security extension of the same model for the perfectly competitive case where the signal either gives information about each of the asset payoffs separately or any number of linear combinations of such payoffs. Caballé and Krishnan (1994) built a multi-security extension where market orders have a direct impact on the asset price, where the market power is concentrated among the insiders.[3] The literature for imperfectly competitive markets contains further extensions of Kyle (1985). Subrahmanyam (1991) solved the model for a single security with risk-averse insiders, and more recently Vitale (2012) developed the multi security case for risk-averse agents.

Rational Inattention is a concept originated in Sims (1998) as a possible source for

[1]It should be noted that in perfectly competitive markets, the introduction of a market maker is equivalent to the market clearing condition. In a non-competitive scenario, this is no longer true.

[2]Unlike the perfectly competitive case, under imperfect competition when agents are risk-neutral, their ability to influence prices causes preferences to move in the mean-variance space.

[3]Informed traders (insiders) can place their orders conditional on their inside information or both the inside information and the price. When the order only depends on the private signal, it is known as a "market order," while the second one as "limit-order."

"stickiness" in markets. Sims (2003) formally defines a method to bound the information set of an agent according to the volume of information he can process. To do so, he quantifies the volume of information the agent is capable to process through the entropy function, used in information theory. As a feature, rational inattentive agents face a capacity constraint to process information, which allows them to resolve a certain amount of uncertainty. As the information-processing capacity increases, the agents resolve more uncertainty. A detailed explanation is provided in the following section.

The first approaches that related the Rational Inattention concept to financial economics are Peng (2005) and Peng and Xiong (2006), who address it as a learning technology, entropy learning, where agents allocate their attention across multiple sources of uncertainty about fundamental values. Later on, van Nieuwerburgh and Veldkamp (2009) and van Nieuwerburgh and Veldkamp (2010) use Admati (1985) to analyze the attention allocation problem considering different learning technologies and test their results against different preferences towards risk. As a result of the introduction of an information-processing constraint van Nieuwerburgh and Veldkamp (2010) provides an explanation for portfolio under-diversification. All these models consider the sources of uncertainty to be ex-ante independent, inducing portfolio selection and prices to inherit independence. Afterwards, Mondria (2010) proposes a more flexible linear information structure proposed in Admati (1985). As a result, he finds that, even if payoffs are ex-ante independent, once the investor chooses a structure given his capacity constraint, the assets display ex-post price comovement.

1.3 Market structure

There are three types of agent who interact in this financial market - insiders, noise traders and market makers. The investors, insiders and noise traders, place a vector of orders in the market, whereas the market maker can only observe the aggregate orders, also called order flow. After she observes the order flow, the market maker sets the price vector. Investors can be either informed or noise traders. There is a monopolistic informed investor who has access to a noisy signal on the vector of asset payoffs, while the uninformed investor places random orders. The model follows the imperfectly competitive market structure in Caballé and Krishnan (1994), where the informed investor submits a market order after he observes some inside information and takes into account the effect that his order has on the asset price. The insider's information structure is as proposed in Admati (1985) and Mondria (2010), where the informed investor observes a series of functions composed by noisy linear combinations of the asset payoffs. Initially, I solve the model for a risk-neutral insider, then I extend the analysis to a risk-averse one.

There are two stages to this model. First, the insider chooses a linear information structure given a capacity constraint. This choice is called attention allocation. In the second stage, signals are available for the insiders, investors place their orders, and the market maker sets the price. After he observes the signal, the informed investor places market orders in order for them to maximize expected profits, while uninformed investors place random liquidity demands. Together, these are called the order flow. The market maker observes the vector of aggregate demands and sets the prices to follow a zero expected profit rule.[4] Finally, payoffs are received.

The characteristics of the insider model an implicit preference for early resolution of uncertainty that is independent of whether or not the insider is risk-averse. The two-stage structure endows the insider with two different levels of uncertainty. In the first stage, the insider chooses the structure that enables him to reduce uncertainty in the second stage. In the trading stage, the insider updates his beliefs after he partially resolves the uncertainty as he observes the signal. Thus, for a Morgenstern-von Neumann utility function $u(w)$ and an information structure $\mathcal{I}$, the expected utility function for the insider is $\mathbb{E}_0 \left[\mathbb{E} \left[u(w) | \mathcal{I} \right] \right]$. For further discussion of this structure, see van Nieuwerburgh and Veldkamp (2010).

1.3.1 Insider information

In the model, we quantify the information flow as proposed in Information Theory. Shannon (1948) defines a measure of unpredictability (uncertainty, in our context), called

[4]The market maker's price rule decision is a zero profit condition due to risk neutrality and Bertrand competition between market makers.

entropy.[5] For example, consider a source that generates a message that gets distorted when it travels through any channel. There is also a receiver that tries to recover as much information as possible from the distorted message. In economics, the message can be any informative variable that agents use as a signal to make decisions. The information structure would be the corresponding channel, and the capacity of this channel determines the precision of the signal. In general, the source is a random variable, and a message is another random variable. In statistics, the information that a random variable contains about another random variable measures the amount of uncertainty that the observed variable can resolve about the unobserved one.

The introduction of rational inattention, Sims (2003) and (2006) solves potential limitation of rational expectations theory. An implicit assumption of rational expectations models is that the channel does not distort information at all, i.e., information can be processed without any noise at no cost. As an alternative approach, Sims defined rational inattention, which used the concept of information-processing constraints proposed in Shannon (1948) to define the "attention" that an agent would allocate to the reduction of noisy content in the information she observes.

There is a close relationship between the entropy function H and information. Let P be the probability function of a random variable, the entropy, $H(\tilde{a}) = -\mathbb{E}\left(\ln P(\tilde{a})\right)$ gives the amount of information required to solve the uncertainty in variable $\tilde{a}$. As a result, it is frequently stated that information is inversely proportional to the probability. Suppose that the distribution of variable $\tilde{a}$ was degenerate. Then, $H(\tilde{a}) = 0$, since the behavior of the variable $\tilde{a}$ is known without the need for any additional information. In certain sense, it is a measure of how surprising or unexpected is the realization of the random variable $\tilde{a}$ is. Similarly, conditional entropy $H\left(\tilde{a}|\tilde{b}\right) = -\mathbb{E}\left(\ln P(\tilde{a}|\tilde{b})\right)$ measures the unresolved uncertainty in the random variable $\tilde{a}$ after variable $\tilde{b}$ is observed. Unsurprisingly, if $\tilde{a}$ and $\tilde{b}$ are independent, both variables are useless for the resolution of any level of uncertainty about the other. That is, $H\left(\tilde{a}|\tilde{b}\right) = H\left(\tilde{a}\right)$ and $H\left(\tilde{b}|\tilde{a}\right) = H\left(\tilde{b}\right)$. Then, the difference between conditional and unconditional entropy is called mutual information and is used

[5]This type of entropy is called Shannon entropy, which was generalized later by Rényi (1961). Although the intuition is somehow similar, Shannon entropy should not be confused with the concept of thermodynamic entropy in statistical mechanics. In general, probability distributions have an associated entropy function denoted by $H(X)$.

as a measure of the informational content of signals.

In the particular setup of this paper, the random vector of payoffs of all assets $\tilde{F}$ plays the role of the source, and the message corresponds to the signal $\tilde{Y}$. A system's mutual indormation is defined as the difference between the conditional and unconditional entropy function, that is

$$I\left(\tilde{F};\tilde{Y}\right)=H\left(\tilde{F}\right)-H\left(\tilde{F}|\tilde{Y}\right),$$

where $H(\cdot)$ is the entropy function.[6]

This mutual information gives the amount of information about the vector of asset payoffs $\tilde{F}$ contained in $\tilde{Y}$. Equivalently, the amount of uncertainty resolved about $\tilde{F}$ after $\tilde{Y}$ becomes available. The capacity constraint κ is gives the upper limit for the uncertainty reduction that an agent can achieve, that is

$$H\left(\tilde{F}\right)-H\left(\tilde{F}|\tilde{Y}\right)\leq\kappa.$$

Let the payoff vector ($\tilde{F}$) and the signals ($\tilde{Y}$) follow a multivariate normal distribution (hereinafter MN) , the information-processing constraint can be re-written as

$$\ln|\Sigma_F|-\ln\left|\Sigma_{F|Y}\right|\leq 2\kappa, \tag{1.1}$$

which can intuitively be interpreted as a noise to signal ratio. A more detailed discussion is provided in Appendix C. Finally, recall that due to the normality of the variables of interest, zero mutual information implies both statistical independence and zero covariance, which may not be true for other parametric distributions.

1.4 A model for risk-neutral insiders

Consider initially a risk-neutral investor who faces a portfolio selection problem over two risky assets. This agent is called an insider. All other investors in the market place random orders for the risky assets since they do not have access to information. These are called noise traders. There is a third agent, a market maker, who observes the orders placed by all investors and sets the prices for the assets. Bertrand price competition exists across

[6]The entropy function for a multivariate normal random vector $X\sim MN\left(\mu_X,\Sigma_X\right)$, where $\dim_{\mathbb{R}}(X)=n$, is given by $H(X)=\frac{1}{2}\ln\left((2\pi e)^n|\Sigma_X|\right)$.

market makers, which results in a zero expected profit pricing rule. The market maker is aware of the existence of the informed investor, but the trading protocol generates coverage for the informed traders. The existence of noise traders provides camouflage for the insider since the market maker cannot distinguish between the orders. As a result, the market maker does not have enough information to practice price discrimination. This creates an informational advantage for the insider, which inevitably results in an informational disadvantage for the noise traders since the pricing rule makes a zero(expected)-sum across investors.

Let $\tilde{F}$ be the two dimensional vector of asset payoffs, and let $r = 1$ be the return of a riskless asset. Both payoffs are measured in units of a consumption good. The uninformed investors generate a vector of random liquidity demands $\tilde{z}$, which follows a multivariate normal distribution with mean $\bar{z}$ and covariance matrix Σ_z, $\tilde{z} \sim MN(\bar{z}, \Sigma_z)$. The insider faces two types of choices in separate stages. In the first stage, an information structure is chosen in the form of a signal, which provides the informational advantage for the portfolio selection. Such a signal has the form

$$\tilde{Y} = C\tilde{F} + \tilde{\varepsilon},$$

with $\tilde{\varepsilon} \sim MN(0, \Sigma_\varepsilon)$ and $\tilde{Y} \sim MN\left(C\bar{F}, \Sigma_Y\right)$, where $\Sigma_Y = C\Sigma_F C^\top + \Sigma_\varepsilon$. The informational content of $\tilde{Y}$ is bounded by an exogenous information-processing capacity κ. In the second stage, the vector of private signals is observed according to the information structure (attention) that is allocated and the vector of market orders $\tilde{x}$ is placed.

The risk-neutral market maker observes the order flow vector

$$\tilde{\omega} \equiv \tilde{x} + \tilde{z},$$

and sets a price such that zero expected profits are made,

$$p\left(\tilde{\omega}\right) = \mathbb{E}\left(\tilde{F}|\tilde{\omega}\right).$$

There is a market for market makers where there is Bertrand competition in asset prices, which prevents the setting of any other price. Let $\tilde{\xi} = \mathbb{E}\left(\tilde{F}|\tilde{Y}\right) - \bar{F}$ be the insider's informational advantage. Moreover, the informational advantage for this insider is

$$\tilde{\xi} = \Sigma_F C^\top \left[C\Sigma_F C^\top + \Sigma_\varepsilon\right]^{-1} \left[\tilde{Y} - C\bar{F}\right], \tag{1.2}$$

which is distributed $MN\left(0, \Sigma_\xi\right)$, where $\Sigma_\xi = \Sigma_F - \Sigma_F C^\top \left[C\Sigma_F C^\top + \Sigma_\varepsilon\right]^{-1} C\Sigma_F$.

The stages of the model are defined as follows: in the first stage, the investor has to choose how to allocate attention (C and Σ_ε) subject to the information-processing constraint (1.1). Then, the information extracted from the signal is incorporated in the previous stage. Finally, given the posterior beliefs, the insider selects a portfolio. The model is solved by backward induction.

Definition 1. For a given capacity level κ, a **noisy rationally-inattentive** equilibrium is an attention allocation (C, Σ_ε), the signal $\tilde{Y}$, a market order $\chi\left(\tilde{Y}\right)$ and a price vector $\tilde{p}(\tilde{\omega})$, where the monopolistic insider maximizes his expected profits subject to the information-processing constraint and the market maker has zero expected profits.

1.4.1 Portfolio selection

The insider selects the portfolio that maximizes the expected profits and conditions the expectations with regard to the information delivered by the previously constructed signal. The insider chooses the optimal asset holdings x to solve the problem

$$\max_{\{x\}} \mathbb{E}\left[W_0 + x^\top(\tilde{F} - p(\tilde{\omega}))|\tilde{Y}\right],$$

which take the form of market orders, i.e., demands that do not depend on the asset price.

The solution of this second stage follows Caballé and Krishnan (1994) for a monopolistic insider. The pricing reaction function of the market maker is a linear function of the order flow $\tilde{\omega}$, and the market orders are linear functions of the informational advantage $\tilde{\xi}$. Note that market orders are, by definition, a function $\tilde{x} = \chi\left(\tilde{Y}\right)$ of the signal. However, since the informational advantage is a linear transformation of the signal, it is equivalent to define market orders as a function of the informational advantage $x\left(\tilde{\xi}\right) \equiv \chi\left(\tilde{Y}\right)$.

More specifically, the pricing function is

$$p\left(\tilde{\omega}\right) = \bar{F} + Q_0 + Q_1\tilde{\omega}, \tag{1.3}$$

where Q_0 is a two-dimensional column vector, and Q_1 is a 2×2 matrix of coefficients, and $\tilde{\omega} = x\left(\tilde{\xi}\right) + \tilde{z}$. Similarly, the vector of market orders is

$$x\left(\tilde{\xi}\right) = D_0 + D_1\tilde{\xi}, \tag{1.4}$$

where D_0 is a two-dimensional column vector and D_1 is a 2×2 matrix of coefficients. The equilibrium in the subgame found by solving the linear system above that satisfies the first order condition. As a result, two key identities are obtained, $D_0 = \underline{0}$, and $Q_1 = \frac{1}{2}D_1^{-1}$. See Appendix D for further details.

1.4.2 Information structure and attention allocation

The first choice that the insider faces is to optimally construct an information structure to observe before he selects a portfolio. The investor's rationality is used to know beforehand that, in the trading stage, a portfolio is selected that is a linear function of the informational advantage and the market maker sets a pricing rule that is linear to the order flow. That is, the investor knows (1.3) and (1.4), as well as the identities mentioned in the previous section. The insider's objective function corresponds to his unconditional expected profits

$$\begin{aligned}
&\mathbb{E}\left[x\left(\tilde{\xi}\right)^{\top}\left(\tilde{F}-p\left(\tilde{\omega}\right)\right)\right]\\
=&\mathbb{E}\left[\left(D_0+D_1\tilde{\xi}\right)^{\top}\left(\tilde{F}-\bar{F}-Q_0-Q_1\left(D_0+D_1\tilde{\xi}+\tilde{z}\right)\right)\right],
\end{aligned}$$

where $\tilde{\xi} = \Sigma_F C^{\top}\left[C\Sigma_F C^{\top}+\Sigma_{\varepsilon}\right]^{-1}\left[\tilde{Y}-C\bar{F}\right]$ is the informational advantage with zero mean and covariance matrix $\Sigma_{\xi} = \Sigma_F - \Sigma_F C^{\top}\left[C\Sigma_F C^{\top}+\Sigma_{\varepsilon}\right]^{-1}C\Sigma_F$. Since the objective function is real valued, the resulting expression

$$D_0^{\top}Q_0+\mathbb{E}\left[\tilde{\xi}^{\top}D_1^{\top}\left(\tilde{F}-\bar{F}\right)-\tilde{\xi}^{\top}D_1^{\top}Q_1D_1\tilde{\xi}-\tilde{\xi}^{\top}D_1^{\top}\left(Q_0+Q_1\tilde{z}\right)\right],$$

satisfies the matrix operating properties for scalars. As a result, $\mathbb{E}\left[\tilde{\xi}^{\top}D_1^{\top}Q_1D_1\tilde{\xi}\right] = Tr\left[D_1^{\top}Q_1D_1\Sigma_{\xi}\right]$, and $\mathbb{E}\left[\tilde{\xi}^{\top}D_1^{\top}\left(Q_0+Q_1\tilde{z}\right)\right]=0$.

From the second stage the insider knows that $D_0 = \underline{0}$, and $Q_1 = \frac{1}{2}D_1^{-1}$. Recall that the matrix of covariances between the signal and the vector of payoffs is $\Sigma_{FY} =$

$\mathbb{E}\left[\left(\tilde{F}-\bar{F}\right)\left(\tilde{Y}-C\bar{F}\right)^{\top}\right]=\Sigma_F C^{\top}$, then

$$\begin{aligned}
\mathbb{E}\left[\tilde{\xi}^{\top}D_1^{\top}\left(\tilde{F}-\bar{F}\right)\right] &= \mathbb{E}\left[\left(\tilde{Y}-C\bar{F}\right)^{\top}\left[C\Sigma_F C^{\top}+\Sigma_{\varepsilon}\right]^{-1}C\Sigma_F D_1^{\top}\left(\tilde{F}-\bar{F}\right)\right] \\
&= Tr\left[\mathbb{E}\left[\left(C\tilde{F}-C\bar{F}+\tilde{\varepsilon}\right)^{\top}\left[C\Sigma_F C^{\top}+\Sigma_{\varepsilon}\right]^{-1}C\Sigma_F D_1^{\top}\left(\tilde{F}-\bar{F}\right)\right]\right] \\
&= Tr\left[C^{\top}\left[C\Sigma_F C^{\top}+\Sigma_{\varepsilon}\right]^{-1}C\Sigma_F D_1^{\top}\mathbb{E}\left[\left(\tilde{F}-\bar{F}\right)\left(\tilde{F}-\bar{F}\right)^{\top}\right]\right] \\
&= Tr\left[C^{\top}\left[C\Sigma_F C^{\top}+\Sigma_{\varepsilon}\right]^{-1}C\Sigma_F D_1^{\top}\Sigma_F\right].
\end{aligned}$$

Altogether, the investor's problem for attention allocation becomes

$$\begin{aligned}
&\max_{C,\Sigma_{\varepsilon}}\frac{1}{2}Tr\left[C^{\top}\left[C\Sigma_F C^{\top}+\Sigma_{\varepsilon}\right]^{-1}C\Sigma_F D_1^{\top}\Sigma_F\right] \\
&\text{subject to } \ln|\Sigma_F|-\ln\left|\Sigma_F-\Sigma_F C^{\top}\left[C\Sigma_F C^{\top}+\Sigma_{\varepsilon}\right]^{-1}C\Sigma_F\right|\leq 2\kappa. \qquad (1.5)
\end{aligned}$$

However, there is still one issue to be addressed in this problem, the multiplicity of equilibria due to the dimension of the vector of signals.

Characterization of the signal

The specific form for the signal used here, is a weighted sum of the true asset payoffs plus some exogenous noise, $\tilde{Y}=C\tilde{F}+\tilde{\varepsilon}$. The number of signals is given by the number of rows in matrix C. The insider is allowed to acquire as many signals as he wants as long as he does not exceed the capacity κ. Suppose that matrix C had more than two rows, then, any nonzero vector added as a third row could be achieved by the linear combination of the first two rows. That is, the rank of matrix C is bounded by the dimension of the asset payoff vector. In other words, the addition of a signal allocates capacity κ into the acquisition of some redundant information. In general, for any m dimensional vector of asset payoffs, the vector of signals $\tilde{Y}$ is at most m dimensional. This model considers the case of $m=2$. Hence, there can only be either one or two signals that supply non-redundant information through vector $\tilde{Y}$.

First, consider the case of a two-dimensional signal. There are infinite values of the coefficient matrix C that generate equally informative signals. Thus, there are just as many indistinguishable equilibria satisfying the maximization conditions in problem (1.5). It is possible to characterize the conditions within all the indistinguishable equilibria, where the noise across signals is uncorrelated, and there is only one matrix for each information level. Such conditions are summarized in a diagonal covariance matrix for the noise term

$$\Sigma_\varepsilon = \begin{bmatrix} \sigma^2_{\varepsilon_1} & 0 \\ 0 & \sigma^2_{\varepsilon_2} \end{bmatrix},$$

and a weighting matrix

$$C = \begin{bmatrix} 1 & c_1 \\ 1 & c_2 \end{bmatrix},$$

where the elements of matrix C were normalized by the elements of the first column. Such a result follows directly from Mondria (2010). See further details in Appendix B.

Let

$$\Sigma_F = \begin{bmatrix} \sigma^2_{F_1} & 0 \\ 0 & \sigma^2_{F_2} \end{bmatrix},$$

be the covariance matrix of asset payoffs, and

$$D_1 = \begin{bmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{bmatrix},$$

is the matrix that contains the marginal effects of the informational advantages over the market orders. Both matrices are exogenous at this stage.

One signal equilibrium

An attention allocation corresponds to the values of weights c_1, c_2 and individual variances $\sigma^2_{\varepsilon_1}$, $\sigma^2_{\varepsilon_2}$. The optimal allocation selects $\sigma^2_{\varepsilon_i}$ first, and then solves for coefficients c_i, $i = 1, 2$. Although equivalent, it is more natural to allow the agent to choose the precision of each signal, $\sigma^{-2}_{\varepsilon_1}, \sigma^{-2}_{\varepsilon_2}$, rather than their variances $\sigma^2_{\varepsilon_1}$, $\sigma^2_{\varepsilon_2}$. Then, the attention allocation problem for the insider becomes

$$\max_{\sigma_{\varepsilon_1}^{-2},\sigma_{\varepsilon_2}^{-2}} \frac{d_{22}\left(\sigma_{F_2}^2\right)^2\left(c_1^2(\sigma_{\varepsilon_2}^2+\sigma_{F_1}^2)-2c_1c_2\sigma_{F_1}^2+c_2^2(\sigma_{\varepsilon_1}^2+\sigma_{F_1}^2)\right)+\sigma_{F_1}^2\sigma_{F_2}^2(d_{12}+d_{21})(c_1\sigma_{\varepsilon_2}^2+c_2\sigma_{\varepsilon_1}^2)}{2\left(\sigma_{\varepsilon_2}^2\left(c_1^2\sigma_{F_2}^2+\sigma_{F_1}^2\right)+\sigma_{F_1}^2\sigma_{F_2}^2(c_1-c_2)^2+\sigma_{\varepsilon_1}^2\left(c_2^2\sigma_{F_2}^2+\sigma_{\varepsilon_2}^2+\sigma_{F_1}^2\right)\right)}$$
$$+\frac{d_{11}\left(\sigma_{F_1}^2\right)^2\left(\sigma_{F_2}^2(c_1-c_2)^2+\sigma_{\varepsilon_1}^2+\sigma_{\varepsilon_2}^2\right)}{2\left(\sigma_{\varepsilon_2}^2\left(c_1^2\sigma_{F_2}^2+\sigma_{F_1}^2\right)+\sigma_{F_1}^2\sigma_{F_2}^2(c_1-c_2)^2+\sigma_{\varepsilon_1}^2\left(c_2^2\sigma_{F_2}^2+\sigma_{\varepsilon_2}^2+\sigma_{F_1}^2\right)\right)} \tag{1.6}$$

$$\text{subject to } \left(\left(c_1^2\sigma_{F_2}^2+\sigma_{F_1}^2\right)\sigma_{\varepsilon_1}^{-2}+\sigma_{F_1}^2\sigma_{F_2}^2(c_1-c_2)^2\sigma_{\varepsilon_1}^{-2}\sigma_{\varepsilon_2}^{-2}+\left(c_2^2\sigma_{F_2}^2+\sigma_{F_1}^2\right)\sigma_{\varepsilon_2}^{-2}\right)=e^{2\kappa}-1.$$

The problem can be transformed into the unconstrained maximization problem,

$$\max_{\sigma_{\varepsilon_1}^{-2},\sigma_{\varepsilon_2}^{-2}} \frac{1}{2e^{2\kappa}\sigma_{\varepsilon_1}^2\sigma_{\varepsilon_2}^2}\Big[d_{22}\left(\sigma_{F_2}^2\right)^2\left(c_1^2(\sigma_{\varepsilon_2}^2+\sigma_{F_1}^2)-2c_1c_2\sigma_{F_1}^2+c_2^2(\sigma_{\varepsilon_1}^2+\sigma_{F_1}^2)\right)$$
$$+\sigma_{F_1}^2\sigma_{F_2}^2(d_{12}+d_{21})(c_1\sigma_{\varepsilon_2}^2+c_2\sigma_{\varepsilon_1}^2)+d_{11}\left(\sigma_{F_1}^2\right)^2\left(\sigma_{F_2}^2(c_1-c_2)^2+\sigma_{\varepsilon_1}^2+\sigma_{\varepsilon_2}^2\right)\Big].$$

Or equivalently,

$$\max_{\sigma_{\varepsilon_1}^{-2},\sigma_{\varepsilon_2}^{-2}} \frac{1}{2e^{2\kappa}}\Big[d_{22}\left(\sigma_{F_2}^2\right)^2\left(c_1^2\left(\sigma_{\varepsilon_1}^{-2}+\sigma_{F_1}^2\sigma_{\varepsilon_1}^{-2}\sigma_{\varepsilon_2}^{-2}\right)-2c_1c_2\sigma_{F_1}^2\sigma_{\varepsilon_1}^{-2}\sigma_{\varepsilon_2}^{-2}+c_2^2\left(\sigma_{\varepsilon_2}^{-2}+\sigma_{F_1}^2\sigma_{\varepsilon_1}^{-2}\sigma_{\varepsilon_2}^{-2}\right)\right)$$
$$+\sigma_{F_1}^2\sigma_{F_2}^2(d_{12}+d_{21})\left(c_1\sigma_{\varepsilon_1}^{-2}+c_2\sigma_{\varepsilon_2}^{-2}\right)+d_{11}\left(\sigma_{F_1}^2\right)^2\left(\sigma_{F_2}^2(c_1-c_2)^2\sigma_{\varepsilon_1}^{-2}\sigma_{\varepsilon_2}^{-2}+\sigma_{\varepsilon_2}^{-2}+\sigma_{\varepsilon_1}^{-2}\right)\Big].$$

Such an expression can be rearranged to obtain

$$\max_{\sigma_{\varepsilon_1}^{-2},\sigma_{\varepsilon_2}^{-2}}\left\{\gamma_1\sigma_{\varepsilon_1}^{-2}+\gamma_2\sigma_{\varepsilon_2}^{-2}+\gamma_{12}\sigma_{\varepsilon_1}^{-2}\sigma_{\varepsilon_2}^{-2}\right\},$$

where $\gamma_1=c_1^2d_{22}\left(\sigma_{F_2}^2\right)^2+c_1\sigma_{F_1}^2\sigma_{F_2}^2(d_{12}+d_{21})+d_{11}\left(\sigma_{F_1}^2\right)^2$, $\gamma_2=c_2^2d_{22}\left(\sigma_{F_2}^2\right)^2+c_2\sigma_{F_1}^2\sigma_{F_2}^2(d_{12}+d_{21})+d_{11}\left(\sigma_{F_1}^2\right)^2$, and $\gamma_{12}=\left[d_{11}\left(\sigma_{F_1}^2\right)^2\sigma_{F_2}^2+d_{22}\left(\sigma_{F_2}^2\right)^2\sigma_{F_1}^2\right](c_1-c_2)^2$. The capacity constraint can be used to obtain an iso-information function that relates all the precision levels of $\sigma_{\varepsilon_2}^{-2}$ with $\sigma_{\varepsilon_1}^{-2}$, for constant values of the mutual information and all other parameters. The iso-information function takes the form

$$\sigma_{\varepsilon_2}^{-2}=\frac{e^{2\kappa}-1-\left(c_1^2\sigma_{F_2}^2+\sigma_{F_1}^2\right)\sigma_{\varepsilon_1}^{-2}}{\left(c_2^2\sigma_{F_2}^2+\sigma_{F_1}^2\right)+\sigma_{F_1}^2\sigma_{F_2}^2(c_1-c_2)^2\sigma_{\varepsilon_1}^{-2}}, \tag{1.7}$$

for

$$\sigma_{\varepsilon_1}^{-2}\in\left[0,\frac{e^{2\kappa}-1}{\left(c_1^2\sigma_{F_2}^2+\sigma_{F_1}^2\right)}\right].$$

The attention allocation problem is then reduced to the identification of the profit-maximizing allocation over the iso-information curve. This takes one dimension away from the chosen attention allocation. As a result, the insider's attention allocation problem solves

$$\max_{\sigma_{\varepsilon_1}^{-2}} \gamma_1 \sigma_{\varepsilon_1}^{-2} + \left(\gamma_2 + \gamma_{12}\sigma_{\varepsilon_1}^{-2}\right) \frac{e^{2\kappa} - 1 - \left(c_1^2\sigma_{F_2}^2 + \sigma_{F_1}^2\right)\sigma_{\varepsilon_1}^{-2}}{\left(c_2^2\sigma_{F_2}^2 + \sigma_{F_1}^2\right) + \sigma_{F_1}^2\sigma_{F_2}^2(c_1 - c_2)^2\sigma_{\varepsilon_1}^{-2}},$$

which is univariate.

Lemma 2. *For all $\kappa > 0$, the agent allocates all attention to one signal, if $c_2(d_{12}+d_{21}) \geq c_2^2 d_{11} + d_{22}$.*

Proof. The second order condition for the problem is

$$\frac{\partial^2}{\left(\partial\sigma_{\varepsilon_1}^{-2}\right)^2} = \frac{2\left(\sigma_{F_1}^2\sigma_{F_2}^2\left(c_1 - c_2\right)^2 e^{2\kappa}\right)\left(\left(c_2^2\sigma_{F_2}^2 + \sigma_{F_1}^2\right)\gamma_{12} - \sigma_{F_1}^2\sigma_{F_2}^2\left(c_1 - c_2\right)^2\gamma_2\right)}{\left(\left(c_2^2\sigma_{F_2}^2 + \sigma_{F_1}^2\right) + \sigma_{F_1}^2\sigma_{F_2}^2(c_1 - c_2)^2\sigma_{\varepsilon_1}^{-2}\right)^3},$$

which is non-negative if and only if

$$\left(c_2^2\sigma_{F_2}^2 + \sigma_{F_1}^2\right)\gamma_{12} \geq \sigma_{F_1}^2\sigma_{F_2}^2\left(c_1 - c_2\right)^2\gamma_2,$$

or equivalently,

$$c_2^2 d_{11} + d_{22} \geq c_2(d_{12} + d_{21}).$$

Hence, under this condition the function is convex and there is a corner solution. Given the identity in (1.7), a corner solution means that only one signal can receive all attention, $\sigma_{\varepsilon_1}^{-2} = \frac{e^{2\kappa}-1}{\left(c_1^2\sigma_{F_2}^2+\sigma_{F_1}^2\right)}$, $\sigma_{\varepsilon_2}^{-2} = 0$ or $\sigma_{\varepsilon_1}^{-2} = 0$ and $\sigma_{\varepsilon_2}^{-2} = \frac{e^{2\kappa}-1}{\left(c_2^2\sigma_{F_2}^2+\sigma_{F_1}\right)}$. Such a condition does not depend on κ, therefore it holds for all $\kappa > 0$. □

The previous result implies that through the allocation of all attention to a single signal, say signal $\tilde{Y}_1$. Then, signal $\tilde{Y}_2$ is not used, irrespectively of the off-diagonal value c_2 that is set. As a result, the signal where the insider allocates all the attention is such that

$$\sigma_{\varepsilon}^{-2} = \frac{e^{2\kappa} - 1}{\left(c^2\sigma_{F_2}^2 + \sigma_{F_1}^2\right)}. \tag{1.8}$$

Alternatively, an equivalent solution eliminates the uninformative dimension from the signal vector. Now the investor faces the following problem to determine the optimal weight c,

$$\max_c \left\{ \frac{d_{11}\left(\sigma_{F_1}^2\right)^2 + c\sigma_{F_1}^2\sigma_{F_2}^2\left(d_{21} + d_{12}\right) + c^2 d_{22}\left(\sigma_{F_2}^2\right)^2}{\sigma_{F_1}^2 + c^2\sigma_{F_2}^2 + \sigma_{\varepsilon}^{-2}} \right\},$$

where σ_{ε}^2 is given by 1.8 and the values of d_{ij} are such that they solve the first order condition for the insider in the first stage.

To compute the specific values, I find d_{ij} and solve $D_1 = \frac{1}{2}Q_1^{-1}$. [7] First, let

$$Q_1 = \left(\frac{1}{\sigma_{F_1}^2 + c^2\sigma_{F_2}^2 + \sigma_\varepsilon^2} \begin{bmatrix} \sigma_{F_1}^2 \left(c^2\sigma_{F_2}^2 + \sigma_\varepsilon^2\right) & -c\sigma_{F_1}^2\sigma_{F_2}^2 \\ -c\sigma_{F_1}^2\sigma_{F_2}^2 & \sigma_{F_2}^2 \left(\sigma_{F_1}^2 + \sigma_\varepsilon^2\right) \end{bmatrix} \right)^{\frac{1}{2}},$$

or, equivalently

$$Q_1 = \frac{\left(\sigma_{F_1}^2 \left(c^2\sigma_{F_2}^2 + \sigma_\varepsilon^2\right) + \sigma_{F_2}^2 \left(\sigma_{F_1}^2 + \sigma_\varepsilon^2\right) - 2s\right)^{-\frac{1}{2}}}{\left(\sigma_{F_1}^2 + c^2\sigma_{F_2}^2 + \sigma_\varepsilon^2\right)^{\frac{1}{2}}} \begin{bmatrix} \sigma_{F_1}^2 \left(c^2\sigma_{F_2}^2 + \sigma_\varepsilon^2\right) + s & -c\sigma_{F_1}^2\sigma_{F_2}^2 \\ -c\sigma_{F_1}^2\sigma_{F_2}^2 & \sigma_{F_2}^2 \left(\sigma_{F_1}^2 + \sigma_\varepsilon^2\right) + s \end{bmatrix},$$

where $s = \left(\sigma_{F_1}^2 \left(c^2\sigma_{F_2}^2 + \sigma_\varepsilon^2\right) \sigma_{F_2}^2 \left(\sigma_{F_1}^2 + \sigma_\varepsilon^2\right) - \left(c\sigma_{F_1}^2\sigma_{F_2}^2\right)^2\right)^{\frac{1}{2}}$. Then,

$$D_1 = \frac{1}{2} \frac{\left(\sigma_{F_1}^2 + c^2\sigma_{F_2}^2 + \sigma_\varepsilon^2\right)^{\frac{1}{2}} \left(\sigma_{F_1}^2 \left(c^2\sigma_{F_2}^2 + \sigma_\varepsilon^2\right) + \sigma_{F_2}^2 \left(\sigma_{F_1}^2 + \sigma_\varepsilon^2\right) - 2s\right)^{\frac{1}{2}}}{\left(\sigma_{F_1}^2 \left(c^2\sigma_{F_2}^2 + \sigma_\varepsilon^2\right) + s\right) \left(\sigma_{F_2}^2 \left(\sigma_{F_1}^2 + \sigma_\varepsilon^2\right) + s\right) - \left(c\sigma_{F_1}^2\sigma_{F_2}^2\right)^2} \begin{bmatrix} \sigma_{F_2}^2 \left(\sigma_{F_1}^2 + \sigma_\varepsilon^2\right) + s & c\sigma_{F_1}^2\sigma_{F_2}^2 \\ c\sigma_{F_1}^2\sigma_{F_2}^2 & \sigma_{F_1}^2 \left(c^2\sigma_{F_2}^2 + \sigma_\varepsilon^2\right) + s \end{bmatrix}.$$

The coefficients in matrix D_1 are determined by the subgame in the second stage. The insider is not aware of the indirect impact that the attention allocation has on the market equilibrium but rather the direct informativeness impact on expected profits. Note that these coefficients are not to be included in the insider's problem since the values are set in equilibrium and not as an agent's optimization directly.

The values for the coefficients in D_1 determine the conditions under which a single signal is unequivocally chosen. The condition $c_2^2 d_{11} + d_{22} \geq c_2(d_{12} + d_{21})$ is satisfied if

$$\sigma_\varepsilon^2 \left(\sigma_{F_1}^2 + c^2\sigma_{F_2}^2\right) + s\left(1 + c^2\right) \geq 0,$$

which is always the case.

It is also possible to follow the approach in Mondria (2010), where the main discussion is on the investor's choice of a symmetric signal $c_1 = c_2 = c^*$. Note that a symmetric

[7] See Appendix D for the obtention of coefficient matrix $D_1 = \frac{1}{2}\left[\Sigma_F - \Sigma_F C^\top \left[C\Sigma_F C^\top + \Sigma_\varepsilon\right]^{-1} C\Sigma_F\right]^{-\frac{1}{2}}$.

signal does not result in redundant information. Therefore, the situation is identical to a one-dimensional signal with weight $C = [1, c^*]$. The following corollary summarizes this result.

Corollary. *A symmetric equilibrium with $c_1 = c_2$, is equivalent to a one-signal equilibrium.*

Proof. The second order condition,

$$\frac{2\left(\sigma_{F_1}^2 \sigma_{F_2}^2 (c_1 - c_2)^2 e^{2\kappa}\right)\left(\left(c_2^2 \sigma_{F_2}^2 + \sigma_{F_1}^2\right)\gamma_{12} - \sigma_{F_1}^2 \sigma_{F_2}^2 (c_1 - c_2)^2 \gamma_2\right)}{\left(\left(c_2^2 \sigma_{F_2}^2 + \sigma_{F_1}^2\right) + \sigma_{F_1}^2 \sigma_{F_2}^2 (c_1 - c_2)^2 \sigma_{\varepsilon_1}^{-2}\right)^3} \geq 0,$$

is satisfied with equality whenever $c_1 = c_2 = c^*$ where there is a continuum of equilibria that can be represented by a single signal $\tilde{Y} = \tilde{F}_1 + c^* \tilde{F}_2 + \tilde{\varepsilon}$, with the same informational content. □

1.4.3 One-dimensional signal and asset prices

We have seen how a rationally inattentive insider can optimally choose an unidimensional signal to select a portfolio. In a noisy rationally inattentive equilibrium, the covariance matrix for prices, $\Sigma_p = \mathbb{E}\left[(p - \mathbb{E}(p))(p - \mathbb{E}(p))^\top\right]$, is given by

$$\Sigma_p = \frac{1}{2}\left(\Sigma_F - \Sigma_F C^\top \left[C \Sigma_F C^\top + \Sigma_\varepsilon\right]^{-1} C \Sigma_F\right). \tag{1.9}$$

The prices are set by a risk-neutral market maker whose zero expected-profit rule uses the order flow as the signal to filter information from the insiders. Note that the information available to the market maker is that which is embedded in the market orders, but distorted through noise trading. Nonetheless, the only source of comovement in the order flow is generated by the insider's information through the market orders. If the insider had chosen an information structure that transmitted the independence of the fundamental values, i.e., a diagonal weighting matrix C, the price covariance matrix Σ_p would be diagonal as well.

Figure 1.1 shows the asset price covariance as a function of different capacity levels, κ, given some values for the diagonal covariance matrix of asset payoffs, Σ_F . It shows that for capacity levels close to zero, i.e., the insider's capacity is not large enough to enable

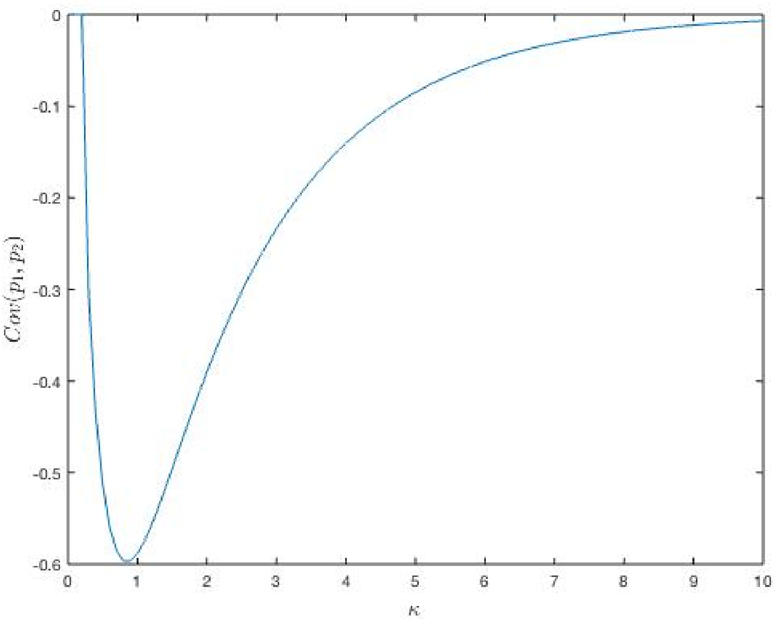

Figure 1.1: Covariance of prices as a function of capacity level κ. The plotted results correspond to parameter values of $\sigma^2_{F_1} = 0.7$ and $\sigma^2_{F_2} = 0.6$.

any information to be processed, the price covariance is zero due to the impossibility of learning about asset payoffs. As the capacity level increases, the covariance rapidly becomes negative as the higher capacity allows the agent to learn simultaneously about both assets, it reaches the largest effect around $\kappa = 0.8$ for this particular case. After this threshold, the capacity allows the investor to identify the information about payoffs more precisely. Therefore, higher capacity leads to lower magnitude comovement. For high capacity levels, the higher precision dominates the effect over the price covariance and, as a result, the covariance disappears.

Proposition 3. *The covariance of prices is negative for all $c > 0$.*

Proof. Let σ_{p12} be the covariance of the prices, that is, the off-diagonal term in matrix Σ_p. Since the asset payoffs $\tilde{F}$ are independent, i.e., Σ_F is diagonal, σ_{p12} is proportional to the negative of the off-diagonal term in the matrix $\Sigma_F C^\top \left[C\Sigma_F C^\top + \Sigma_\varepsilon\right]^{-1} C\Sigma_F$. In fact, $\sigma_{p12} = -c\sigma^2_{F_1}\sigma^2_{F_2} \times a$ where $a > 0$. □

The covariance of asset prices shown in Figure 1.1 depicts how minimal levels of capacity, κ, generate zero covariance since no information can be processed under this capacity level. As the insider starts learning on a single signal, the covariance of the informational advantage becomes negative, and this leads to the negative correlation of market orders and, therefore, negative price covariance. This result is consistent with the negative covariance in asset prices described in Mondria (2010) for certain degrees of risk aversion.

1.4.4 Trading and independent information

As mentioned earlier, there are multiple equilibria in the model. We have already discussed an equilibrium with correlated prices. The next step is to show that the only possible source of price comovement is the information structure. Consider the risk-neutral market structure described above. Now, assume that the insider is allowed to learn from each asset separately. That is, the C matrix can only be diagonal. Moreover, given the properties of mutual information, it can be normalized as an identity matrix. Therefore, the signal becomes $\tilde{Y} = \tilde{F} + \tilde{\varepsilon}$. Similarly, the informational advantage is

$$\tilde{\xi} = \Sigma_F \left(\Sigma_F + \Sigma_\varepsilon\right)^{-1} \left[\tilde{Y} - \bar{F}\right],$$

where $\Sigma_F \left(\Sigma_F + \Sigma_\varepsilon\right)^{-1}$ is a diagonal matrix. More precisely, the informational advantages are orthogonal to each other

$$\tilde{\xi}_i = \frac{\sigma^2_{F_i}}{\sigma^2_{F_i} + \sigma^2_{\varepsilon_i}} \left(\tilde{F}_i - \bar{F}_i + \tilde{\varepsilon}_i\right),$$

for $i = 1, 2$. Unsurprisingly, the effects over the market orders are that each market order depends on its own informational advantage only. That is, the matrix $D_1 = \left[\Sigma_F - \Sigma_F \left[\Sigma_F + \Sigma_\varepsilon\right]^{-1} \Sigma_F\right]^{-\frac{1}{2}}$ is diagonal. The portfolio selection for this investor after he observes the signal is

$$x_i\left(\tilde{\xi}_i\right) = \left(\sigma^{-2}_{F_i} + \sigma^{-2}_{\varepsilon_i}\right)^{\frac{1}{2}} \tilde{\xi}_i,$$

for $i = 1, \ldots, n$. Note that the market orders response to the informational advantage is increasing in the corresponding signal precision $\sigma^{-2}_{\varepsilon_i}$. The pricing function for each market is

$$p_i\left(\tilde{\omega}_i\right) = \bar{F}_i + \frac{1}{2}\left(\frac{\sigma^2_{F_i}\sigma^2_{\varepsilon_i}}{\sigma^2_{F_i} + \sigma^2_{\varepsilon_i}}\right)^{\frac{1}{2}} \tilde{\omega}_i,$$

for $i = 1, \ldots, n$ where all prices are independent. The information-processing constraint in (1.1) becomes

$$\ln \prod_{i=1}^{n} \left(\frac{\sigma^2_{F_i}}{\sigma^2_{\varepsilon_i}} + 1\right) \leq 2\kappa, \tag{1.10}$$

which is the product of the mutual information for each of the asset payoffs with their corresponding signal.

Before the introduction of the aggregate information-processing constraint, this market structure resembles n separate Kyle (1985) type monopolistic insiders under imperfect information, each trading a different asset. Nonetheless, after the information choice is made endogenous according to an entropy-learning approach, the insider's problem is no longer equivalent to a set of n separate choices. A similar analysis was performed by van Nieuwerburgh and Veldkamp (2010) for perfectly competitive markets where they consider different information acquisition technologies and preferences toward risk are considered. The results under risk aversion are similar since the solution to the nonlinear system conformed by the coefficient matrices is diagonal as well. Thus, in this model the only possible source of price comovement is the informational channel.

In short, this section has shown that even without risk aversion, once imperfectly competitive markets are assumed, a negative covariance arises for a risk-neutral agent. So, this covariance is the result of a hedging strategy when there is very limited information available. In the following section, the insider is assumed to be risk-averse.

1.5 Insider trading under risk aversion

The previous section characterized the effects of introducing an information-processing constraint to a risk-neutral monopolistic insider. The non-competitive setup defined here allows full characterization of the equilibria for a risk-neutral insider, and in particular, the effects on the covariance of asset prices in the equilibrium with a unidimensional signal. However, this analytical framework does not permit modeling of the insider's preferences towards risk.

Facing an information-processing constraint such as the one introduced in Section 3 acts as an additional source of uncertainty. This is the case because the signal's lack of precision increases the uncertainty experienced by the insider. In order to analyze how a direct preference towards risk affects the market outcome, the insider must be allowed to be risk-averse. This section considers the case of a monopolistic insider whose preferences exhibit constant absolute risk aversion (CARA). The risk-neutral case is a particular one of a risk-averse insider where the coefficient of risk aversion is zero.

1.5.1 Portfolio selection and risk aversion

The model considered here follows a risk-averse version of Caballé and Krishnan (1994) as in Vitale (2012). The agents interacting in the market for two securities are the same as in section 4, an informed monopolistic investor, random liquidity (noise) traders and a risk-neutral market maker setting the prices. The informed and uninformed investors simultaneously place their orders for the two assets, and the market maker sets the price after observing the order flow.

The monopolistic insider has access to a signal $\tilde{Y}$ about the asset payoffs and selects a his portfolio by maximizing his expected utility conditional to this information, and. His preferences towards risk are described by the CARA utility function$u\left(W\right) = -\exp\left(-\rho W\right)$.[8] Given a signal $\tilde{Y}$, the insider's expected utility of selecting a portfolio x is

$$\begin{aligned} E\left(u\left(\tilde{W}\left(x\right)\right)|\tilde{Y}\right) &= -\mathbb{E}\left[\exp\left(-\rho\left(W_0 + x^\top(\tilde{F} - p(\tilde{\omega}))\right)\right)|\tilde{Y}\right] \\ &= -\exp\left(-\rho E\left[W_0 + x^\top(\tilde{F} - p(\tilde{\omega}))|\tilde{Y}\right] + \frac{\rho^2}{2}V\left[W_0 + x^\top(\tilde{F} - p(\tilde{\omega}))|\tilde{Y}\right]\right). \end{aligned} \tag{1.11}$$

Then, the insider selects the portfolio that solves

$$\max_{\{x\}} \mathbb{E}\left[W_0 + x^\top(\tilde{F} - p(\tilde{\omega}))|\tilde{Y}\right] - \frac{\rho}{2}\mathbb{V}\left[W_0 + x^\top(\tilde{F} - p(\tilde{\omega}))|\tilde{Y}\right]. \tag{1.12}$$

Note that in this static context maximizing (1.11)and (1.12) is equivalent since $g(a) = -\exp(-a)$ is increasing in a.[9] Vitale (2012) shows in a more general model that after introducing constant absolute risk aversion, the pricing function remains linear in the order flow and the portfolio selected is linear in terms of informational advantages. Hence, the functional forms $x\left(\tilde{\xi}\right) = \mathcal{B}_0 + \mathcal{B}_1\tilde{\xi}$ and $p\left(\tilde{\omega}\right) = \bar{F} + \mathcal{A}_0 + \mathcal{A}_1\tilde{\omega}$ can be assumed to

[8] The preferences towards risk are also referred to as preferences for an early resolution of uncertainty.

[9] An implicit assumption that is being made here is that I am transforming the function through $f(a) = -\ln(-a)$. This has no major implication for static environments but says that should I face a dynamic environment the agent has preferences for early resolution of uncertainty.

characterize the solution. In fact, by setting $\rho = 0$ in (1.12) the risk-neutral solution can be achieved. The coefficient matrices $\mathcal{A}_0$, $\mathcal{A}_1$, $\mathcal{B}_0$ and $\mathcal{B}_1$ are given by the solution of the following nonlinear system:

$$\mathcal{A}_0 = -\mathcal{A}_1 \bar{z},$$
$$\mathcal{B}_0 = 0,$$
$$\mathcal{A}_1 = \left[I - \Sigma_z \left[\Sigma_F C^\top \left[C \Sigma_F C^\top + \Sigma_\varepsilon \right]^{-1} C \Sigma_F \mathcal{B}_1^\top \right]^{-1} \right]^{-1},$$

$$\begin{aligned}
\mathcal{B}_1^{-1} = 2 & \left[I - \Sigma_z \left(\mathcal{B}_1^\top \right)^{-1} \left[\Sigma_z C^\top \left[C \Sigma_F C^\top + \Sigma_\varepsilon \right]^{-1} C \Sigma_z \right]^{-1} \right]^{-1} \\
& + \rho \left[\Sigma_F - \left[\Sigma_z C^\top \left[C \Sigma_F C^\top + \Sigma_\varepsilon \right]^{-1} C \Sigma_z \right]^{-1} C \Sigma_F \right. \\
& + \left[I - \left[\Sigma_z C^\top \left[C \Sigma_F C^\top + \Sigma_\varepsilon \right]^{-1} C \Sigma_z \right]^{-1} \mathcal{B}_1{}^{-1} \Sigma_z \right]^{-1} \Sigma_z \\
& \left. \left[I - \Sigma_z \left(\mathcal{B}_1^\top \right)^{-1} \left[\Sigma_z C^\top \left[C \Sigma_F C^\top + \Sigma_\varepsilon \right]^{-1} C \Sigma_z \right]^{-1} \right]^{-1} \right],
\end{aligned}$$

which does not have an analytical solution. However, this system provides the conditions that any equilibrium must satisfy. The insider, as in section 4, is aware of the linearity of pricing functions and market orders, when the attention allocation problem is solved.

1.5.2 Attention allocation

In the first stage, the insider needs to design the optimal information structure for a given capacity constraint κ. Once again, the signal takes the form $\tilde{Y} = C\tilde{F} + \tilde{\varepsilon}$. The attention allocation problem that the insider faces corresponds to the allocation of weights on matrix C and precision parameters on the covariance matrix Σ_ε such that the insider's expected profits are maximized.

At first, the monopolist can anticipate the market maker's reaction function. Thus, in the first stage the insider chooses the information structure that maximizes the following expected utility

$$\max_{\{C,\Sigma_\varepsilon\}} \mathbb{E}\left[\mathbb{E}\left[x^\top\left(\tilde{\xi}\right)\left(\tilde{F}-\left(\bar{F}-\mathcal{A}_1\bar{z}+\mathcal{A}_1\left(x\left(\tilde{\xi}\right)+\tilde{z}\right)\right)\right)|\tilde{Y}\right]\right.$$
$$\left.-\frac{\rho}{2}\mathbb{V}\left[x^\top\left(\tilde{\xi}\right)\left(\tilde{F}-\left(\bar{F}-\mathcal{A}_1\bar{z}+\mathcal{A}_1\left(x\left(\tilde{\xi}\right)+\tilde{z}\right)\right)\right)|\tilde{Y}\right]\right]$$
$$\text{subject to } \ln|\Sigma_F|-\ln\left|\Sigma_F-\Sigma_F C^\top\left[C\Sigma_F C^\top+\Sigma_\varepsilon\right]^{-1}C\Sigma_F\right|\leq 2\kappa.$$

Or, equivalently

$$\max_{\{C,\Sigma_\varepsilon\}} \mathbb{E}\left[\left(\mathcal{B}_1\tilde{\xi}\right)^\top\left[\tilde{\xi}-\mathcal{A}_0-\mathcal{A}_1\bar{z}\right]-\left(\mathcal{B}_1\tilde{\xi}\right)^\top\mathcal{A}_1\left(\mathcal{B}_1\tilde{\xi}\right)\right.$$
$$\left.-\frac{\rho}{2}\left[\left(\mathcal{B}_1\tilde{\xi}\right)^\top\mathbb{V}\left[\tilde{F}|\tilde{Y}\right]\left(\mathcal{B}_1\tilde{\xi}\right)+\left(\mathcal{B}_1\tilde{\xi}\right)^\top\mathcal{A}_1^\top\Sigma_z\mathcal{A}_1\left(\mathcal{B}_1\tilde{\xi}\right)\right]\right]$$
$$\text{subject to } \ln|\Sigma_F|-\ln\left|\Sigma_F-\Sigma_F C^\top\left[C\Sigma_F C^\top+\Sigma_\varepsilon\right]^{-1}C\Sigma_F\right|\leq 2\kappa,$$

where the values for the conditional mean, $\mathbb{E}\left[\tilde{F}|\tilde{Y}\right]$, and variance, $\mathbb{V}\left[\tilde{F}|\tilde{Y}\right]$, and the informational advantage $\tilde{\xi}=\mathbb{E}\left[\tilde{F}|\tilde{Y}\right]-\bar{F}$, are known (see Appendix A). In the end, the objective function follows the quadratic form,

$$\max_{\{C,\Sigma_\varepsilon\}} \mathbb{E}\left[\tilde{\xi}^\top\mathcal{B}_1^\top\left(I-\mathcal{A}_1\mathcal{B}_1\right.\right.$$
$$\left.\left.-\frac{\rho}{2}\left[\left[\Sigma_F-\Sigma_F C^\top\left[C\Sigma_F C^\top+\Sigma_\varepsilon\right]^{-1}C\Sigma_F\right]\mathcal{B}_1+\mathcal{A}_1^\top\Sigma_z\mathcal{A}_1\mathcal{B}_1\right]\right)\tilde{\xi}\right],$$

which is of the type $\mathbb{E}\left[\tilde{\xi}^\top\mathbf{A}\tilde{\xi}\right]$, where $\tilde{\xi}^\top\mathbf{A}\tilde{\xi}$ is a scalar. Therefore, the expected utility can be computed as the trace of a product matrix, and the attention allocation problem is reduced to

$$\max_{\{C,\Sigma_\varepsilon\}} Tr\left[\mathbf{A}\Sigma_\xi\right]$$

$$\text{subject to } \left(\left(c_1^2\sigma_{F_2}^2+\sigma_{F_1}^2\right)\sigma_{\varepsilon_1}^{-2}+\sigma_{F_1}^2\sigma_{F_2}^2(c_1-c_2)^2\sigma_{\varepsilon_1}^{-2}\sigma_{\varepsilon_2}^{-2}+\left(c_2^2\sigma_{F_2}^2+\sigma_{F_1}^2\right)\sigma_{\varepsilon_2}^{-2}\right)=e^{2\kappa}-1,$$

where $\Sigma_\xi=\Sigma_F-\Sigma_F C^\top\left[C\Sigma_F C^\top+\Sigma_\varepsilon\right]^{-1}C\Sigma_F$. This is equivalent to (1.5). Moreover, the constraint remains unchanged for risk-averse insiders (1.7), since the informativeness of the signal does not depend on the insider's preferences.

Following the same procedure as in Section 4, the condition to allocate attention to a single signal is determined by the second order condition. Then, we proceed to check when is the objective function convex in the precision parameters $\sigma_{\varepsilon_1}^{-2}$ and $\sigma_{\varepsilon_2}^{-2}$, and how the condition changes as the the risk-aversion parameter varies.

The second-order condition of the previous problem is

$$\frac{\partial^2}{\partial \sigma_{\varepsilon_1}^{-2}} = \frac{\left(\sigma_{F_1}^2 \sigma_{F_2}^2 \left(c_1 - c_2\right)^2 e^{2\kappa}\right) \left(\sigma_{F_1}^2 \sigma_{F_2}^2 \left(c_1 - c_2\right)^2 \varphi_1 - \left(c_2^2 \sigma_{F_2}^2 + \sigma_{F_1}^2\right) \varphi_2\right) - \frac{1}{2} \rho \varphi_3}{\left(\left(c_2^2 \sigma_{F_2}^2 + \sigma_{F_1}^2\right) + \sigma_{F_1}^2 \sigma_{F_2}^2 (c_1 - c_2)^2 \sigma_{\varepsilon_1}^{-2}\right)^3},$$

where $\varphi_1, \varphi_2, \varphi_3 > 0$. Such a condition is decreasing in terms of the coefficient of risk aversion, ρ, and hence there is a value ρ^* large enough for the second-order derivative to be negative. As a result, risk-averse agents may not choose to allocate attention to only one signal under a scenario where the risk-neutral investor chooses to allocate all attention to one signal. Section 4 discusses how the choice of a one dimensional signal leads to a hedging strategy that induces a negative correlation in the market orders. Now, we find that as risk aversion grows, an insider is less likely to choose one signal. In the following subsection we characterize the effect of risk aversion on price comovement.

1.5.3 Implications for price comovement

Once the insider has allocated the attention, the signal is observed and he places the market orders and the market maker sets the prices. In equilibrium, the covariance matrix for prices is

$$\Sigma_p = \mathcal{A}_1 \mathcal{A}_1^\top + \mathcal{A}_1 \mathcal{B}_1 \left[\Sigma_F - \Sigma_F C^\top \left[C \Sigma_F C^\top + \Sigma_\varepsilon\right]^{-1} C \Sigma_F\right] \mathcal{B}_1^\top \mathcal{A}_1^\top. \tag{1.13}$$

The price covariance matrix in (1.9) is a particular case of (1.13).Thus, the asset price covariance is always negative for the risk-averse insider as well, since the only source of covariance comes from the negative correlation of the informational advantage. However, introducing risk aversion scales down the effect. In particular, the price covariance is proportional to the off-diagonal term in the covariance matrix of the informational advantages $\frac{-c\sigma_{F_1}^2 \sigma_{F_2}^2}{\sigma_{F_1}^2 + c^2 \sigma_{F_2}^2 + \sigma_\varepsilon^2}$, that is, $\sigma_{p12} \propto -c\sigma_{F_1}^2 \sigma_{F_2}^2$. Moreover, this term determines the sign of the price covariance.

Figure 2.2 shows how the covariance changes as the risk aversion increases, compared to the risk-neutral case discussed in the previous section. The results for a perfectly informed

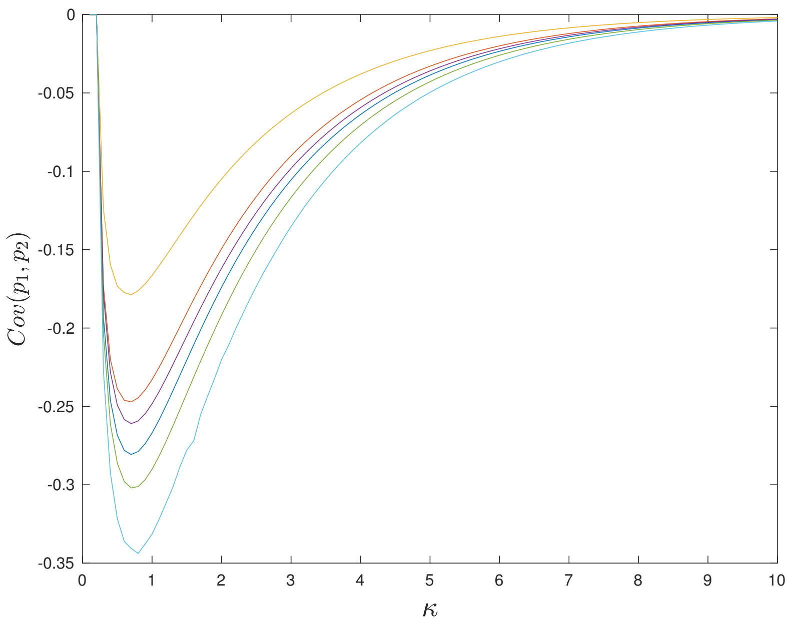

Figure 1.2: Covariance of prices as a function of capacity level κ for different levels of risk aversion . The plotted results correspond to parameter values of $\sigma^2_{F_1} = 0.7$, $\sigma^2_{F_2} = 0.6$ and $\rho = 0.5$, 1, 5, 10, 20 and 50.

insider, $\kappa \to \infty$, and a completely uninformed insider, $\kappa = 0$, remain unchanged when risk aversion is introduced. A perfectly informed agent converges to zero price covariance, since the asset payoffs are observable in the signal. A completely uninformed insider does not induce price covariance because he is not able to produce an informative signal.

The response of the price covariance to the risk aversion goes in a different direction. Figure 2.2 shows how the risk-aversion coefficient scales down the magnitude of the covariance. Thus, the covariance in market orders is negative as a hedging response to the (lack of) information. Furthermore, as the risk aversion grows, the marginal effect of the informational advantages over the market orders is scaled down, reducing the magnitude of the effect over the price covariance. Finally, consider an infinitely risk-averse insider, $\rho \to \infty$, the marginal effect is zero of any informational advantage in the market orders is annulled by risk aversion, i.e., an infinitely risk-averse insider never trades unless there is perfect information. In this case, prices are determined only by noise trading, which is uncorrelated across markets, which makes the prices uncorrelated as well.

1.6 Concluding remarks

This paper provides a characterization of asset price covariance induced by a rational investor's optimal information choice. Asset price covariance arises in equilibrium in the context of a portfolio selection of assets whose fundamental values are statistically independent. In addition, this result is partially due to the market power that inside information gives to an investor. Then, I show that, under normality assumption, a different explanation for the asset price covariance cannot be found in a market where noise trading is uncorrelated across assets and asset payoffs are statistically independent. In the end, a limited information-processing capacity induces the insider to favor an information structure where aggregate precision dominates other alternatives. Thus, the insider prefers to have one signal with the highest possible precision, and that is informative about all assets, over asset-specific signals. As a result, the information set fails to identify the origin of an unexpected shock to a specific asset, which causes an impact in the whole vector of market orders. Later this shock is reflected on both asset prices.

In general, when there is an information-processing constraint, the pricing decision is made after the correlated order flow is observed. Therefore, the market maker observes a noisy spread of the insider's information set, and the prices reflect this information. Since the price covariance can be traced back to the information set, I characterize asset price covariance as the information processing constraint slacks. In the analysis, I observe that the magnitude of the price covariance converges to zero for minimal capacity levels. It then grows rapidly as the investor starts to become informed. Finally, I show that after the processing capacity crosses a threshold, an increase in the capacity reduces the covariance's magnitude until it disappears to zero for larger levels of information-processing capacity.

Additionally, the paper provides some insight into the effect of risk aversion over the price covariance. Risk aversion scales down the magnitude of market orders, i.e., the insider displays hedging behavior in portfolio selection, but the quantities are lower as risk aversion increases. As a result, the magnitude of the covariance diminishes as the risk aversion increases. Further analysis may include the effect of short-lived and long-lived information in a dynamic extension of this model.

Appendix

A. Moments of $\tilde{F}|\tilde{Y}$

Given that the random vectors $\tilde{F}$ and $\tilde{Y}$ are MN, then $\begin{bmatrix} \tilde{F} \\ \tilde{Y} \end{bmatrix} \sim MN\left(\begin{pmatrix} \bar{F} \\ C\bar{F} \end{pmatrix}, \begin{bmatrix} \Sigma_F & \Sigma_{FY}^\top \\ \Sigma_{FY} & C\Sigma_F C^\top + \Sigma_\varepsilon \end{bmatrix}\right)$. The expected value of vector $\tilde{F}$ given $\tilde{Y}$ is

$$\mathbb{E}\left(\tilde{F}|\tilde{Y}\right) = \bar{F} + \Sigma_{FY}\left[C\Sigma_F C^\top + \Sigma_\varepsilon\right]^{-1}\left[\tilde{Y} - C\bar{F}\right].$$

And the conditional covariance matrix of $\tilde{F}$ given $\tilde{Y}$is

$$\begin{aligned}
\mathbb{V}\left(\tilde{F}|\tilde{Y}\right) \equiv \Sigma_{F|Y} &= \Sigma_F - \Sigma_{FY}\Sigma_Y^{-1}\Sigma_{FY}^\top \\
&= \Sigma_F - \Sigma_{FY}\left[C\Sigma_F C' + \Sigma_\varepsilon\right]^{-1}\Sigma_{FY}^\top.
\end{aligned}$$

Also recall that $\Sigma_{FY} = \mathbb{E}\left[\left(\tilde{F} - \bar{F}\right)\left(\tilde{Y} - C\bar{F}\right)^\top\right]$ and since $\tilde{Y} = C\tilde{F} + \varepsilon$,

$$\begin{aligned}
\Sigma_{FY} &= \mathbb{E}\left[\left(\tilde{F} - \bar{F}\right)\left(C\tilde{F} + \varepsilon - C\bar{F}\right)^\top\right] \\
&= \mathbb{E}\left[\left(\tilde{F} - \bar{F}\right)\left(\tilde{F}^\top C^\top - \bar{F}^\top C^\top + \varepsilon^\top\right)\right] \\
&= \mathbb{E}\left[\left(\tilde{F} - \bar{F}\right)\tilde{F}^\top C^\top - \left(\tilde{F} - \bar{F}\right)\bar{F}^\top C^\top + \left(\tilde{F} - \bar{F}\right)\varepsilon^\top\right] \\
&= \mathbb{E}\left[\left(\tilde{F} - \bar{F}\right)\tilde{F}^\top C^\top - \left(\tilde{F} - \bar{F}\right)\bar{F}^\top C^\top\right] \\
&= \mathbb{E}\left[\left(\tilde{F} - \bar{F}\right)\left(\tilde{F}^\top - \bar{F}^\top\right)C^\top\right] \\
&= \mathbb{E}\left[\left(\tilde{F} - \bar{F}\right)\left(\tilde{F} - \bar{F}\right)^\top\right]C^\top \\
&= \Sigma_F C^\top,
\end{aligned}$$

so the conditional moments can be re-written as

$$\begin{aligned}
\mathbb{E}\left(\tilde{F}|\tilde{Y}\right) &= \bar{F} + \Sigma_F C^\top\left[C\Sigma_F C^\top + \Sigma_\varepsilon\right]^{-1}\left[\tilde{Y} - C\bar{F}\right] \\
\mathbb{V}\left(\tilde{F}|\tilde{Y}\right) \equiv \Sigma_{F|Y} &= \Sigma_F - \Sigma_F C^\top\left[C\Sigma_F C^\top + \Sigma_\varepsilon\right]^{-1}C\Sigma_F.
\end{aligned}$$

The same analysis can be used to derive other conditional moments throughout the text.

B. Normalization of the signal

There are multiple signals with the same informational content. Normalization is required to avoid multiplicity of equilibria. This section follows Admati (1985); Mondria (2010) to normalize the covariance matrix Σ_ε into a diagonal covariance matrix. Let $\Sigma_\varepsilon = P\Lambda P^\top$ where Λ is diagonal, and $P^\top P = I$, I can re-write the signal so that

$$\begin{aligned}\tilde{Y}^* &= P^{-1}\tilde{Y}\\ &= P^{-1}C\tilde{F} + P^{-1}\tilde{\varepsilon}\\ &= C^*\tilde{F} + \tilde{\varepsilon}^*,\end{aligned}$$

where $\Sigma_{\varepsilon^*} = \Lambda$. Note that information is invariant to any linear invertible transformation, therefore $I\left(\tilde{F};\tilde{Y}\right) = I\left(\tilde{F};\tilde{Y}^*\right)$ so the capacity constraint (1.1) remains unchanged. Also note that the term $C^\top\Sigma_\varepsilon^{-1}C = C^{*\top}\Lambda^{-1}C^*$. Hence any equilibrium that satisfies this transformation is indistinguishable.

Similarly, Mondria (2010) also normalizes the weight matrix C and shows that provided a diagonal non-singular matrix Γ and provided a diagonal covariance matrix Σ_ε (which can be obtained by the normalization described above)

$$\begin{aligned}\tilde{Y}^* &= \Gamma\tilde{Y}\\ &= \Gamma C\tilde{F} + \Gamma\tilde{\varepsilon}\\ &= C^*\tilde{F} + \tilde{\varepsilon}^*,\end{aligned}$$

where Σ_{ε^*} is diagonal since Γ is diagonal. And just as before, one can see that any equilibrium that follows this transformation is indistinguishable since $C^\top\Sigma_\varepsilon^{-1}C = C^{*\top}\left[\Gamma\Sigma_\varepsilon\Gamma\right]^{-1}C^* = C^{*\top}\Sigma_{\varepsilon^*}^{-1}C^*$. So the same argument as before can be used. This permits normalization using either column. Generically, it is normalized by the first column.

C. Capacity constraint for the 2×2 case and uncorrelated fundamentals and liquidity demands

This section considers an illustrative case to understand how a capacity constraint operates in the construction of a signal for a two dimensional vector. [10] The capacity constraint under the normality assumption is

$$\ln |\Sigma_F| - \ln \left| \Sigma_F - \Sigma_F C^\top \left[C \Sigma_F C^\top + \Sigma_\varepsilon \right]^{-1} C \Sigma_F \right| \leq 2\kappa.$$

Let $\Sigma_F = \sigma_F^2 I$ and Σ_ε diagonal. Then, the capacity constraint becomes

$$\begin{aligned} \ln \left| \sigma_F^2 I \right| - \ln \left| \sigma_F^2 I - \sigma_F^2 C^\top \left[C C^\top + \frac{1}{\sigma_F^2} \Sigma_\varepsilon \right]^{-1} C \right| &\leq 2\kappa \\ \ln \left(\sigma_F^2 \right)^2 - \ln \left[\left(\sigma_F^2 \right)^2 \left| I - C^\top \left[C C^\top + \frac{1}{\sigma_F^2} \Sigma_\varepsilon \right]^{-1} C \right| \right] &\leq 2\kappa \\ - \ln \left| I - C^\top \left[C C^\top + \frac{1}{\sigma_F^2} \Sigma_\varepsilon \right]^{-1} C \right| &\leq 2\kappa. \end{aligned}$$

The previous expression can be re-written as

$$\begin{aligned} - \ln \left(\left| C C^\top + \frac{1}{\sigma_F^2} \Sigma_\varepsilon \right|^{-1} \left| C C^\top + \frac{1}{\sigma_F^2} \Sigma_\varepsilon - C C^\top \right| \right) &\leq 2\kappa \\ \ln \frac{\left| \sigma_F^2 C C^\top + \Sigma_\varepsilon \right|}{|\Sigma_\varepsilon|} &\leq 2\kappa, \end{aligned}$$

which follows from the properties of determinants.[11] Note that for this expression it follows automatically that for any finite value of κ it is not possible to eliminate the noise of the signal. [12]

[10] The unidimensional case is exposed in detail in Wiederholt (2010)

[11] In particular that $\det (X + AB) = \det (X) \det \left(I_n + B X^{-1} A \right)$.

[12] If $n = 1$, then it follows immediately and the constraint becomes the most widely used as described in Wiederholt (2010): $\frac{c \sigma_F^2}{\sigma_\varepsilon^2} \leq e^{2\kappa} - 1$. This expression also provides the intuition for the indistinguishability of equilibria.

D. Finding the coefficient matrices for price and orders in the second stage

Risk-neutral insider

Note that the problem in the second stage is

$$\max_{\{x\}} \mathbb{E}\left[x^\top(\tilde{F} - (Q_0 + Q_1(x + \tilde{z})))|\tilde{Y}\right],$$

where the first order conditions are

$$\begin{aligned} x :& \mathbb{E}\left[\tilde{F}|\tilde{Y}\right] - D_0 - D_1\bar{z} = 2D_1 x \\ x =& \frac{1}{2}[D_1]^{-1}\left[\mathbb{E}\left[\tilde{F}|\tilde{Y}\right] - D_0 - D_1\bar{z}\right]. \end{aligned} \tag{1.14}$$

Substitute $\mathbb{E}\left[\tilde{F}|\tilde{Y}\right] = \bar{F} + \Sigma_F C^\top \left[C\Sigma_F C^\top + \Sigma_\varepsilon\right]^{-1}\left[\tilde{Y} - C\bar{F}\right]$. See Appendix A. The market order solves the previous equation and assuming a linear equilibrium it becomes

$$x = \frac{1}{2}D_1^{-1}\left[\bar{F} + \underbrace{\Sigma_F C^\top \left[C\Sigma_F C^\top + \Sigma_\varepsilon\right]^{-1}\left[\tilde{Y} - C\bar{F}\right]}_{\tilde{\xi}} - D_0 - D_1\bar{z}\right] \tag{1.15}$$

$$x(\tilde{\xi}) = \frac{1}{2}D_1^{-1}\tilde{\xi} + \frac{1}{2}D_1^{-1}\left[\bar{F} - D_0\right] - \frac{1}{2}\bar{z}. \tag{1.16}$$

Then, since the equilibrium functional form of $x\left(\tilde{\xi}\right)$ is known, it can be combined with the first order condition to obtain

$$Q_0 + Q_1\tilde{\xi} = \frac{1}{2}D_1^{-1}\tilde{\xi} + \frac{1}{2}D_1^{-1}\left[\bar{F} - D_0\right] - \frac{1}{2}\bar{z},$$

which allows us to define the following relations between coefficients:

$$\begin{aligned} Q_0 &= \frac{1}{2}D_1^{-1}\left[\bar{F} - D_0\right] - \frac{1}{2}\bar{z} \\ Q_1 &= \frac{1}{2}D_1^{-1}. \end{aligned}$$

The market maker now takes this order to set the price according to

$$\begin{aligned} \tilde{p} &= p(\tilde{\omega}) \\ &= \mathbb{E}\left(\tilde{F}|\tilde{\omega}\right), \end{aligned}$$

given that market makers are risk-neutral and face Bertrand competition. Note that since $\tilde{\omega} = x\left(\tilde{\xi}\right)+\tilde{z}$, then $\tilde{\omega} \sim MN\left(Q_0+\bar{z}, Q_1\left[\Sigma_F - \Sigma_F C^\top\left[C\Sigma_F C^\top+\Sigma_\varepsilon\right]^{-1} C\Sigma_F\right] B_1^\top + \Sigma_z\right)$ therefore $\tilde{\omega}$ and $\tilde{F}$ are follow a multivariate normal distribution where $\Sigma_{F\omega} = \left[\Sigma_F - \Sigma_F C^\top\left[C\Sigma_F C^\top+\Sigma_\varepsilon\right]^{-1} C\Sigma_F\right] Q_1^\top$. Then,

$$
\begin{aligned}
p(\tilde{\omega}) = \mathbb{E}\left(\tilde{F}|\tilde{\omega}\right) = \bar{F} + &\left[\Sigma_F - \Sigma_F C^\top\left[C\Sigma_F C^\top+\Sigma_\varepsilon\right]^{-1} C\Sigma_F\right] Q_1^\top \\
&\left[Q_1\left[\Sigma_F - \Sigma_F C^\top\left[C\Sigma_F C^\top+\Sigma_\varepsilon\right]^{-1} C\Sigma_F\right] Q_1^\top + \Sigma_z\right]^{-1}(\tilde{\omega} - Q_0 - \bar{z}).
\end{aligned}
$$

As a result, the following equality holds:

$$
\begin{aligned}
D_1 &= \left[\Sigma_F - \Sigma_F C^\top\left[C\Sigma_F C^\top+\Sigma_\varepsilon\right]^{-1} C\Sigma_F\right] Q_1^\top \left[Q_1\left[\Sigma_F - \Sigma_F C^\top\left[C\Sigma_F C^\top+\Sigma_\varepsilon\right]^{-1} C\Sigma_F\right] Q_1^\top + \Sigma_z\right]^{-1} \\
&= \left[Q_1\left[\Sigma_F - \Sigma_F C^\top\left[C\Sigma_F C^\top+\Sigma_\varepsilon\right]^{-1} C\Sigma_F\right]\left[\left[\Sigma_F - \Sigma_F C^\top\left[C\Sigma_F C^\top+\Sigma_\varepsilon\right]^{-1} C\Sigma_F\right] Q_1^\top\right]^{-1}\right. \\
&\left.+\Sigma_z\left[\left[\Sigma_F - \Sigma_F C^\top\left[C\Sigma_F C^\top+\Sigma_\varepsilon\right]^{-1} C\Sigma_F\right] Q_1^\top\right]^{-1}\right]^{-1} \\
&= \left[I + \Sigma_z\left[Q_1^\top\right]^{-1}\left[\Sigma_F - \Sigma_F C^\top\left[C\Sigma_F C^\top+\Sigma_\varepsilon\right]^{-1} C\Sigma_F\right]^{-1}\right]^{-1}.
\end{aligned}
$$

So,

$$
\begin{aligned}
D_1 &= \left[I + \Sigma_z\left[Q_1'\right]^{-1}\left[\Sigma_F - \Sigma_F C^\top\left[C\Sigma_F C^\top+\Sigma_\varepsilon\right]^{-1} C\Sigma_F\right]^{-1}\right]^{-1} \\
D_0 &= \bar{F} - D_1 Q_0 - D_1\bar{z}.
\end{aligned}
$$

The equality $Q_0 = \frac{1}{2}D_1^{-1}\left[\bar{F} - D_0\right] - \frac{1}{2}\bar{z}$ is substituted in the previous equation, $D_0 = \bar{F} - D_1\bar{z}$ to obtain $Q_0 = 0$. Recall that $D_1 = \frac{1}{2}Q_1^{-1}$

$$
\begin{aligned}
\left[2Q_1\right]^{-1} &= \left[Q_1 + \Sigma_z\left[Q_1^\top\right]^{-1}\left[\Sigma_F - \Sigma_F C^\top\left[C\Sigma_F C^\top+\Sigma_\varepsilon\right]^{-1} C\Sigma_F\right]^{-1}\right]^{-1} \\
Q_1 &= \Sigma_z(Q_1')^{-1}\left[\Sigma_F - \Sigma_F C^\top\left[C\Sigma_F C^\top+\Sigma_\varepsilon\right]^{-1} C\Sigma_F\right]^{-1},
\end{aligned}
$$

and substituting into the previous expression for D_1 to get

$$D_1 = \left[2\Sigma_z \left[Q_1^\top\right]^{-1} \left[\Sigma_F - \Sigma_F C^\top \left[C\Sigma_F C^\top + \Sigma_\varepsilon\right]^{-1} C\Sigma_F\right]^{-1}\right]^{-1}$$
$$D_1^{-1} = 4\Sigma_z D_1 \left[\Sigma_F - \Sigma_F C^\top \left[C\Sigma_F C^\top + \Sigma_\varepsilon\right]^{-1} C\Sigma_F\right]^{-1}$$

$$\frac{1}{4} D_1^{-1} \left[\Sigma_F - \Sigma_F C^\top \left[C\Sigma_F C^\top + \Sigma_\varepsilon\right]^{-1} C\Sigma_F\right] = \Sigma_z D_1,$$

given that Σ_z is a covariance matrix

$$\frac{1}{4} \left[\Sigma_F - \Sigma_F C^\top \left[C\Sigma_F C^\top + \Sigma_\varepsilon\right]^{-1} C\Sigma_F\right] = D_1 \Sigma_z D_1$$
$$\frac{1}{4}\Sigma_z^{\frac{1}{2}} \left[\Sigma_F - \Sigma_F C^\top \left[C\Sigma_F C^\top + \Sigma_\varepsilon\right]^{-1} C\Sigma_F\right] \Sigma_z^{\frac{1}{2}} = \Sigma_z^{\frac{1}{2}} D_1 \Sigma_z^{\frac{1}{2}} \Sigma_z^{\frac{1}{2}} D_1 \Sigma_z^{\frac{1}{2}}.$$

Hence,

$$D_1 = \frac{1}{2}\Sigma_z^{-\frac{1}{2}} \left[\Sigma_z^{\frac{1}{2}} \left[\Sigma_F - \Sigma_F C^\top \left[C\Sigma_F C^\top + \Sigma_\varepsilon\right]^{-1} C\Sigma_F\right] \Sigma_z^{\frac{1}{2}}\right]^{\frac{1}{2}} \Sigma_z^{-\frac{1}{2}}.$$

The matrix $\Sigma_z^{\frac{1}{2}}$ corresponds to the unique symmetric positive definite square root of Σ_z, and D_1 is the unique symmetric matrix that solves the equation. In this case D_1 can be obtained through the principal square root matrix on the l.h.s. of the equation. Additionally, the expression for the coefficients of the market orders becomes much simpler if the noise traders' covariance matrix is $\Sigma_z = \sigma_z^2 I$, as was pointed out in Caballé and Krishnan (1994).

So, the price and order that result in this stage are

$$p^s(\tilde{\omega}) = \bar{F} - \frac{1}{2}\Sigma_z^{-\frac{1}{2}} \left[\Sigma_z^{\frac{1}{2}} \left[\Sigma_F - \Sigma_F C^\top \left[C\Sigma_F C^\top + \Sigma_\varepsilon\right]^{-1} C\Sigma_F\right] \Sigma_z^{\frac{1}{2}}\right]^{\frac{1}{2}} \Sigma_z^{-\frac{1}{2}} \bar{z} +$$
$$\frac{1}{2}\Sigma_z^{-\frac{1}{2}} \left[\Sigma_z^{\frac{1}{2}} \left[\Sigma_F - \Sigma_F C^\top \left[C\Sigma_F C^\top + \Sigma_\varepsilon\right]^{-1} C\Sigma_F\right] \Sigma_z^{\frac{1}{2}}\right]^{\frac{1}{2}} \Sigma_z^{-\frac{1}{2}} \tilde{\omega}$$
$$x^s(\tilde{\xi}) = \Sigma_z^{\frac{1}{2}} \left[\Sigma_z^{\frac{1}{2}} \left[\Sigma_F - \Sigma_F C^\top \left[C\Sigma_F C^\top + \Sigma_\varepsilon\right]^{-1} C\Sigma_F\right] \Sigma_z^{\frac{1}{2}}\right]^{-\frac{1}{2}} \Sigma_z^{\frac{1}{2}} \tilde{\xi},$$

and the equilibrium price is

$$p^*\left(\tilde{\xi},\tilde{z}\right) = \bar{F} - \frac{1}{2}\Sigma_z^{-\frac{1}{2}}\left[\Sigma_z^{\frac{1}{2}}\left[\Sigma_F - \Sigma_F C^\top\left[C\Sigma_F C^\top + \Sigma_\varepsilon\right]^{-1} C\Sigma_F\right]\Sigma_z^{\frac{1}{2}}\right]^{\frac{1}{2}}\Sigma_z^{-\frac{1}{2}}\bar{z} +$$
$$\frac{1}{2}\Sigma_z^{-\frac{1}{2}}\left[\Sigma_z^{\frac{1}{2}}\left[\Sigma_F - \Sigma_F C^\top\left[C\Sigma_F C^\top + \Sigma_\varepsilon\right]^{-1} C\Sigma_F\right]\Sigma_z^{\frac{1}{2}}\right]^{\frac{1}{2}}\Sigma_z^{-\frac{1}{2}}$$
$$\left(\Sigma_z^{\frac{1}{2}}\left[\Sigma_z^{\frac{1}{2}}\left[\Sigma_F - \Sigma_F C^\top\left[C\Sigma_F C^\top + \Sigma_\varepsilon\right]^{-1} C\Sigma_F\right]\Sigma_z^{\frac{1}{2}}\right]^{-\frac{1}{2}}\Sigma_z^{\frac{1}{2}}\tilde{\xi} + \tilde{z}\right)$$
$$= \bar{F} + \frac{1}{2}\tilde{\xi} + \frac{1}{2}\Sigma_z^{-\frac{1}{2}}\left[\Sigma_z^{\frac{1}{2}}\left[\Sigma_F - \Sigma_F C^\top\left[C\Sigma_F C^\top + \Sigma_\varepsilon\right]^{-1} C\Sigma_F\right]\Sigma_z^{\frac{1}{2}}\right]^{\frac{1}{2}}\Sigma_z^{-\frac{1}{2}}\left(\tilde{z} - \bar{z}\right),$$

which is used to compute the expected profit of the first stage.

Risk-averse insider

The second stage problem for the risk-averse insider is

$$\max_{\{x\}} E\left[W_0 + x^\top(\tilde{F} - p(\tilde{\omega}))|\tilde{Y}\right] - \frac{\rho}{2}V\left[W_0 + x^\top(\tilde{F} - p(\tilde{\omega}))|\tilde{Y}\right].$$

Due to linearity of the equilibrium it is known that $p\left(\tilde{\omega}\right) = \mathcal{A}_0 + \mathcal{A}_1\tilde{\omega}$, and the conditional expectation and variance can be substituted in and computed. The problem becomes

$$\max_{\{x\}} x^\top\left[E\left[\tilde{F}|\tilde{Y}\right] - \mathcal{A}_0 - \mathcal{A}_1\bar{z}\right] - x^\top\mathcal{A}_1 x - \frac{\rho}{2}\left[x^\top V\left[\tilde{F}|\tilde{Y}\right]x + x^\top\mathcal{A}_1^\top\Sigma_z\mathcal{A}_1 x\right].$$

And the first order conditions are

$$x: E\left[\tilde{F}|\tilde{Y}\right] - \mathcal{A}_0 - \mathcal{A}_1\bar{z} - 2\mathcal{A}_1 x - \rho\left[V\left[\tilde{F}|\tilde{Y}\right] + \mathcal{A}_1^\top\Sigma_z\mathcal{A}_1\right]x$$
$$= 0,$$

given the values for $E\left[\tilde{F}|\tilde{Y}\right]$ and $V\left[\tilde{F}|\tilde{Y}\right]$ from A1, and $\tilde{\xi} = E\left[\tilde{F}|\tilde{Y}\right] - \bar{F}$

$$x = \left[2\mathcal{A}_1 + \rho\left[\Sigma_F - \Sigma_F C^\top\left[C\Sigma_F C^\top + \Sigma_\varepsilon\right]^{-1} C\Sigma_F + \mathcal{A}_1^\top\Sigma_z\mathcal{A}_1\right]\right]^{-1}\tilde{\xi}$$
$$+\left[2\mathcal{A}_1 + \rho\left[\Sigma_F - \Sigma_F C^\top\left[C\Sigma_F C^\top + \Sigma_\varepsilon\right]^{-1} C\Sigma_F + \mathcal{A}_1^\top\Sigma_z\mathcal{A}_1\right]\right]^{-1}\left[\bar{F} - \mathcal{A}_0 - \mathcal{A}_1\bar{z}\right].$$

The coefficient matrices solve the following system:

$$\mathcal{B}_0 = \mathcal{B}_1\left[\bar{F} - \mathcal{A}_0 - \mathcal{A}_1\bar{z}\right],$$
$$\mathcal{B}_1 = \left[2\mathcal{A}_1 + \rho\left[\Sigma_F - \Sigma_F C^\top\left[C\Sigma_F C^\top + \Sigma_\varepsilon\right]^{-1} C\Sigma_F + \mathcal{A}_1'\Sigma_z\mathcal{A}_1\right]\right]^{-1},$$
$$\mathcal{A}_0 = \bar{F} - \mathcal{A}_1\left[\mathcal{B}_0 + \bar{z}\right],$$
$$\mathcal{A}_1 = \left[I - \Sigma_z\mathcal{B}_1^{-1}\left[\Sigma_F C^\top\left[C\Sigma_F C^\top + \Sigma_\varepsilon\right]^{-1} C\Sigma_F\right]^{-1}\right]^{-1}.$$

From the expressions for $\mathcal{A}_0$ and $\mathcal{B}_0$ together, $\mathcal{B}_0 = 0$ and $\mathcal{A}_0 = \bar{F} - \mathcal{A}_1 \bar{z}$.

Chapter 2

Bid-Ask spread in a rationally inattentive multi-security market

2.1 Introduction

The presence of insiders in a financial market offers a potential explanation for the spread between bid and ask prices. In this paper, I analyze the cross-market effect of the existence of an upper limit on insiders' capacity for information processing. I consider a model that features the endogenous choice of a noisy information structure for the investor. The choice of signal determines the investor's actions when placing a bid or ask order for each asset. When the signal chosen by the insider is correlated, investment decisions exhibit a structural link across markets. Insiders are the only investors with access to relevant information about asset payoffs, and are hence the ones who can supply true information to the market about them. Alternatively, there are also noise traders who place random liquidity orders in each market. Finally, market makers compete à la Bertrand by setting asset prices as they process the orders placed by investors. Market makers face an adverse selection problem with regard to insiders since they cannot identify the profiles of the investors that place each order. As in Glosten and Milgrom's (1985) model, in this paper I address the bid-ask spread as a purely informational phenomenon in all markets. Moreover, the correlated information that insiders observe induces correlated actions across markets, which generates a structural correlation in the formation of prices.

Markets operate as market makers set asset bid and ask prices to equilibrate the supply and the demand for assets. The spread arises in response to the potential losses to the market maker due to insiders. However, the specifics on how market makers trade can change from market to market, depending on pre-defined trading protocols. I construct a model that can be used to study two possible trading protocols. The first requires agents to submit the order for each asset to a different market maker. Therefore, market makers behave as specialists for each asset, as in the New York Stock Exchange (NYSE), and cannot see the orders that each investor places for other assets. The second protocol allows investors to place their orders for all assets with one market maker, whereby the market maker has additional information that can be used for screening purposes. From now on, I refer to market makers who trade only one type of asset as specialists and use market makers only for price-setting agents who trade portfolios.

In finance, any investor who is at an informational advantage beforehand is called an insider. However, the quality of the information available to the insider has a significant impact on the market outcome, but information can be costly to acquire and process and brokerage and financial firms spend a substantial amount of money on doing this. This model provides a tool to analyze the effects of having limited resources to process information.

For the analysis, I focus on the effect that limited access to affordable information has on insider choice, and how this choice affects price formation across markets. I set the asset payoffs to be ex-ante statistically independent to isolate the effect of the informational channel on bid and ask prices. Specifically, I show how the endogenous choice of an information structure can prevent the price vector from inheriting the statistical independence on the vector of asset payoffs—the model departs from a static bivariate version of the Glosten and Milgrom (1985) model. A risk-neutral specialist, or a market maker, sets the prices, and there is a group of insiders who observe an informative signal on the vector of asset payoffs.

The structure of the paper is the following. Section 2 introduces related literature related to market microstructure, information acquisition, and rational inattention. Section 3 describes how the market operates in this model. Later on, Section 4 presents the model where one specialist processes the orders for each asset. Since the market outcome, does not necessarily hold when the insiders face a capacity constraint for processing information, in Section 5, I present a model with a different trading protocol where one market maker processes the orders for both assets. This alternative trading protocol provides a sufficient condition to lose statistical independence in asset price formation, even when insiders are perfectly informed. Finally, Section 6 summarizes some concluding remarks.

2.2 Related literature

There has been some recent research regarding the endogenous choice of an information structure and its impact on financial markets. For instance, Mondria (2010) introduces the endogenous construction of a signal in a portfolio selection problem to explain price comovement. Specifically, he applies an information-processing constraint to Admati's (1985) portfolio selection problem, to characterize price correlation analytically. In his model, insiders are allowed to choose a noisy multidimensional signal as a linear combination of the true asset payoffs. This flexible structure allows for particular cases, such as producing independent signals for each of the assets, having one informative signal as a linear combination of all asset payoffs, or any linear combination between asset prices. He solves the model for two assets and shows how insiders prefer one informative signal over all other alternatives.

All previous studies that introduced the information-processing constraint to a portfolio selection problem impose ex-ante to have separate signals for each of the assets. In such problems, the constraint gives thresholds for allocating attention over some assets. Hence, the investor only chooses to observe signals on a limited number of assets. The first studies to apply this concept were Peng (2005) and Peng and Xiong (2006), calling it entropy learning. Later on, van Nieuwerburgh and Veldkamp (2009) and van Nieuwerburgh and Veldkamp (2010) use Admati's (1985) model and introduce the information-processing constraint by assigning one signal to each asset payoff. As a result, van Nieuwerburgh and Veldkamp (2010) find that this constraint leads to portfolio under-diversification. In a similar analysis, van Nieuwerburgh and Veldkamp (2009) explain the home bias puzzle. These studies impose ex-ante independence on the information; therefore, all prices remain statistically independent.

The optimal choice of information structure is a topic that has been revisited lately in the rational inattention literature. Rational inattention was initially introduced in Sims (1998) as a possible source of "stickiness" in markets, later to be formally defined in Sims (2003), featuring analytical tools developed for information theory.In models with rationally inattentive agents, the amount of information that an agent can process is finite and is called information-processing capacity. The agent faces a trade-off, allocating capacity to the precision of different signals to find the most profitable allocation; this is called attention allocation. The intuition of the attention allocation process is to set an upper limit on the uncertainty that an agent can resolve through signals. Other literature on rational inattention has adopted the perspective of choice theory, to which de Oliveira (2013) provides an axiomatic approach to it, while Matějka and McKay (2015) build an alternative explanation to the multinomial model following the results in Jung *et al.* (2015).[1]

In finance, investors profit from learning about the fundamental values of assets before trading—one who can do this before trading is called an insider. In this paper, I depart from the canonical models for insider trading, Kyle (1985) and Glosten and Milgrom (1985), where there are two types of investors, namely insiders and others who have no prior knowledge and submit random orders to the market, what are called noise traders. A third agent, a market maker, processes the orders and sets the prices after updating their beliefs. The market orders, placed by the insider, contain information about the asset payoffs. Noise trading prevents the market maker from perfectly filtering the information out of the order flow. The fundamental difference between the two models is the set of actions performed by the insider. Kyle's model allows a continuous domain for the market order, which makes the model suitable for analysis of the competition structure in the market. Analogously, in Glosten and Milgrom (1985), traders' are allowed to either buy or sell a single unit of the asset: the specialist randomly calls traders to submit their orders

[1] There is a different, more traditional, approach to understanding the concept of informativeness that was initially introduced as information chosen under uncertain conditions by Blackwell (1951, 1953), with sufficient conditions for ordering structures according to their informativeness without providing a quantitative measure for information. See de Oliveira (2018) for further references to Blackwell's theorem.

and sets the bid and ask prices. Both models would explain the difference between bid and ask prices, which the literature addresses as bid-ask spreads, considering this a purely informational phenomenon. Krishnan (1992) shows that there is equivalence between the results of Kyle (1985) and Glosten and Milgrom (1985). Admati (1985) proposes a multi-security version of Kyle's model for risk-averse insiders who select portfolios under perfect competition. Caballé and Krishnan (1994) develop a portfolio selection model for a risk-neutral insider under imperfect competition. A common factor to all insider trading models is that insiders are the only source of relevant information supplied to the market. Hence, the informational content of prices relies on insider information and the market maker's ability to filter it.

2.3 Information and market structure

2.3.1 Information

The measure of information flows is a topic that is mostly studied in information theory. Any analysis derived from information theory requires the identification of four elements: a source, a message, a channel, and a receiver. At one extreme, there is a source that generates a message (or signal). Then, as the message travels through a channel, it gets distorted by external facts. At the other extreme, there is a receiver who tries to recover as much information as possible after observing the distorted message. The flow of information depends on the channel's capacity. The standard measure to determine how distorted a message is on arrival at the receiver is Shannon entropy.

Shannon (1948) uses entropy as a measure of unpredictability (uncertainty, in our context).[2] Sims (2003) adapts Shannon entropy to quantify information flows in economics. Here, a message can be any relevant information, that is, an informative signal used to make decisions. The information structure would be the corresponding channel, and this channel's capacity determines the precision of the signal. The receiver then updates her

[2]This type of entropy is called Shannon entropy, which was generalized later by Rényi (1961). Not to be confused with the concept of thermodynamic entropy in statistical mechanics, although the intuition is somewhat similar. In general, probability distributions have an associated entropy function denoted by $H(X)$.

beliefs about the source about the source based on observation of the message. The information flow is the amount of uncertainty about the source that the message resolves for the receiver. This is called mutual information.

Sims (2003) and (2006) defines rational inattention as an agent's decision to optimally allocate zero precision to a particular signal. The theory implements the information-processing constraints proposed in Shannon (1948) to determine the allocation of "attention" to the signals that the agent wants to observe more precisely. The information-processing constraint is an upper bound to the mutual information between source and message.

Let $\tilde{x}_1, \tilde{x}_2$ be two random variables. The mutual information of a system is defined as the difference between the conditional and unconditional entropy function, that is

$$I\left(\tilde{x}_1; \tilde{x}_2\right) = H\left(\tilde{x}_1\right) - H\left(\tilde{x}_1 | \tilde{x}_2\right),$$

where I denotes the mutual information function, $H\left(\tilde{x}_i\right) = -\mathbb{E}\left[\ln\left(\Pr(\tilde{x}_i)\right)\right]$ is the entropy function and $H\left(\tilde{x}_1 | \tilde{x}_2\right) = -\mathbb{E}\left[\ln\left(\Pr(\tilde{x}_1)\right) | \tilde{x}_2\right]$ is the conditional entropy function. The mutual information measures the information about $\tilde{x}_1$ contained in $\tilde{x}_2$. An equivalent interpretation states that the mutual information gives the amount of uncertainty about $\tilde{x}_1$ that can be resolved as $\tilde{x}_2$ is observed. Suppose that two variables $\tilde{x}_1$ and $\tilde{x}_2$ share information but only $\tilde{x}_2$ is observable. The observable variable can deliver some information about the outcome of variable $\tilde{x}_1$. Moreover, variable $\tilde{x}_2$ is a signal for $\tilde{x}_1$ and the mutual information measures the informational content of this signal. A capacity constraint κ sets an upper bound for the uncertainty resolution generated by the signal. In most the cases, the capacity constraint binds.

2.3.2 Market structure

In this market, investors trade two assets whose payoffs, denoted by vector $\tilde{F} = \left[\tilde{F}_1, \tilde{F}_2\right]'$, are independently distributed. The asset payoffs represent the liquidation value of each asset after it is purchased. The random vector $\tilde{F}$ has support $\{0, 1\} \times \{0, 1\}$, which indicates that the possible outcomes for each of the payoffs are either 1, a high payoff, or 0, a low one. The probability of each asset having a high payoff is independent across assets. Let $\bar{F}_i$ be the probability of a high payoff on the i-th asset, then the distribution of each asset payoff is a Bernoulli, $\tilde{F}_i \sim B\left(\bar{F}_i\right)$. Thus, the expected payoff for the i-th asset is $\bar{F}_i$.

The set of investors is divided into two profiles: informed and noise traders. On the one hand, there are informed agents who can observe a multivariate signal $\tilde{Y}$ on the true payoff $\tilde{F}$ before they place their orders in the market, $\tilde{x}\left(\tilde{Y}\right)$ and which represent a fraction a of all traders. On the other hand, uninformed traders place random orders in the market $\tilde{z}$ and they are a $(1-a)$ share of the mass of investors. For simplicity and without loss of generality, I set noise traders to buy or sell a unit of an asset with equal probability. Traders are only allowed to buy/sell one unit of each asset once they are called to trade, i.e., the support of orders $\tilde{x}$ and $\tilde{z}$ is $\{-1.1\} \times \{-1, 1\}$.

The goal of the upcoming sections is to build the information structure of an informed investor endogenously. There is a critical difference between insiders, as partially informed investors, and noise traders. An insider, as part of his optimal information structure, is able to choose not to be informed about a specific asset. In contrast, noise traders can never become informed traders. Thus, an insider who optimally decides to be uninformed about one asset is still an insider, since it is his choice to be uninformed. Therefore, the two profiles form disjoint groups.

The trading protocol should be defined to characterize the market structure. In this paper, a trading protocol refers to the particular way traders are supposed to submit their orders. The protocol defines the price-setting agent's profile. I consider two possible trading protocols: the first supposes that a specialist processes the orders for each asset. A specialist randomly calls an investor to place an order for one asset, and another specialist does likewise for the second asset. The second trading protocol supposes a market maker is processing the orders for both assets. The difference between the two depends on the information that the price-setting trader observes, and is discussed in detail below. Both specialists and market makers set asset prices through Bertrand competition, which results in a zero expected profit rule. In this model, the specialists/market makers are ex-ante uninformed. They then update their beliefs from the information they extract from the submitted orders. A specialist/market maker randomly calls an investor to place an order, which means that all investors have the same probability of being called to trade. The selected trader belongs to the informed group with probability a or to the uninformed with probability $(1-a)$.Even if the specialist/market maker is aware of the existence of informed traders, the anonymity of order submission means investor profiles are not revealed. As a result, the specialist/market maker cannot practice price discrimination and sets the price according to what she observes, a bid or ask order; one price for all bid orders and another for all ask ones. The orders that a specialist on market i receives are represented by variable $\tilde{\omega}_i$, the order flow, for $i = 1, 2$. Such a variable is a market order for asset i, $\tilde{x}_i\left(\tilde{Y}\right)$ with probability a, and a liquidity trade $\tilde{z}_i$ with probability $(1-a)$, for $i = 1, 2$. Similarly, the orders that the market maker receives are represented by variable $\tilde{\omega}$, the order flow. Such a variable is a market order $\tilde{x}\left(\tilde{Y}\right)$ with probability a and a liquidity trade $\tilde{z}$ with probability $(1-a)$. I summarize the behavior of the order flow by means of the following mixture notation: $\tilde{\omega} = a \circ \tilde{x}\left(\tilde{Y}\right) \oplus (1-a) \circ \tilde{z}$.[3] The order flow is the only source of information that the specialist or the market maker can use to update her beliefs. Hence, the zero expected profit pricing condition becomes $p_i\left(\tilde{\omega}_i\right) = \mathbb{E}\left[\tilde{F}_i|\tilde{\omega}_i\right]$, $i = 1, 2$, for specialists, and $p\left(\tilde{\omega}\right) = \mathbb{E}\left[\tilde{F}|\tilde{\omega}\right]$ for market makers. Since the pricing rule set by the specialist/market maker is a conditional expectation of the payoffs, the price vector indirectly contains the insiders' information structure. If there are no insiders to trade in the market, the order flow is orthogonal to the asset payoffs. Therefore, both specialists and market makers are unable to update their beliefs and the prices coincide with the unconditional expected payoff.

One may wrongly assume that trading to a specialist is equivalent to trading to a market maker, but specialists and market makers have different information sets. Recall that insiders are either informed or noise traders for both assets. Therefore, a specialist trading one asset does not know whether the same investor is called to place an order for the other asset. This protects the identity of investors across markets, as discussed in section 4. In contrast, when market makers are trading, their information sets contain information about the vector of orders. They know that an insider is an insider in both markets; similarly, if they call a noise trader, the investor is noise trading in both markets. The order flow comes from the same investor profile in both markets. Section 5 solves the

[3]Given three random variables $\tilde{w}_1$, $\tilde{w}_2$ and $\tilde{y}$, the notation $\tilde{y} = c \circ \tilde{w}_1 \oplus (1-c) \circ \tilde{w}_2$ indicates that the variable $\tilde{y}$ takes the value of variable $\tilde{w}_1$ with probability c and $\tilde{w}_2$ with probability $(1-c)$.

model when a market maker trades.

2.4 A specialist for each market

This section analyzes the effects of insider behavior on prices when there is a specialist in each market who processes the orders. The first part introduces a benchmark model to describe insider behavior when they have access to the true asset payoffs before submitting their market orders. The second part explains how a limited information-processing capacity can generate correlation in the formation of bid and ask prices, as well as the effect on the informational content of prices.

2.4.1 Perfect information

Consider the case where there is a a specialist for each market and informed traders that have perfect information about the payoff of each asset. This market is equivalent to two separate Glosten and Milgrom (1985) models for each asset. The specialist on each market cannot infer anything about the order placed for the other asset. Since payoffs are independent, insiders have perfect information, and the specialist only observes the order flow for one asset, so trading actions are independent in both markets.

The informed trader submits an order to buy a unit of asset i, $x_i = 1$, whenever he knows that the payoff is high and to sell a unit, $x_i = -1$, when the payoff is low, for $i = 1, 2$. Given a generic asset i, the ask and bid prices for this asset would be given by

$$\begin{aligned}\tilde{p}_i(1) &= \mathbb{E}\left[\tilde{F}_i|\omega_i = 1\right] = \Pr(F_i = 1|\omega_i = 1)\\ &= \frac{\Pr(F_i = 1)}{\Pr(\omega_i = 1)}\Pr(\omega_i = 1|F_i = 1) = \frac{\bar{F}_i\left(a + \frac{1-a}{2}\right)}{a\bar{F}_i + \frac{(1-a)}{2}}\\ &= \frac{\bar{F}_i(a+1)}{2a\bar{F}_i + (1-a)},\end{aligned}$$

and

$$\begin{aligned}\tilde{p}_i(-1) &= \mathbb{E}\left[\tilde{F}_i|\omega_i=-1\right] = \Pr\left(F_i=1|\omega_i=-1\right) \\ &= \frac{\Pr\left(F_i=1\right)}{\Pr\left(\omega_i=-1\right)}\Pr\left(\omega_i=-1|F_i=1\right) = \frac{\frac{1-a}{2}\bar{F}_i}{a\left(1-\bar{F}_i\right)+\frac{(1-a)}{2}} \\ &= \frac{(1-a)\bar{F}_i}{2a\left(1-\bar{F}_i\right)+(1-a)}.\end{aligned}$$

The pricing functions show how different is the price from the true value of the asset payoff. If the asset gives a high payoff, an insider buys a unit of the asset, yet, the price is lower than 1. Otherwise, if it gives a low payoff, the price is greater than 0. The difference between the price and the payoff is the profit made by the insider. Note that the pricing function follows the conditional expectation of the specialist, who faces higher uncertainty than the insiders, due to never knowing if the observed order corresponds to an insider or a noise trader. Thus, the existence of noise traders prevents specialists from perfectly filtering the true asset payoffs from the order flow. If there are only insiders, specialists can perfectly discriminate prices, $p(1)=1$ and $p(-1)=0$, and insiders cannot profit from their informational advantage. Alternatively, if there are only noise traders, no useful information can be filtered from the order flow, and the market maker sets a single price $p(1)=p(-1)=\bar{F}_i$. The resulting asset price is the probability of the asset delivering a high payoff.

2.4.2 Imperfectly informed insiders

Consider now the case where insiders do not know the exact outcome of the payoff. Instead, they can observe a signal that contains information about it. Hence, insiders are constrained to placing their orders based on noisy signals $\tilde{Y}_s$, $s=1,2$, at most one per asset.[4] Note that I do not constrain the signals to be ex-ante independent, or assume that each signal can only contain information about one asset. As a result, the signal I construct can be used to endogenously address ex-ante asset payoff regardless of the insider's perspective. The noisy signal takes the form of a Bernoulli random variable that

[4]Recall that for any random vector $\tilde{F}$ of dimension n there can be at most n linearly independent informative signals.

delivers external noise with probability δ_s, that with probability $(1-\delta_s)\lambda_s$ delivers the payoff of asset 1, and with probability $(1-\delta_s)(1-\lambda_s)$ the payoff of asset 2. The behavior of the signal is summarized by $\tilde{Y}_s = (1-\delta_s) \circ \left(\lambda_s \circ \tilde{F}_1 \oplus (1-\lambda_s) \circ \tilde{F}_2\right) \oplus \delta_s \circ \tilde{\varepsilon}_s$, where $\tilde{\varepsilon}_s = \frac{1}{2} \circ 0 \oplus \frac{1}{2} \circ 1$ represents the external noise as an exogenous Bernoulli variable, with probability $\frac{1}{2}$, for $s = 1, 2$.[5] The parameters λ_s can be viewed as the intensity of $\tilde{F}_1$ on the signal $\tilde{Y}_s$, while the parameter δ_s can be viewed as the intensity of the noise, or equivalently, the general garbling parameter. Its complement $(1-\delta_s)$ is proportional to the precision of the signal.[6] Since the asset payoffs are statistically independent, when the variable $\tilde{Y}_s$, $s = 1, 2$ serves as a signal for $\tilde{F}_1$, the signal delivers the desired outcome with probability $(1-\delta_s)\lambda_s$. So is the case when the variable $\tilde{Y}_s$ serves as signal for $\tilde{F}_2$. Let $\tilde{Y} = \left[\tilde{Y}_1, \tilde{Y}_2\right]'$ be the vector of signals. Finally, the probability of external noise is assumed to be symmetric for all signals in $\tilde{Y}$, $\delta_1 = \delta_2 = \delta$. Figure 2.1 graphically represents of the signal on a tree diagram.

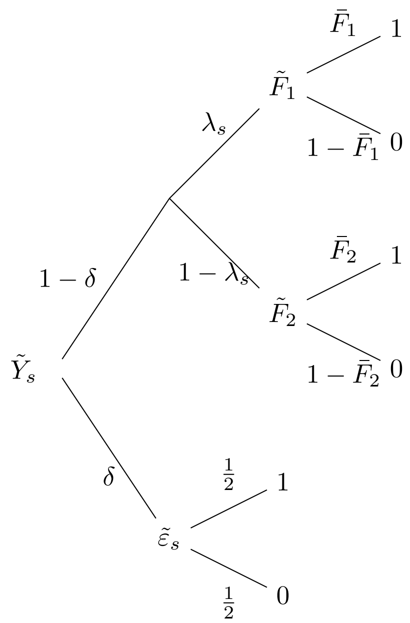

Figure 2.1: Tree diagram for the signal $\tilde{Y}_s = (1-\delta) \circ \left(\lambda_s \circ \tilde{F}_1 \oplus (1-\lambda_s) \circ \tilde{F}_2\right) \oplus \delta \circ \tilde{\varepsilon}_s$.

The signal $\tilde{Y}_s$, as depicted in Figure 2.1, has terminal nodes, or final outcomes, with

[5]By repeatedly applying lottery decomposition property, any binary random variable $\tilde{l}$ with support $\{l_1, l_2\}$ can be represented as $\tilde{l} = q_1 \circ l_1 \oplus (1-q_1) \circ l_2$, where $q_1 = \Pr[l = l_1]$.

[6]I use garbling as proposed by Marschak and Miyasawa (1968) to interpret Blackwell informativeness, i.e., a mean preserving spread of an informative signal. Therefore, the garbling parameter makes the signal less informative.

two possible values, 0 and 1. Nature defines the paths to the terminal nodes given the probabilities δ, λ_s, $\bar{F}_1$, and $\bar{F}_2$. For illustrative purposes, suppose that Nature moves sequentially to define the paths to nodes $\tilde{F}_1$, $\tilde{F}_2$, and $\tilde{\varepsilon}_s$. Initially, Nature chooses the node $\tilde{\varepsilon}_s$ with probability δ, otherwise she plays again. If the node $\tilde{\varepsilon}_s$ is not chosen, Nature delivers node $\tilde{F}_1$ with probability λ_s, and node $\tilde{F}_2$ otherwise. After she chooses a node $\tilde{F}_1$, $\tilde{F}_2$ or $\tilde{\varepsilon}_s$, the probabilities of an outcome of "1" are $\bar{F}_1$, $\bar{F}_2$ or $\frac{1}{2}$ respectively. Hence, the unconditional probability of a high outcome of the signal is $\Pr(Y_s = 1) = (1-\delta)\left(\lambda_s \bar{F}_1 + (1-\lambda_s)\bar{F}_2\right) + \frac{1}{2}\delta$. Therefore, given $F_1 = 1$, the probability that the signal delivers a high outcome is $\Pr(Y_s = 1 | F_1 = 1) = (1-\delta)\left(\lambda_s + (1-\lambda_s)\bar{F}_2\right) + \frac{1}{2}\delta$, which is higher than the unconditional probability. The intuition behind the difference is that, should Nature choose node $\tilde{F}_1$, the outcome is "1" with probability 1, as opposed to the unconditional case where the probability is $\bar{F}_1$. The same argument applies to all other conditional probabilities. Table 2.1 summarizes the conditional distribution of signal $\tilde{Y}_s$.

			$\tilde{Y}_s$	
			1	0
$\tilde{F}$	$\tilde{F}_1$	1	$(1-\delta)\left(\lambda_s + (1-\lambda_s)\bar{F}_2\right) + \frac{1}{2}\delta$	$(1-\delta)(1-\lambda_s)\left(1-\bar{F}_2\right) + \frac{1}{2}\delta$
		0	$(1-\delta)(1-\lambda_s)\bar{F}_2 + \frac{1}{2}\delta$	$(1-\delta)\left(\lambda_s + (1-\lambda_s)\left(1-\bar{F}_2\right)\right) + \frac{1}{2}\delta$
	$\tilde{F}_2$	1	$(1-\delta)\left(\lambda_s \bar{F}_1 + (1-\lambda_s)\right) + \frac{1}{2}\delta$	$(1-\delta)\lambda_s\left(1-\bar{F}_1\right) + \frac{1}{2}\delta$
		0	$(1-\delta)\lambda_s \bar{F}_1 + \frac{1}{2}\delta$	$(1-\delta)\left(\lambda_s\left(1-\bar{F}_1\right) + (1-\lambda_s)\right) + \frac{1}{2}\delta$

Table 2.1: Conditional distribution of $\tilde{Y}_s$ given $\tilde{F}$.

Market orders

Let $X = \{-1, 1\} \times \{-1, 1\}$ be the support of the market orders, $x \in X$. Given the signal vector $\tilde{Y}$, an insider must choose the market orders $x\left(\tilde{Y}\right) \in X$ to maximize the (conditional) expected profits, i.e.,

$$x\left(\tilde{Y}\right) = \arg\max_{x \in X} \mathbb{E}\left[\sum_{i=1}^{2} x_i \left(\tilde{F}_i - p_i(\omega_i)\right) \middle| \tilde{Y}\right]. \tag{2.1}$$

Note that in (2.1), the term $p_i(x_i)$ corresponds to the bid/ask price for asset i depending on the agent's choice. Each price is a conditional expectation that corresponds to a garbled transformation of the signal. Hence, it is always closer to the unconditional mean $\bar{F}_i$ than the insider's conditional expectation $\mathbb{E}\left[\tilde{F}_i|\tilde{Y}\right]$. From now on, let signal $\tilde{Y}_s$, where $\lambda_s \geq \lambda_{s'}$, $s, s' = 1, 2$ and $s \neq s'$, be $\tilde{Y}_1$, since it is the signal that contains the most information about $\tilde{F}_1$. Analogously, $\tilde{Y}_2$ is the one that contains most information about $\tilde{F}_2$. Thus, the market orders are such that the most informative signal determines the action, i.e., the signal that delivers the most information about each state determines the insider's bid/ask choice. The resulting market orders are

$$x_1\left(\tilde{Y}\right) = \begin{cases} 1 & \text{if } Y_1 = 1 \\ -1 & \text{if } Y_1 = 0, \end{cases}$$

and

$$x_2\left(\tilde{Y}\right) = \begin{cases} 1 & \text{if } Y_2 = 1 \\ -1 & \text{if } Y_2 = 0. \end{cases}$$

Proof that these values are the optimal choices for the insider follows by contradiction. Suppose that the insider observes signal $\tilde{Y}_2$ before he selects the market order for the first asset. The insider's corresponding informational advantage is smaller because the probability of node $\tilde{F}_1$ is higher in signal $\tilde{Y}_1$ than in signal $\tilde{Y}_2$.[7] Then, suppose that the insider had chosen $x_i = -1$ when $Y_i = 1$, his expected profits increase automatically if he deviated to $x_i = 1$, since $\mathbb{E}\left[\tilde{F}_i - p_i(1)|Y_i = 1\right] > \mathbb{E}\left[p_i(-1) - \tilde{F}_i|Y_i = 1\right]$, where the second term is negative.

The insider's market order behavior is a response to increasing expected profits in the informational advantage. The following definition characterizes the meaning of the informativeness of signals in this particular model.

Definition. A signal $\tilde{Y}_s$, for $s = 1, 2$, is said to be informative if, conditionally on observing it, the probability of observing the implied outcome increases.

The previous definition for informative signals implies the following conditions:

$$\Pr\left(F_1 = 1|Y_1 = 1\right) \geq \Pr\left(F_1 = 1|Y_2 = 1\right)\left(> \bar{F}_1\right),$$

[7] The informational advantage is the gap between insider's expectation over the specialist's expectation.

$$\Pr\left(F_2 = 1|Y_2 = 1\right) \geq \Pr\left(F_2 = 1|Y_1 = 1\right) \left(> \bar{F}_2\right),$$

and, analogously

$$\Pr\left(F_1 = 0|Y_1 = 0\right) \geq \Pr\left(F_1 = 0|Y_2 = 0\right) \left(> 1 - \bar{F}_1\right),$$

$$\Pr\left(F_2 = 0|Y_2 = 0\right) \geq \Pr\left(F_2 = 0|Y_1 = 0\right) \left(> 1 - \bar{F}_2\right).$$

In general, let $\bar{F}_{i|s,1} := \Pr\left(F_i = 1|Y_s = 1\right)$, $\left(1 - \bar{F}_{i|s,0}\right) := \left(1 - \Pr\left(F_i = 1|Y = 0\right)\right) \equiv \Pr\left(F_i = 0|Y_s = 0\right)$ for $i, s = 1, 2$.

Given that the garbling parameter that weights the external noise is symmetric for the two signals, the only difference lies in the distribution of weights inside the informative part of each signal. Moreover, the behavior induced by each signal in the market order has an additional implication. Each signal has two sources of noise; the first one is represented by garbling and is common to both signals; the second one is the irrelevant information for each choice. For example, in signal $\tilde{Y}_1$, the informational content about $\tilde{F}_2$ is as good as noise in the choice of x_1. Similarly, any information that $\tilde{Y}_2$ contains about $\tilde{F}_1$ can be considered noise. The weights $(1 - \lambda_1)$ and λ_2 are the redundant information parameters. Therefore, the optimal weight distribution plays an important role in the precision analysis. Proposition 1 determines the corresponding values for the weights, such that the signals are informative and they are consistent with the assigned name.

Proposition 4. *Let $\tilde{Y}_1$ and $\tilde{Y}_2$ be two informative signals for the asset payoffs $\tilde{F}_1$ and $\tilde{F}_2$. Signal $\tilde{Y}_1$ is at least as informative as the signal $\tilde{Y}_2$ for $\tilde{F}_1$, and the signal $\tilde{Y}_2$ is at least as informative as the signal $\tilde{Y}_1$ for $\tilde{F}_2$ if and only if $\lambda_2 \leq \lambda_1$.*

Proof. Following the informativeness conditions

$$\frac{(1-\delta)\left(\lambda_1 + (1-\lambda_1)\bar{F}_2\right) + \frac{1}{2}\delta}{(1-\delta)\left(\lambda_1\bar{F}_1 + (1-\lambda_1)\bar{F}_2\right) + \frac{1}{2}\delta} \geq \frac{(1-\delta)\left(\lambda_2 + (1-\lambda_2)\bar{F}_2\right) + \frac{1}{2}\delta}{(1-\delta)\left(\lambda_2\bar{F}_1 + (1-\lambda_2)\bar{F}_2\right) + \frac{1}{2}\delta}, \tag{2.2}$$

$$\frac{(1-\delta)\left(\lambda_2\bar{F}_1 + (1-\lambda_2)\right) + \frac{1}{2}\delta}{(1-\delta)\left(\lambda_2\bar{F}_1 + (1-\lambda_2)\bar{F}_2\right) + \frac{1}{2}\delta} \geq \frac{(1-\delta)\left(\lambda_1\bar{F}_1 + (1-\lambda_1)\right) + \frac{1}{2}\delta}{(1-\delta)\left(\lambda_1\bar{F}_1 + (1-\lambda_1)\bar{F}_2\right) + \frac{1}{2}\delta}, \tag{2.3}$$

$$\begin{aligned} &\frac{(1-\delta)\left(\lambda_1 + (1-\lambda_1)\left(1-\bar{F}_2\right)\right) + \frac{1}{2}\delta}{(1-\delta)\left(\lambda_1\left(1-\bar{F}_1\right) + (1-\lambda_1)\left(1-\bar{F}_2\right)\right) + \frac{1}{2}\delta} \\ &\quad \geq \frac{(1-\delta)\left(\lambda_2 + (1-\lambda_2)\left(1-\bar{F}_2\right)\right) + \frac{1}{2}\delta}{(1-\delta)\left(\lambda_2\left(1-\bar{F}_1\right) + (1-\lambda_2)\left(1-\bar{F}_2\right)\right) + \frac{1}{2}\delta}, \text{ and} \end{aligned} \tag{2.4}$$

$$\frac{(1-\delta)\left(\lambda_2\left(1-\bar{F}_1\right)+(1-\lambda_2)\right)+\frac{1}{2}\delta}{(1-\delta)\left(\lambda_2\left(1-\bar{F}_1\right)+(1-\lambda_2)\left(1-\bar{F}_2\right)\right)+\frac{1}{2}\delta} \tag{2.5}$$
$$\geq \frac{(1-\delta)\left(\lambda_1\left(1-\bar{F}_1\right)+(1-\lambda_1)\right)+\frac{1}{2}\delta}{(1-\delta)\left(\lambda_1\left(1-\bar{F}_1\right)+(1-\lambda_1)\left(1-\bar{F}_2\right)\right)+\frac{1}{2}\delta}.$$

Inequalities (2.2) and (2.4) hold for $0 \leq \delta \leq 1$, $0 \leq \lambda_1 \leq 1$, $0 \leq \lambda_2 \leq \lambda_1$. Analogously, inequalities (2.3) and (2.5) imply that for any value $0 \leq \delta \leq 1$, then $0 \leq \lambda_1 \leq 1$ and $\lambda_2 \leq \lambda_1 \leq 1$. □

Proposition 1 states that the most informative signals are those that increase the insider's predicting power regarding the true asset payoffs and that there is no signal that can be strictly more informative for both asset payoffs. This informativeness condition is also sufficient to prevent moral hazard problems arising when the insider chooses the signal.

Pricing

A specialist prices separately in each market, according to her zero expected profit rule $p_i(\tilde{\omega}_i) = \mathbb{E}\left[\tilde{F}_i|\tilde{\omega}_i\right]$. Thus, when she observes the order flow, the specialist does not know whether she observes the market order or noise trading. Due to Bertrand competition, the pricing rule becomes a learning rule; it is the updated belief given the order flow. Besides, the probability distribution of market orders is given by the distribution of the signals. There is one-to-one mapping between signals and market orders, which means that the market orders inherit the stochastic properties of the signal. Therefore, the specialist anticipates this relationship in the order flow and performs Bayesian updating to obtain the new distribution. Note that $p_i(1) = \mathbb{E}\left[\tilde{F}_i|\omega_i = 1\right] = \Pr\left[F_i = 1|\omega_i = 1\right]$ and $p_i(-1) = \mathbb{E}\left[\tilde{F}_i|\omega_i = -1\right] = \Pr\left[F_i = 1|\omega_i = -1\right]$ because all variables follow a Bernoulli distribution. Then, the corresponding bid and ask prices are

$$\begin{aligned}
p_1(1) &= \Pr[F_1 = 1|\omega_1 = 1] \\
&= \frac{\Pr[\omega_1 = 1|F_1 = 1]\Pr[F_1 = 1]}{\Pr[\omega_1 = 1]} \\
&= \frac{a\Pr[Y_1 = 1|F_1 = 1] + \frac{1-a}{2}}{a\Pr[Y_1 = 1] + \frac{1-a}{2}}\bar{F}_1 \\
&= \frac{a(1-\delta)\left(\lambda_1 + (1-\lambda_1)\bar{F}_2\right) + \frac{1}{2}(a\delta + 1 - a)}{a(1-\delta)\left(\lambda_1\bar{F}_1 + (1-\lambda_1)\bar{F}_2\right) + \frac{1}{2}(a\delta + 1 - a)}\bar{F}_1,
\end{aligned}$$

$$\begin{aligned}
p_1(-1) &= \frac{\Pr[\omega_1 = -1|F_1 = 1]\Pr[F_1 = 1]}{\Pr[\omega_1 = -1]} \\
&= \frac{a\Pr[Y_1 = 0|F_1 = 1] + \frac{1-a}{2}}{a\Pr[Y_1 = 0] + \frac{1-a}{2}}\bar{F}_1 \\
&= \frac{a(1-\delta)(1-\lambda_1)\left(1-\bar{F}_2\right) + \frac{1}{2}(a\delta + 1 - a)}{a(1-\delta)\left(\lambda_1\left(1-\bar{F}_1\right) + (1-\lambda_1)\left(1-\bar{F}_2\right)\right) + \frac{1}{2}(a\delta + 1 - a)}\bar{F}_1,
\end{aligned}$$

$$\begin{aligned}
p_2(1) &= \frac{a(1-\delta)\left(\lambda_2\bar{F}_1 + (1-\lambda_2)\right) + \frac{1}{2}(a\delta + 1 - a)}{a(1-\delta)\left(\lambda_2\bar{F}_1 + (1-\lambda_2)\bar{F}_2\right) + \frac{1}{2}(a\delta + 1 - a)}\bar{F}_2, \text{ and} \\
p_2(-1) &= \frac{a(1-\delta)\lambda_2\left(1-\bar{F}_1\right) + \frac{1}{2}(a\delta + 1 - a)}{a(1-\delta)\left(\lambda_2\left(1-\bar{F}_1\right) + (1-\lambda_2)\left(1-\bar{F}_2\right)\right) + \frac{1}{2}(a\delta + 1 - a)}\bar{F}_2.
\end{aligned}$$

These prices reflect the new source of noise introduced to the market by noise traders. Thus, if the insiders are able to perfectly filter the signals, i.e., $\delta = 0$, $\lambda_1 = 1$ and $\lambda_2 = 0$, the resulting prices converge to those obtained under perfect information in the previous subsection. Moreover, if the specialist can identify the insiders, the prices would reflect the insider information, and the informational advantage disappears. Nevertheless, the prices cannot reveal the true asset payoffs if the insiders are unable to eliminate all the external noise from the signal.

Signal construction

An insider was defined as an investor who is allowed to choose the information before the trading decision. Since the information should be available to the investor at the time of the trade, information is chosen during a stage before trading takes place. In this stage, only informed investors are involved in optimally designing the signal, provided that information acquisition is costly. Consider the informational content of the signal $\tilde{Y}$ about $\tilde{F}$, i.e., $I\left(\tilde{F};\tilde{Y}\right) = H\left(\tilde{F}\right) - H\left(\tilde{F}|\tilde{Y}\right)$, where $H\left(\tilde{F}\right) = -\mathbb{E}\left[\ln\left(\Pr(\tilde{F})\right)\right]$, and $H\left(\tilde{F}|\tilde{Y}\right) = -\mathbb{E}\left[\ln\left(\Pr(\tilde{F})\right)|\tilde{Y}\right]$. The two signals take the form $\tilde{Y}_s = (1-\delta) \circ \left(\lambda_s \circ \tilde{F}_1 \oplus (1-\lambda_s) \circ \tilde{F}_2\right) \oplus \delta \circ \tilde{\varepsilon}_s$, for $s = 1, 2$, where a key feature is that $\tilde{\varepsilon}_1$ and $\tilde{\varepsilon}_2$ are independent so there cannot be an exogenous source of comovement introduced ex-ante. I determine the conditions, if any, where the signals deliver redundant information.

Given the signal, the insider's conditional expectations of the asset payoffs are

$$\bar{F}_{1|1,1} = \frac{(1-\delta)\left(\lambda_1 + (1-\lambda_1)\bar{F}_2\right) + \frac{1}{2}\delta}{(1-\delta)\left(\lambda_1\bar{F}_1 + (1-\lambda_1)\bar{F}_2\right) + \frac{1}{2}\delta}\bar{F}_1,$$

$$\bar{F}_{2|2,1} = \frac{(1-\delta)\left(\lambda_2\bar{F}_1 + (1-\lambda_2)\right) + \frac{1}{2}\delta}{(1-\delta)\left(\lambda_2\bar{F}_1 + (1-\lambda_2)\bar{F}_2\right) + \frac{1}{2}\delta}\bar{F}_2,$$

$$\left(1-\bar{F}_{1|1,0}\right) = \frac{(1-\delta)\left(\lambda_1 + (1-\lambda_1)\left(1-\bar{F}_2\right)\right) + \frac{1}{2}\delta}{(1-\delta)\left(\lambda_1\left(1-\bar{F}_1\right) + (1-\lambda_1)\left(1-\bar{F}_2\right)\right) + \frac{1}{2}\delta}\left(1-\bar{F}_1\right),$$

$$\left(1-\bar{F}_{2|2,0}\right) = \frac{(1-\delta)\left(\lambda_2\left(1-\bar{F}_1\right) + (1-\lambda_2)\right) + \frac{1}{2}\delta}{(1-\delta)\left(\lambda_2\left(1-\bar{F}_1\right) + (1-\lambda_2)\left(1-\bar{F}_2\right)\right) + \frac{1}{2}\delta}\left(1-\bar{F}_2\right).$$

The reduction of garbling comes at a high cost, so the second set of conditions to be imposed on the signal must guarantee affordability for the investor. The main goal is to reduce the noise as much as possible, to make the conditional probabilities shown above as large as possible, provided the investor's capacity (budget) to reduce uncertainty. The insider aims to acquire a perfectly informative signal, that is $\delta = 0$, $\lambda_1 = 1$ and $\lambda_2 = 0$. However, the cost of such a signal is too high to compensate for each investor's expected profits. In fact, information theory predicts that the resolution of uncertainty comes at an infinite cost when variables are continuous.[8]

To do so, let the cost of each signal be a strictly increasing linear function $g : \mathbb{R}_+ \to \mathbb{R}_+$ $g'(\cdot) > 0, g''(\cdot) = 0$ of the information it provides, i.e., $c_1(\delta, \lambda_1) = g\left(I\left(\tilde{F};\tilde{Y}_1\right)\right)$, and

[8]The cost of the acquisition of perfect information about continuously distributed asset payoffs can be infinite. Take the case when payoffs follow a Gaussian distribution; perfect information is equivalent to the reduction of the signal's variance to zero. The information measure is inversely proportional to the probability of an event. This relation holds by construction, see Wiederholt (2010) for the definition under normality. Therefore, the resolution of uncertainty in a normal random variable, i.e., the reduction of the variance to zero, requires an infinitely large amount of information. Moreover, any nondecreasing cost function for information generates an infinite cost for the uncertainty resolution of the asset payoffs.

$c_2(\delta,\lambda_2) = g\left(I\left(\tilde{F};\tilde{Y}_2\right)\right)$. Proposition 4 implies $I\left(\tilde{F}_i;\tilde{Y}_i\right) \geq I\left(\tilde{F}_i;\tilde{Y}_j\right)$, $i,j = 1,2$ and $i \neq j$. Moreover, the mutual information function $I\left(\tilde{F}_1;\tilde{Y}_1\right)$ is increasing in probability λ_1 and $I\left(\tilde{F}_2;\tilde{Y}_2\right)$ is decreasing in λ_2 and both are decreasing in δ. This result follows from the informativeness condition stated above.[9] The mutual information between each signal and the corresponding asset payoff given by

$$
\begin{aligned}
I\left(\tilde{F}_1;\tilde{Y}_1\right) =& H\left(\tilde{F}_1\right) - H\left(\tilde{F}_1|\tilde{Y}_1\right) \\
=& -\mathbb{E}\left[\log\left(\Pr\left(\tilde{F}_1\right)\right)\right] - \sum_{i=0}^{1} \Pr\left(Y_1 = i\right) H\left(\tilde{F}_1|Y_1 = i\right)
\end{aligned}
$$

$$
\begin{aligned}
&= \left((1-\delta)\left(\lambda_1 + (1-\lambda_1)\bar{F}_2\right) + \frac{1}{2}\delta\right)\log\left(\frac{(1-\delta)\left(\lambda_1 + (1-\lambda_1)\bar{F}_2\right) + \frac{1}{2}\delta}{(1-\delta)\left(\lambda_1\bar{F}_1 + (1-\lambda_1)\bar{F}_2\right) + \frac{1}{2}\delta}\right) \\
&+ \left((1-\delta)(1-\lambda_1)\left(1-\bar{F}_2\right) + \frac{1}{2}\delta\right)\log\left(\frac{(1-\delta)(1-\lambda_1)\left(1-\bar{F}_2\right) + \frac{1}{2}\delta}{(1-\delta)\left(\lambda_1\left(1-\bar{F}_1\right) + (1-\lambda_1)\left(1-\bar{F}_2\right)\right) + \frac{1}{2}\delta}\right) \\
&-\bar{F}_1\log\left(\bar{F}_1\right)
\end{aligned}
$$

and

$$
\begin{aligned}
I\left(\tilde{F}_2;\tilde{Y}_2\right) =& \left((1-\delta)\left(\lambda_2\bar{F}_1 + (1-\lambda_2)\right) + \frac{1}{2}\delta\right)\log\left(\frac{(1-\delta)\left(\lambda_2\bar{F}_1 + (1-\lambda_2)\right) + \frac{1}{2}\delta}{(1-\delta)\left(\lambda_2\bar{F}_1 + (1-\lambda_2)\bar{F}_2\right) + \frac{1}{2}\delta}\right) \\
&+ \left((1-\delta)\lambda_2\left(1-\bar{F}_1\right) + \frac{1}{2}\delta\right)\log\left(\frac{(1-\delta)\lambda_2\left(1-\bar{F}_1\right) + \frac{1}{2}\delta}{(1-\delta)\left(\lambda_2\left(1-\bar{F}_1\right) + (1-\lambda_2)\left(1-\bar{F}_2\right)\right) + \frac{1}{2}\delta}\right) \\
&- \bar{F}_2\log\left(\bar{F}_2\right).
\end{aligned}
$$

See appendix A.

Both mutual information functions are increasing in the conditional probabilities of an accurate prediction of the outcome. Figure 2.2 shows the behavior of the mutual information in the case of signal $\tilde{Y}_1$. The plot depicts how the capacity requirement increases dramatically when either $\delta \to 0$ or $\lambda_1 \to 1$. The black lines over the surface are level curves that illustrate the rapid increase in the capacity as less garbling is introduced.

[9] Simply take the derivatives of the conditional probabilities to obtain $\frac{\partial\bar{F}_{1|1,1}}{\partial\delta} < 0$, $\frac{\partial\bar{F}_{2|2,1}}{\partial\delta} < 0$, $\frac{\partial\left(1-\bar{F}_{1|1,0}\right)}{\partial\delta} < 0$ and $\frac{\partial\left(1-\bar{F}_{2|2,0}\right)}{\partial\delta} < 0$, and $\frac{\partial\bar{F}_{1|1,1}}{\partial\lambda_s}$, $\frac{\partial\bar{F}_{2|2,1}}{\partial\lambda_s} > 0$,$\frac{\partial\left(1-\bar{F}_{1|1,0}\right)}{\partial\lambda_{-s}}$ and $\frac{\partial\left(1-\bar{F}_{2|2,0}\right)}{\partial\lambda_{-s}} < 0$.

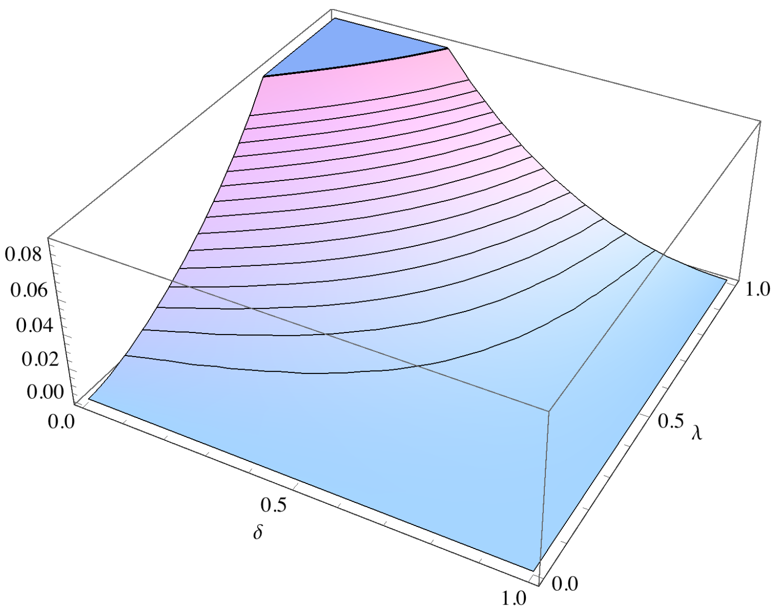

Figure 2.2: $I\left(\tilde{F}_1;\tilde{Y}_1\right)$ for different parameter values of δ and λ_1. *Note: The surface was plotted for* $\bar{F}_1 = 0.7$, $\bar{F}_2 = 0.5$, *the shape of the surface does not exhibit significant changes as the value of the parameters is changed.*

Since the marginal cost of information is assumed to be constant, the total cost is $c(\delta,\lambda_1,\lambda_2) = g\left(I\left(\tilde{F};\tilde{Y}_1\right) + I\left(\tilde{F};\tilde{Y}_2\right)\right)$. However, each signal delivers information about both asset payoffs, and the insider purchases some amount of duplicated information. Recall that $I\left(\tilde{F};\tilde{Y}_i\right) = I\left(\tilde{F}_i;\tilde{Y}_i\right) + I\left(\tilde{F}_j;\tilde{Y}_i\right)$ for $i,j = 1,2$ and $i \neq j$, where $I\left(\tilde{F}_j;\tilde{Y}_i\right)$ $i,j = 1,2$ and $i \neq j$ is information that has been purchased and will not be used. Furthermore, Proposition 7 rules out any possibility of moral hazard while the investor purchases the signal.

Definition. The *iso-cost* function is the function $\delta^c(\lambda_1,\lambda_2,\bar{\kappa})$ that solves the relation $g(\bar{\kappa}) = c(\delta^c,\lambda_1,\lambda_2)$ for δ^c, where $\bar{c} = g(\bar{\kappa})$.[10]

The behavior of the mutual information functions above characterizes the following properties of the *iso-cost function* :

1. $\frac{\partial\delta(\lambda_1,\lambda_2,\bar{\kappa})}{\partial\bar{\kappa}} < 0$, because both mutual information functions are decreasing in δ,
2. $\frac{\partial\delta(\lambda_1,\lambda_2,\bar{\kappa})}{\partial\lambda_1} > 0$, because $\frac{\partial I\left(\tilde{F};\tilde{Y}_1\right)}{\partial\lambda_1} > 0$, and
3. $\frac{\partial\delta(\lambda_1,\lambda_2,\bar{\kappa})}{\partial\lambda_2} < 0$, because $\frac{\partial I\left(\tilde{F};\tilde{Y}_2\right)}{\partial\lambda_2} < 0$.

[10] An alternative approach is to find the iso-cost as a function of the monetary value of the capacity $\bar{\kappa}$. Both are equivalent since there is a one-to-one relationship between the two.

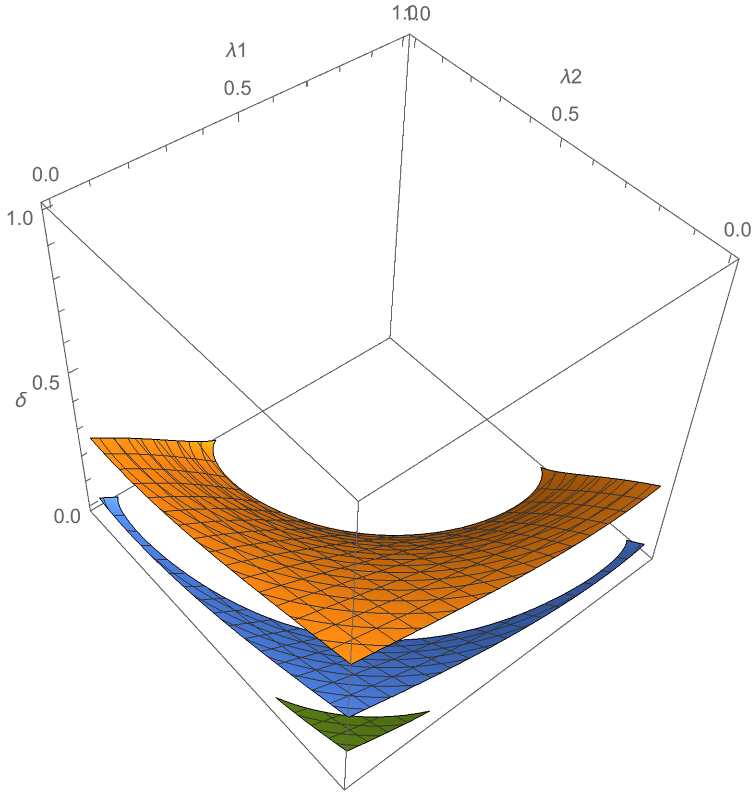

Figure 2.3: This figure contains iso-cost surfaces.

Figure 2.3 depicts these properties over the iso-cost surfaces. The plot contains the full combination of weights and garbling that generate the same cost.

Proposition 5. *Let $\mathring{\kappa} \in \bar{\mathbb{R}}^{+}$ be such that $\delta\left(1, 0, \mathring{\kappa}\right) = 0$. Then, there is no other level of capacity $\bar{\kappa} < \mathring{\kappa}$ that can produce perfectly informative signals, i.e., $\delta\left(1, 0, \bar{\kappa}\right) = 0$.*

Proof. This follows directly from the fact that the *iso-cost* function is strictly decreasing in κ (Property 1). □

This result is sufficient to see that for agents that face a high information-processing constraint, i.e., a low κ, there is an interior solution, which means that the optimal signals that the insider chooses contains garbling and redundant noise. This choice of information with redundant noise is informationally equivalent to having had ex-ante correlated noise. Either way, the information structure cannot generate independent information for both assets. As a result, this choice of information generates price comovement.

As mentioned before, the second stage problem for the insider is equivalent to solving a separate problem for each asset. However, the mutual information function has some useful properties to characterize the cost function, and affect both markets. For example, take the chain rule of mutual information,

$$I\left(\tilde{F};\tilde{Y}_1,\tilde{Y}_2\right)=I\left(\tilde{F};\tilde{Y}_1\right)+I\left(\tilde{F};\tilde{Y}_2|\tilde{Y}_1\right) \tag{2.6}$$

$$\leq I\left(\tilde{F};\tilde{Y}_1\right)+I\left(\tilde{F};\tilde{Y}_2\right), \tag{2.7}$$

which additionally holds with equality when $\tilde{Y}_1$ and $\tilde{Y}_2$ are independent. In fact, this condition shows that the insiders waste some capacity as they purchase redundant information. The mutual information $I\left(\tilde{F};\tilde{Y}_2|\tilde{Y}_1\right)$ determines the informational gain of observing $\tilde{Y}_2$ given $\tilde{Y}_1$. Since the signals are not independent it is smaller than $I\left(\tilde{F};\tilde{Y}_2\right)$ (see the Venn diagram example). Now consider the other extreme case, $\tilde{Y}_1$ and $\tilde{Y}_2$ are fully correlated, $\tilde{Y}_1=\tilde{Y}_2=\tilde{Y}^*$, then (2.6) becomes

$$I\left(\tilde{F};\tilde{Y}^*\right)\leq 2I\left(\tilde{F};\tilde{Y}^*\right),$$

or, equivalently, if the agent buys two signals the second would be completely redundant. This means that the cost function has a kink at the point $\lambda_1=\lambda_2$,

$$c(\delta,\lambda_1,\lambda_2)=\begin{cases} g\left(I\left(\tilde{F};\tilde{Y}_1\right)+I\left(\tilde{F};\tilde{Y}_2\right)\right) & \lambda_1>\lambda_2\\ g\left(I\left(\tilde{F};\tilde{Y}_1\right)\right) & \lambda_1=\lambda_2.\end{cases}$$

If a fixed capacity level $\bar{\kappa}$ is set, then, $\delta\left(\lambda,\lambda,\bar{\kappa}\right)<\delta\left(\lambda_1,\lambda_2,\bar{\kappa}\right)$ for all $\lambda_1>\lambda_2$ since $I\left(\cdot\right)$ is decreasing in δ.

Example: Venn diagram

To summarize the signal construction process, consider this illustrative example. Figure 2.4 uses Venn diagrams to show the information required to explain each variable $H\left(\cdot\right)$ as well as the mutual information between the variables. The blue shaded area represents the amount of information required to explain asset payoff $\tilde{F}_1$, the red shaded area for $\tilde{F}_2$, light green $\tilde{Y}_1$, and darker green $\tilde{Y}_2$. Given that $\tilde{Y}_1$ and $\tilde{Y}_2$ are signals, this area represents the information that the insider has. The intersections represent the mutual information, for example the intersection between the blue and light green areas represents the mutual information between $\tilde{F}_1$ and $\tilde{Y}_1$, $I\left(\tilde{F}_1;\tilde{Y}_1\right)$, or equivalently, the amount of information revealed about $\tilde{F}_1$ once the signal $\tilde{Y}_1$ is revealed. Note that the blue and red areas do not intersect since the asset payoffs are independent. Also note that both green areas intersect the red and blue areas indicating that both signals are informative for both asset payoffs. However, it shows that $I\left(\tilde{F}_i;\tilde{Y}_i\right) \geq I\left(\tilde{F}_i;\tilde{Y}_j\right)$. Finally, note that the intersection between the two green areas shows that there is some capacity allocated to the introduction of redundant information so it is best for all the capacity to be devoted to the construction of a single signal that is informative for both asset payoffs, i.e., only one green area for both.

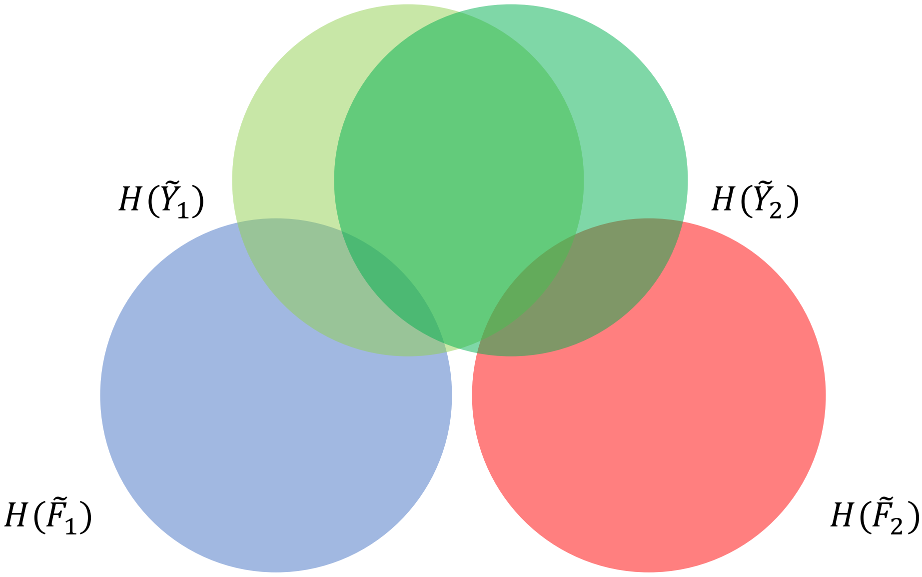

Figure 2.4: Venn diagram for mutual information between the asset payoffs $(\tilde{F})$ and the signals $(\tilde{Y})$

First stage: Attention allocation

After the signal structure has been determined, the insider chooses the distribution of weights within the signal in order to maximize his expected revenue. Then, the insider's optimal attention allocation solves the following expected profit maximization

$$\max_{\delta,\lambda_1,\lambda_2} \pi\left(\delta,\lambda_1,\lambda_2\right)$$

such that

$$c\left(\delta,\lambda_1,\lambda_2\right) \leq g\left(\bar{\kappa}\right),$$

where

$$\pi\left(\delta,\lambda_1,\lambda_2\right)=\Pr\left(Y_1=1\right)\left(\bar{F}_{1|1,1}-p_1\left(1\right)\right)+\left(1-\Pr\left(Y_1=1\right)\right)\left(p_1(-1)-\bar{F}_{1|1,0}\right) \quad (2.8)$$
$$+\Pr\left(Y_2=1\right)\left(\bar{F}_{2|2,1}-p_2\left(1\right)\right)+\left(1-\Pr\left(Y_2=1\right)\right)\left(p_2(-1)-\bar{F}_{2|2,0}\right).$$

There is an interior solution to this problem, at least for values $0<\kappa<\bar{\kappa}$ and that gives the weights that an optimal signal should assign. The response of the expected profit function to the information structure parameters characterizes the solution. Proposition 9 describes the response of the expected profit function to garbling and redundant noise.

Proposition 6. *The expected profit function $\pi\left(\delta,\lambda_1,\lambda_2\right)$ is decreasing both in garbling and redundant noise*

$$\frac{\partial\pi\left(\delta,\lambda_1,\lambda_2\right)}{\partial\delta}<0,$$

$$\frac{\partial\pi\left(\delta,\lambda_1,\lambda_2\right)}{\partial\left(1-\lambda_1\right)}<0,\ \textit{and}$$

$$\frac{\partial\pi\left(\delta,\lambda_1,\lambda_2\right)}{\partial\lambda_2}<0.$$

Proof. Due to the informativeness constraints and from the second stage optimization, as long as the signal is informative, the informational advantages $\left(\bar{F}_{1|1,1}-p_1\left(1\right)\right)$, $\left(p_1(-1)-\bar{F}_{1|1,0}\right)$, $\left(\bar{F}_{2|2,1}-p_2\left(1\right)\right)$ and $\left(p_2(-1)-\bar{F}_{2|2,0}\right)$ are strictly positive. Furthermore, they are decreasing in terms of garbling, i.e.,

$$\frac{\partial\left(\bar{F}_{1|1,1}-p_1\left(1\right)\right)}{\partial\delta},\frac{\partial\left(p_1(-1)-\bar{F}_{1|1,0}\right)}{\partial\delta},\frac{\partial\left(\bar{F}_{2|2,1}-p_2\left(1\right)\right)}{\partial\delta},\frac{\partial\left(p_2(-1)-\bar{F}_{2|2,0}\right)}{\partial\delta}<0,$$

$$\frac{\partial\left(\bar{F}_{1|1,1}-p_1\left(1\right)\right)}{\partial 1-\lambda_1},\frac{\partial\left(p_1(-1)-\bar{F}_{1|1,0}\right)}{\partial 1-\lambda_1}<0,$$

and

$$\frac{\partial\left(\bar{F}_{2|2,1}-p_2\left(1\right)\right)}{\partial\lambda_2},\frac{\partial\left(p_2(-1)-\bar{F}_{2|2,0}\right)}{\partial\lambda_2}<0.$$

The results follow from the lower response of prices to an increase in garbling, with respect to the insider's expectations, due to the additional noise that liquidity traders generate for specialists. □

Furthermore, the response of the profit function to garbling is higher in magnitude than the response to either of the redundant noise parameters $(1-\lambda_1)$ and $\lambda_{2.}$. Since δ is the probability of the extraction of exogenous noise in the signal, and λ_i is a conditional probability, given that the outcome is not noise, then the overall effect of changes in δ are higher than the effect of either of the λ_i $i=1,2$. In fact, all marginal effects of the weight λ_i are scaled down by $(1-\delta)$.

Definition. For all $0<\iota<\mathring{\kappa}$, the ι-th iso-information surface is formed by all points $(\delta^\iota,\lambda_1^\iota,\lambda_2^\iota)\in(0,1)\times(0,1)\times(0,1)$ such that $I\left(\tilde{F};\tilde{Y}_1,\tilde{Y}_2\right)|_{\left(\delta^\iota,\lambda_1^\iota,\lambda_2^\iota\right)}=\iota$.

The iso-information surface contains all the combinations of garbling δ and weights λ_i, $i=1,2$ that deliver informationally equivalent signals about the vector $\tilde{F}$. Therefore, the iso-information surface provides a measure of informational efficiency. The following definition characterizes informationally efficient allocations.

Definition. An attention allocation $(\delta,\lambda_1,\lambda_2)$ is said to be informationally efficient if and only if its cost is a linear transformation $g(\iota)$ of the informational content ι.

In general, all informationally efficient allocations are allocations such that their image over the iso-cost function is a linear transformation $g(.)$ of their images over the iso-information surface. Proposition 7 gives the relation between efficiency and independence of signals.

Proposition 7. *All informationally efficient allocations* $(\delta^\iota,\lambda_1^\iota,\lambda_2^\iota)\in(0,1)\times(0,1)\times(0,1)$ *generate independent signals.*

Proof. It follows from the definition of informationally efficient allocations and the chain rule in formula (2.6). □

Informational efficiency is incompatible with any allocation where $\delta=1$, since it generates uninformative signals. Furthermore, informational efficiency reduces the feasible values of parameters to, $\lambda_1=1,\lambda_2=0$ or $\lambda_1=\lambda_2=\lambda$, with $\lambda\in(0,1)$. The above informativeness constraints, allow for the expected profit function to be redefined as an implicit function of iso-profit surfaces.

Definition. The *iso-profit* function is given by $\delta^{\pi}(\lambda_1, \lambda_2, \bar{\pi})$, where

$$\pi\left(\delta^{\pi}(\lambda_1, \lambda_2, \bar{\pi}), \lambda_1, \lambda_2\right) = \bar{\pi}.$$

From equation (2.8), the iso-profit is a continuous function. Also, note that monotonicity alongside with the previous assumption of informativeness in Definition 1 imply $\frac{\partial \pi}{\partial \delta^{\pi}} < 0$.[11] Lemma 8 shows the existence of a solution to the insider's problem with a unidimensional signal. The proof is divided in three parts, first it shows that such values exist, then that this allocation is informationally efficient, and finally that higher expected profits cannot be achieved in a neighborhood around the optimal values. Informational efficiency can be understood intuitively on the Venn diagram example.

Lemma 8. *For a finite processing capacity* $0 < \bar{\kappa} < \mathring{\kappa}$*, there is some* $\delta^*, \lambda^* > 0$ *such that* $(\delta^*, \lambda^*, \lambda^*)$ *is informationally efficient and expected profits are maximized at* $(\delta^*, \lambda^*, \lambda^*)$.

Proof. The existence is provided by the iso-cost curve, which is defined on $\lambda_1 = \lambda_2$ for all $0 < \bar{\kappa} < \mathring{\kappa}$.

Informational efficiency comes directly from Proposition 6.

Suppose that expected profits were not maximized.Due to construction of the cost function, $\delta^*(\lambda^*, \lambda^*, \bar{\kappa}) < \delta(\lambda^*, \lambda_2, \bar{\kappa})$, $\lambda_2 < \lambda^*$. Now, let

$$\begin{aligned} C = & N_{\varepsilon}(\delta^*, \lambda^*, \lambda^*) \cap \{(\delta, \lambda_1, \lambda_2) \\ & \in (0,1) \times (0,1) \times (0,1) \,\text{st}\, c(\delta, \lambda_1, \lambda_2) \\ & = g(\bar{\kappa})\}, \end{aligned}$$

be the set of all affordable allocations within an $\varepsilon > 0$ range of $(\delta^*, \lambda^*, \lambda^*)$. Now, consider the objective function evaluated in the solution, $\pi(\delta^*, \lambda^*, \lambda^*)$ and the following deviations from the equilibrium

$$\begin{aligned} \pi\left(\delta^*, \lambda^*, \lambda_2^{'}\right) &> \pi(\delta^*, \lambda^*, \lambda^*) \text{ iff } \lambda_2^{'} < \lambda^*, \\ \pi\left(\delta^*, \lambda_1^{'}, \lambda^*\right) &> \pi(\delta^*, \lambda^*, \lambda^*) \text{ iff } \lambda_1^{'} > \lambda^{*,} \\ \pi\left(\delta^{'}, \lambda^*, \lambda^*\right) &> \pi(\delta^*, \lambda^*, \lambda^*) \text{ iff } \delta^{'} < \delta^*. \end{aligned}$$

Which follows from Proposition 6. These conditions imply that if an allocation in C delivers higher expected profits, it should either satisfy $\delta' < \delta^*$, $\lambda_1' > \lambda^*$, or $\lambda_2' < \lambda^*$.

[11] Alternatively, compute the derivative to confirm the sign.

Consider the case where $\lambda_2' < \lambda^*$, if it is a solution to the problem, it should be over the same iso-cost surface, since $\left(\delta^*, \lambda^*, \lambda_2'\right) \notin C$ then, either $\lambda_1' < \lambda^*$ or $\delta' > \delta^*$, or both.

Moreover, I consider the allocation $(\delta', \lambda_1', \lambda_2')$ where expected profits are higher. If the insider faces additional garbling through the reduction of λ_1 to match λ_2', the resulting gain in expected profits from the reduction of δ dominates the effect from the reduction of λ_1. Over the iso-cost surface

$$\begin{aligned}\lim_{\varepsilon \to 0^+} \delta\left(\lambda_1, \lambda_1 - \varepsilon, \bar{\kappa}\right) &= \delta\left(\lambda_1, \lambda_1, \frac{\bar{\kappa}}{2}\right) \\ &> \delta\left(\lambda_1, \lambda_1, \bar{\kappa}\right),\end{aligned}$$

and

$$\begin{aligned}\lim_{\varepsilon \to 0^+} \delta\left(\lambda_2 + \varepsilon, \lambda_2, \bar{\kappa}\right) &= \delta\left(\lambda_2, \lambda_2, \frac{\bar{\kappa}}{2}\right) \\ &> \delta\left(\lambda_2, \lambda_2, \bar{\kappa}\right).\end{aligned}$$

Which is contradicts Proposition 5. Therefore, expected profits are maximized at $(\delta^*, \lambda^*, \lambda^*)$.

□

The last part of the proof relies on the fact that the effect of parameter δ dominates the effect of the changes in λ_1 and λ_2. Hence, as long as there is a possibility to reduce δ while λ changes, this choice dominates any other effect. Lemma 8 states that for a given capacity, there is a point along the iso-cost where δ and $\bar{\kappa}$ are such that $\lambda_1 = \lambda_2 = \lambda$.

2.4.3 Price comovement and market orders.

The takeaway of the previous subsection, specifically Lemma 8, is the willingness of an insider to sacrifice one dimension of the information structure in exchange for overall accuracy. Such behavior is exhibited by insiders who face a very tight capacity constraint.

In consequence, the prices set in the market would be

$$p_1(1) = \frac{a(1-\delta)\left(\lambda + (1-\lambda)\bar{F}_2\right) + \frac{1}{2}(a\delta + 1 - a)}{a(1-\delta)\left(\lambda\bar{F}_1 + (1-\lambda)\bar{F}_2\right) + \frac{1}{2}(a\delta + 1 - a)}\bar{F}_1,$$
$$p_1(-1) = \frac{a(1-\delta)(1-\lambda)\left(1-\bar{F}_2\right) + \frac{1}{2}(a\delta + 1 - a)}{a(1-\delta)\left(\lambda\left(1-\bar{F}_1\right) + (1-\lambda)\left(1-\bar{F}_2\right)\right) + \frac{1}{2}(a\delta + 1 - a)}\bar{F}_1,$$

$$p_2(1) = \frac{a(1-\delta)\left(\lambda\bar{F}_1 + (1-\lambda)\right) + \frac{1}{2}(a\delta + 1 - a)}{a(1-\delta)\left(\lambda\bar{F}_1 + (1-\lambda)\bar{F}_2\right) + \frac{1}{2}(a\delta + 1 - a)}\bar{F}_2, \text{ and}$$
$$p_2(-1) = \frac{a(1-\delta)\lambda\left(1-\bar{F}_1\right) + \frac{1}{2}(a\delta + 1 - a)}{a(1-\delta)\left(\lambda\left(1-\bar{F}_1\right) + (1-\lambda)\left(1-\bar{F}_2\right)\right) + \frac{1}{2}(a\delta + 1 - a)}\bar{F}_2,$$

and the market orders are

$$x\left(\tilde{Y}\right) = \begin{cases} (1,1)' & \text{if } Y = 1 \\ (-1,-1)' & \text{if } Y = 0 \end{cases}.$$

This section has shown the difference in price formation when the insiders are perfectly informed and when they face information constraints. An information-processing constraint can explain the existence of market anomalies because, in pursuit of precision, the insider optimally decides to constrain his action space. In that sense, there is not external force that could generate direct price comovement.

2.5 A market maker

As mentioned in Section 2.3, the trading protocol can influence the behavior of agents in different ways. In particular, a trading protocol where a market maker processes the orders faces a broader information set than two separate specialists do. Conversely, a specialist does not have enough information about the investor's profile when observing the order flow for only one asset. In this section, the trading protocol requires a market maker to trade both assets, i.e., to review market orders for both assets together. This trading protocol is less common than specialists trading in stock markets. The model considered for this analysis considers the same investor profiles as in the previous section, insiders and noise traders, who both submit their orders to a market maker. The protocol allows the market maker to update her beliefs about the investor profiles through the order flow. As a result, in the pricing rule, the market maker cross-subsidizes for potential losses across markets. This section concludes with a comparative statics exercise to examine the differences in an insider's choices of information when trading before specialists and market makers.

2.5.1 Perfect information

Insiders observe the asset payoff vector before it is public information to the other investors and place the market order after they update their beliefs. Noise traders place orders for the two assets simultaneously, and contrary to specialists, the market maker observes an order flow $\tilde{\omega} \in \{(1,1),(1,-1),(-1,1),(-1,-1)\}$. Since the insiders have perfect information, the market orders are

$$x\left(\tilde{F}\right) = \begin{cases} (1,1)' & \text{if } F = (1,1)' \\ (1,-1)' & \text{if } F = (1,0)' \\ (-1,1)' & \text{if } F = (0,1)' \\ (-1,-1)' & \text{if } F = (0,0)'. \end{cases}$$

Simultaneously, noise traders place random liquidity orders, in which all outcomes are equally likely, that is $\tilde{z} = \frac{1}{4}\circ(1,1)' \oplus \frac{1}{4}\circ(1,-1)' \oplus \frac{1}{4}\circ(-1,1)' \oplus \frac{1}{4}\circ(-1,-1)'$. Once the agents are called to trade, the market maker observes the order flow, $\tilde{\omega} = a \circ x\left(\tilde{F}\right) \oplus (1-a)\circ\tilde{z}$, and announces the corresponding price vector. The prices she sets for $\omega = (1,1)$ are

$$\begin{aligned}
p(1,1) &= \mathbb{E}\left[\tilde{F}|\omega = (1,1)\right] \\
&= \begin{bmatrix} 1 \\ 1 \end{bmatrix} \Pr\left(F = (1,1)' \,|\omega = (1,1)'\right) + \begin{bmatrix} 1 \\ 0 \end{bmatrix} \Pr\left(F = (1,0)' \,|\omega = (1,1)'\right) \\
&+ \begin{bmatrix} 0 \\ 1 \end{bmatrix} \Pr\left(F = (0,1)' \,|\omega = (1,1)'\right) \\
&= \begin{bmatrix} 1 \\ 1 \end{bmatrix} \frac{\Pr\left(\omega = (1,1)' \,|F = (1,1)'\right) \Pr\left(F = (1,1)'\right)}{\Pr\left(\omega = (1,1)'\right)} \\
&+ \begin{bmatrix} 1 \\ 0 \end{bmatrix} \frac{\Pr\left(\omega = (1,1)' \,|F = (1,0)'\right) \Pr\left(F = (1,0)'\right)}{\Pr\left(\omega = (1,1)'\right)} \\
&+ \begin{bmatrix} 0 \\ 1 \end{bmatrix} \frac{\Pr\left(\omega = (1,1)' \,|F = (0,1)'\right) \Pr\left(F = (0,1)'\right)}{\Pr\left(\omega = (1,1)'\right)}
\end{aligned}$$

$$\begin{aligned}
&= \begin{bmatrix} \frac{\left(a+\frac{1-a}{4}\right)\bar{F}_1\bar{F}_2}{a\bar{F}_1\bar{F}_2+\frac{1-a}{4}} + \frac{\frac{1-a}{4}\bar{F}_1\left(1-\bar{F}_2\right)}{a\bar{F}_1\bar{F}_2+\frac{1-a}{4}} \\ \frac{\left(a+\frac{1-a}{4}\right)\bar{F}_1\bar{F}_2}{a\bar{F}_1\bar{F}_2+\frac{1-a}{4}} + \frac{\frac{1-a}{4}\left(1-\bar{F}_1\right)\bar{F}_2}{a\bar{F}_1\bar{F}_2+\frac{1-a}{4}} \end{bmatrix} = \begin{bmatrix} \frac{\left(a+\frac{1-a}{4}\right)\bar{F}_1\bar{F}_2+\frac{1-a}{4}\bar{F}_1\left(1-\bar{F}_2\right)}{a\bar{F}_1\bar{F}_2+\frac{1-a}{4}} \\ \frac{\left(a+\frac{1-a}{4}\right)\bar{F}_1\bar{F}_2+\frac{1-a}{4}\left(1-\bar{F}_1\right)\bar{F}_2}{a\bar{F}_1\bar{F}_2+\frac{1-a}{4}} \end{bmatrix} \\
&= \begin{bmatrix} \frac{\bar{F}_1\left(a\bar{F}_2+\frac{1-a}{4}\right)}{a\bar{F}_1\bar{F}_2+\frac{1-a}{4}} \\ \frac{\left(a\bar{F}_1+\frac{1-a}{4}\right)\bar{F}_2}{a\bar{F}_1\bar{F}_2+\frac{1-a}{4}} \end{bmatrix} = \begin{bmatrix} 1 - \frac{\left(1-\bar{F}_1\right)(1-a)}{4a\bar{F}_1\bar{F}_2+1-a} \\ 1 - \frac{(1-a)\left(1-\bar{F}_2\right)}{4a\bar{F}_1\bar{F}_2+1-a} \end{bmatrix}.
\end{aligned}$$

Similarly,

$$p(1,-1) = \begin{bmatrix} 1 - \frac{\left(1-\bar{F}_1\right)(1-a)}{4a\bar{F}_1\left(1-\bar{F}_2\right)+1-a} \\ \frac{(1-a)\bar{F}_2}{4a\bar{F}_1\left(1-\bar{F}_2\right)+1-a} \end{bmatrix},$$

$$p(-1,1) = \begin{bmatrix} \frac{\bar{F}_1(1-a)}{4a\left(1-\bar{F}_1\right)\bar{F}_2+1-a} \\ 1 - \frac{(1-a)\left(1-\bar{F}_2\right)}{4a\left(1-\bar{F}_1\right)\bar{F}_2+1-a} \end{bmatrix}, \text{ and}$$

$$p(-1,-1) = \begin{bmatrix} \frac{\bar{F}_1(1-a)}{4a\left(1-\bar{F}_1\right)\left(1-\bar{F}_2\right)+1-a} \\ \frac{(1-a)\bar{F}_2}{4a\left(1-\bar{F}_1\right)\left(1-\bar{F}_2\right)+1-a} \end{bmatrix}.$$

These prices differ from the prices that the specialists set. Contrary to specialists, a market maker sets the asset prices conditional on the order flow in all markets. he market maker cross subsidizes for the adverse selection problem as she sets state-contingent price vectors. Compared to specialists, the informational advantage that market makers have is the ability to observe an investor's order for both assets. Recall that an insider is an investor who has access to information for both assets. Therefore, the market maker updates her beliefs beliefs in awareness that an order submitted by an insider contains an informational advantage about both assets. In response, the prices set by the market makers have greater informational content than those of specialists. There are two exceptions where specialists and market makers converge to the same prices, when there are no insiders, $a = 0$, and when all traders are informed $a = 1$. Both situations fully reveal the investors' profiles to the specialist as well as to the market maker.

In this section, I have presented two fundamental differences in the structure of asset prices, derived from the trading protocol. First, I pointed out that prices are not independent across markets, that is, the ask (bid) price for asset 1 changes according to the order submitted for asset 2. As a result, the bid-ask spread is contingent on what is traded in other markets. This feature implies that asset payoffs are not conditionally independent on the order flow, but does not contradict the ex-ante independence of asset payoffs and instead describes a different informational channel that generates asset price comovement. Here, the comovement arises from the market maker's ability to extract additional information on the identity of the trader based on the order vector. As a result, the market maker cross subsidizes for her expected losses across markets, which makes the prices of each asset contingent on the action taken in the other market, i.e., $p_i\left(x_i|x_j = 1\right) \neq p_i\left(x_i|x_j = -1\right)$, $i, j = 1, 2$ and $i \neq j$. The price comovement is generated by the market maker's information structure. When specialists trade, instead of market makers, the only possible source of information-based asset price comovement is through correlated trading. Secondly, I find that prices have a higher informational content when they are set by a market maker, i.e., the asset prices are closer to 1 when the asset has a high payoff and closer to zero when the payoff is low.

2.5.2 Imperfectly informed insiders

Having analyzed the market outcome when the price-setting agents are market makers, I will now discuss the implications of insiders that trade with a limited information set. In this section, insiders trade but cannot observe the true payoff vector $\tilde{F}$. Instead, they use a given information-processing capacity $\bar{\kappa}$ to build informative signals about it. The insiders' portfolio selection problem does not depend on the trading protocol. Thus, this process is identical to the one described in Section 4.2. Furthermore, the insiders' optimal learning process is unaffected by the order submission protocol, and there are at most two informative signals $\tilde{Y}_s = (1-\delta) \circ \left(\lambda_s \circ \tilde{F}_1 \oplus (1-\lambda_s) \circ \tilde{F}_2\right) \oplus \delta \circ \tilde{\varepsilon}_s$, none of which is strictly more informative for both assets. When there are equally informative signals, the insiders are indifferent between the two signals for both assets, i.e., they uses a single signal for both assets. As a result, the market orders are the vector $x\left(\tilde{Y}\right)$ such that,

$$x\left(\tilde{Y}\right) = \arg\max_{x \in X} \mathbb{E}\left[\sum_{i=1}^{2} x_i \left(\tilde{F}_i - p_i(\omega)\right) |\tilde{Y}\right]. \tag{2.9}$$

The problem described in (2.9) is equivalent to (2.1) from the insiders perspective, where the only difference is the reaction of asset prices to the order flow. The resulting market orders are

$$x\left(\tilde{Y}\right) = \begin{cases} (1,1)' & \text{if } Y = (1,1)' \\ (1,-1)' & \text{if } Y = (1,0)' \\ (-1,1)' & \text{if } Y = (0,1)' \\ (-1,-1)' & \text{if } Y = (0,0)'. \end{cases}$$

Sine the informativeness of the signal remains unchanged for the insiders with respect to the previous section, the informativeness conditions stated in Proposition 7 do not depend on the trading protocol.

Bid and ask prices

In order to post the prices, the market maker knows that the order flow is $\tilde{\omega} = a \circ x\left(\tilde{F}\right) \oplus (1-a) \circ \tilde{z}$, and prices according to the zero expected profit rule $p(\tilde{\omega}) = \mathbb{E}\left[\tilde{F}|\tilde{\omega}\right]$. Recall that the market orders $\tilde{x}\left(\tilde{Y}\right)$ inherit the statistical properties of $\tilde{Y}$, So

$$p(1,1) = \mathbb{E}\left[\tilde{F}|\omega = (1,1)\right]$$
$$= \begin{bmatrix} \frac{a\left(\Pr\left(Y=(1,1)'|F=(1,1)'\right)\bar{F}_2+\Pr\left(Y=(1,1)'|F=(1,0)'\right)\left(1-\bar{F}_2\right)\right)+\frac{1}{4}(1-a)}{a\Pr\left(Y=(1,1)'\right)+\frac{1}{4}(1-a)}\bar{F}_1 \\ \frac{a\left(\Pr\left(Y=(1,1)'|F=(1,1)'\right)\bar{F}_1+\Pr\left(Y=(1,1)'|F=(0,1)'\right)\left(1-\bar{F}_1\right)\right)+\frac{1}{4}(1-a)}{a\Pr\left(Y=(1,1)'\right)+\frac{1}{4}(1-a)}\bar{F}_2 \end{bmatrix},$$

$$p(1,-1) = \begin{bmatrix} \frac{a\left(\Pr\left(Y=(1,0)'|F=(1,1)'\right)\bar{F}_2+\Pr\left(Y=(1,0)'|F=(1,0)'\right)\left(1-\bar{F}_2\right)\right)+\frac{1}{4}(1-a)}{a\Pr\left(Y=(1,0)'\right)+\frac{1}{4}(1-a)}\bar{F}_1 \\ \frac{a\left(\Pr\left(Y=(1,0)'|F=(1,1)'\right)\bar{F}_1+\Pr\left(Y=(1,0)'|F=(0,1)'\right)\left(1-\bar{F}_1\right)\right)+\frac{1}{4}(1-a)}{a\Pr\left(Y=(1,0)'\right)+\frac{1}{4}(1-a)}\bar{F}_2 \end{bmatrix},$$

$$p(-1,1) = \begin{bmatrix} \frac{a\left(\Pr\left(Y=(0,1)'|F=(1,1)'\right)\bar{F}_2+\Pr\left(Y=(0,1)'|F=(1,0)'\right)\left(1-\bar{F}_2\right)\right)+\frac{1}{4}(1-a)}{a\Pr\left(Y=(0,1)'\right)+\frac{1}{4}(1-a)}\bar{F}_1 \\ \frac{a\left(\Pr\left(Y=(0,1)'|F=(1,1)'\right)\bar{F}_1+\Pr\left(Y=(0,1)'|F=(0,1)'\right)\left(1-\bar{F}_1\right)\right)+\frac{1}{4}(1-a)}{a\Pr\left(Y=(0,1)'\right)+\frac{1}{4}(1-a)}\bar{F}_2 \end{bmatrix}, \text{ and}$$

$$p(-1,-1) = \begin{bmatrix} \frac{a\left(\Pr\left(Y=(0,0)'|F=(1,1)'\right)\bar{F}_2+\Pr\left(Y=(0,0)'|F=(1,0)'\right)\left(1-\bar{F}_2\right)\right)+\frac{1}{4}(1-a)}{a\Pr\left(Y=(0,0)'\right)+\frac{1}{4}(1-a)}\bar{F}_1 \\ \frac{a\left(\Pr\left(Y=(0,0)'|F=(1,1)'\right)\bar{F}_1+\Pr\left(Y=(0,0)'|F=(0,1)'\right)\left(1-\bar{F}_1\right)\right)+\frac{1}{4}(1-a)}{a\Pr\left(Y=(0,0)'\right)+\frac{1}{4}(1-a)}\bar{F}_2 \end{bmatrix}.$$

See appendix B for the detailed conditional probabilities.

Attention allocation

The insider allocates attention to maximize his expected profit function, given the pricing reaction functions in the trading stage. The first-stage expected profit function is

$$\begin{aligned}\pi^{Port}(\delta,\lambda_1,\lambda_2) = {} & \Pr(Y=(1,1))\left(\left(\bar{F}_{1|(1,1)} - p_1(1,1)\right) + \left(\bar{F}_{2|(1,1)} - p_2(1,1)\right)\right) \\ & + \Pr(Y=(1,0))\left(\left(\bar{F}_{1|(1,0)} - p_1(1,0)\right) + \left(p_2(1,0) - \bar{F}_{2|(1,0)}\right)\right) \\ & + \Pr(Y=(0,1))\left(\left(p_1(0,1) - \bar{F}_{1|(0,1)}\right) + \left(\bar{F}_{2|(0,1)} - p_2(0,1)\right)\right) \\ & + \Pr(Y=(0,0))\left(\left(p_1(1,0) - \bar{F}_{1|(0,0)}\right) + \left(p_2(1,0) - \bar{F}_{2|(0,0)}\right)\right),\end{aligned}$$

where $\bar{F}_{i|(j,l)} = \Pr(F_i = 1|Y = (j,l))$, $i = 1,2$ and $j,l = 0,1$. This expected profits can be re-expressed in terms of the probabilities of the signals $\tilde{Y}_1$ and $\tilde{Y}_2$ separately, given the way that the insider chooses to act in the second stage, i.e.,

$$\begin{aligned}\pi^{Port}(\delta,\lambda_1,\lambda_2) = {} & \Pr(Y_1=1)\left(\bar{F}_{1|1,1}-\bar{Y}_{2|1}p_1(1,1)-\left(1-\bar{Y}_{2|1}\right)p_1(1,0)\right)\\ & +\Pr(Y_2=1)\left(\bar{F}_{2|2,1}-\bar{Y}_{1|1}p_2(1,1)-\left(1-\bar{Y}_{1|1}\right)p_2(0,1)\right)\\ & +\Pr(Y_1=0)\left(\bar{Y}_{2|0}p_1(0,1)+\left(1-\bar{Y}_{2|0}\right)p_1(0,0)-\bar{F}_{1|1,0}\right)\\ & +\Pr(Y_2=0)\left(\bar{Y}_{1|0}p_2(1,0)+\left(1-\bar{Y}_{1|0}\right)p_2(0,0)-\bar{F}_{2|2,0}\right),\end{aligned} \tag{2.10}$$

where $\bar{Y}_{s|j} = \Pr(Y_s=1|Y_r=j) = (1-\delta)\left(\lambda_s\bar{F}_{1|r,j}+(1-\lambda_s)\bar{F}_{2|r,j}\right)+\frac{1}{2}\delta,\ s,r = 1,2,$ $s\neq r$ and $j = 0,1$. The expected profit function in (2.10) resembles (2.8), where the informational advantages are proportional to the weighted average between the bid and ask prices for each asset. Note that the informativeness of the signals guarantees positive expected profits since, $\bar{F}_{1|1,1} > p_1(1,1), p_1(1,0)$, $\bar{F}_{2|2,1} > p_2(1,1), p_2(0,1)$, and $\bar{F}_{1|1,0} < p_1(0,1), p_1(0,0)$, $\bar{F}_{2|2,0} < p_2(1,0)$, $p_2(0,0)$. This is rather similar to the behavior observed in the previous section. The following proposition summarizes the effect of the weights of the uninformative elements of the signal on the insider's expected profits.

Proposition 9. *The expected profit function under the portfolio trading protocol* $\pi^{Port}(\delta,\lambda_1,\lambda_2)$ *is decreasing in both garbling, and redundant noise*

$$\frac{\partial\pi^{Port}(\delta,\lambda_1,\lambda_2)}{\partial\delta} < 0,$$

$$\frac{\partial\pi^{Port}(\delta,\lambda_1,\lambda_2)}{\partial(1-\lambda_1)} < 0,$$

and

$$\frac{\partial\pi^{Port}(\delta,\lambda_1,\lambda_2)}{\partial\lambda_2} < 0.$$

Proof. See Proposition 6. □

The expected profit function under the portfolio trading protocol is decreasing in garbling as in the previous section. The next step is to examine the effect on the expected profits of both protocols. Proposition 7 shows that the finer information set available to market makers reduces the insider's expected profits in comparison to the expected profits achieved from trading with specialists. More specifically, the insiders' informational advantages are less valuable when they trade portfolios since this trading protocol discloses additional information about the insiders. Therefore, the expected profits are lower for the insiders.

Proposition 10. *For any given information structure* $(\delta, \lambda_1, \lambda_2)$*, where* $\delta \in [0,1]$ *and* $0 < \lambda_2 \leq \lambda_1$*, specialist trading generates higher expected profits for the insider compared to marker maker trading,* $\pi(\delta, \lambda_1, \lambda_2) \geq \pi^{Port}(\delta, \lambda_1, \lambda_2)$ *.*

Proof. Consider the cases of perfect information introduced at the beginning of sections 2.4 and 2.5. It is straightforward to see that the constraint does not bind since the informational advantage in every state is higher when specialists trade the assets. Now consider an information structure $(\delta, \lambda_1, \lambda_2)$, and compare the informational advantages in every possible state for both (2.8) and (2.10). The analysis is is performed through the comparison of the bid and ask prices because the expected payoffs only depend on the information structure and not on the trading protocol. Generically, for asset 1 the bid and ask prices set by a specialists satisfy

$$p_1(1) \leq \bar{Y}_{2|1} p_1(1,1) + \left(1 - \bar{Y}_{2|1}\right) p_1(1,0)$$

and

$$p_1(-1) \geq \bar{Y}_{2|0} p_1(0,1) + \left(1 - \bar{Y}_{2|0}\right) p_1(0,0).$$

Following a similar procedure for the two assets and signal outcomes, the inequalities hold to the benefit of the insider when specialists trade. Hence, $\pi(\delta, \lambda_1, \lambda_2) \geq \pi^{Port}(\delta, \lambda_1, \lambda_2)$. □

The previous proposition states that the informational gains from inside information are slightly diminished by the additional information that the market maker obtains due to observing the whole order flow. However, there is still one unexplored case in the previous proposition, the acquisition of a single signal, $\lambda_1 = \lambda_2 = \lambda$.

If $\lambda_1 = \lambda_2 = \lambda$, then (2.10) can be re-written as

$$\pi^{Port}(\delta, \lambda, \lambda) = \Pr(Y=1)\left(\bar{F}_{1|1} - p_1(1,1) + \bar{F}_{2|1} - p_2(1,1)\right)$$
$$+ (1 - \Pr(Y=1))\left(p_1(0,0) - \bar{F}_{1|0} + p_2(0,0) - \bar{F}_{2|0}\right),$$

and (2.8)

$$\pi(\delta,\lambda,\lambda) = \Pr(Y=1)\left(\bar{F}_{1|1} - p_1(1) + \bar{F}_{2|1} - p_2(1)\right)$$
$$+ (1-\Pr(Y=1))\left(p_1(-1) - \bar{F}_{1|0} + p_2(-1) - \bar{F}_{2|0}\right),$$

where $\bar{F}_{i|j} = \Pr(F_i = 1|Y=j)$. The previous result follows from the acquisition of a single signal, which implies that the market orders are perfectly correlated across markets. Moreover, the corresponding prices for asset 1 set by the specialists and market makers respectively are

$$p_1(1) = \frac{a(1-\delta)\left(\lambda + (1-\lambda)\bar{F}_2\right) + \frac{1}{2}(a\delta + 1 - a)}{a(1-\delta)\left(\lambda\bar{F}_1 + (1-\lambda)\bar{F}_2\right) + \frac{1}{2}(a\delta + 1 - a)}\bar{F}_1,$$
$$p_1(1,1) = \frac{a(1-\delta)\left(\lambda + (1-\lambda)\bar{F}_2\right) + \frac{1}{2}\delta a + \frac{1}{4}(1-a)}{a(1-\delta)\left(\lambda\bar{F}_1 + (1-\lambda)\bar{F}_2\right) + \frac{1}{2}\delta a + \frac{1}{4}(1-a)}\bar{F}_1,$$

where $p_1(1,1) > p_1(1)$. In general, $p_i(1,1) > p_i(1)$ and $p_i(-1,-1) < p_i(-1)$ for $i = 1,2$. The market maker knows that when insiders choose to buy only one signal their market orders are perfectly correlated. Therefore, whenever a market maker observes an order flow that contains opposite orders for the two assets, she knows that there is no inside information behind this order. As a result, she sets $p_1(1,-1) = \bar{F}_1$, and $p_2(-1,1) = \bar{F}_2$. The following lemma shows that under portfolio trading protocol, there is an informationally efficient equilibrium with only one signal.

Lemma 11. *For a finite processing capacity* $0 < \bar{\kappa} < \mathring{\kappa}$*, there exist some* $\delta^{**}, \lambda^{**} > 0$ *such that* $(\delta^{**}, \lambda^{**}, \lambda^{**})$ *is informationally efficient and expected profits are maximized at* $(\delta^{**}, \lambda^{**}, \lambda^{**})$.

Proof. Since Proposition 9 indicates that the expected profit function is decreasing in the garbling and redundant noise parameters just as it was under the asset trading protocol, and the iso-cost function is not conditional on the trading protocol. Then, the same conditions as in Lemma 8 are satisfied. □

Lemma 11 implies that an informationally constrained insider optimally chooses only one signal. Hence, the market orders are perfectly correlated through two different sources of information-induced price comovement. This paper concludes with a proposed extension to the model where there is a previous stage in which insiders purchase an expected profit-maximizing quantity of information. In this extension, I analyze the effect on prices of a regulatory change in the trading protocol.

Extension: a change of protocol.

This section has shown that under identical market conditions and different trading protocols, the insider's expected profits are always lower when the market maker trades portfolios. These models allow for the capacity level κ to be endogenously determined through a profit maximization problem. However, I do not address this specific problem above since it is not relevant to the discussion.

For this extension, I assume that there is a competitive market for information where the insiders can purchase the signals. Once the signal(s) is(are) purchased, the optimal values for the parameters in the signals solve the aforementioned implicit functions of κ. This endogenous choice of the κ occurs in a previous stage, stage 0, and the other two stages were discussed previously. In stage 1 the agent chooses the information structure given κ, and in stage 2 trading takes place. For the optimal level of information-processing capacity, in stage 0, the insider chooses a capacity level such that its expected marginal revenue is equal to the marginal cost. In stage 0, the expected revenue for the insider corresponds to his value function in stage 1.

Let $\Pi(\kappa)$, be the stage 1 value function under a protocol where specialists trade, and let $\Psi(\kappa)$ be the stage 1 value function when market makers trade portfolios. Also, define the value $\mathring{\kappa}$ that $\Pi'(\kappa), \Psi'(\kappa) = 0$ for all $\kappa > \mathring{\kappa}$. Additionally, consider the cost function $g(\kappa)$ with a constant marginal cost of information, where $g'(\cdot) > 0$ and $g''(\cdot) = 0$.

Initially, the insider chooses the information-processing capacity before trading under an asset trading protocol, $\kappa^* = \arg\max_{\kappa} \Pi(\kappa) - g(\kappa)$. Let this particular insider face market conditions such that $\kappa^* = \mathring{\kappa}$, and $\Pi(\kappa^*) - g(\kappa^*) = \nu$, i.e., his optimal capacity level is just high enough to purchase a perfectly informative signal. After the insider has chosen his optimal capacity κ^*, there is a regulatory change in this market and the specialists become market makers and start trading portfolios.

If there is a change in the trading protocol, the expected losses for the insider are $\Pi(\kappa^*) - \Psi(\kappa^*)$. Consider $\Pi(\kappa^*) - \Psi(\kappa^*) < \nu$, and $\Psi(\kappa^*) - g'(\kappa^*) < 0$. Then, there must be a new optimal capacity $\kappa^{**} = \arg\max_\kappa \Psi(\kappa) - g(\kappa) < \kappa^* = \mathring{\kappa}$, which means that the same insider does not find κ^* optimal anymore. As a result, an insider whose level of expected profits permitted the acquisition of perfect information when specialists were trading cannot reach the profit threshold for the acquisition of perfect information anymore. In other words, this is an insider whose expected profits are just enough to optimally acquire perfect information when a specialist is trading assets. The lower informational advantages implied by market makers trading portfolios reduce the expected profits. Therefore, the acquisition of perfect information is no longer optimal for the insider described here.

2.6 Concluding remarks

I use a bivariate model for insider trading to provide a rational explanation for the information-based comovement of prices across financial assets. This model allows the insider to characterize the optimal information structure he accesses during the trading decisions. As a result, it provides a potential explanation as to why, in the pursuit of higher precision, insiders are willing to lose the statistical independence of the information acquired. Moreover, market orders inherit the acquired information's statistical properties, and with it, the loss of statistical independence induces perfectly correlated market orders. The results are obtained through the characterization of the expected profit function and its response to changes in the garbling and redundant noise parameters. Through the model, I find a relevant mechanism that links the expected profit function to the mutual information function.

The bivariate feature of the model extends the explanation of the bid-ask spread as an informational phenomenon beyond the quality of the insider's information. The bivariate choice allows different characterizations of the information set that the price-setting agent faces. Thus, I can explain correlation in the formation of the bid-ask spread as a result of the insider's access to correlated information or a trading protocol that discloses some additional information about the insiders.

Throughout the paper, I analyze two sources of price comovement. First, I consider traders with low processing capacity who supply optimally correlated information through their market orders. Second, I find that market makers cross-subsidize across markets to compensate for the adverse selection problem that insiders pose. Both explanations reveal the existence of an informational channel as a source of artificial price comovement.

This model is suitable to analyze markets for assets whose payoffs are not high enough to compensate for the cost of buying precise information about it. Consequently, traders group some assets into a joint index, based on which they place their orders. The market orders are not affected directly by the protocol, but it does affect the expected profits and the choice of information. Possible extensions to this model include the extension of the time horizon to analyze the evolution and persistence of information-based market anomalies in financial markets and their effect on the bid-ask spread.

Appendix

B Mutual information of signals

This appendix contains the corresponding expression for the mutual information function when the signal is a mixture of Bernoulli variables. From the definition of mutual information function $I\left(\tilde{F}_i|\tilde{Y}_i\right)$ stated in Section 4.2, we have

$$\begin{aligned}
I\left(\tilde{F}_1;\tilde{Y}_1\right) &= H\left(\tilde{F}_1\right) - H\left(\tilde{F}_1|\tilde{Y}_1\right) \\
&= -E\left[\log\left(\Pr\left(\tilde{F}_1\right)\right)\right] - \sum_{i=0}^{1} \Pr\left(Y_1 = i\right) H\left(\tilde{F}_1|Y_1 = i\right) \\
&= -\bar{F}_1 \log\left(\bar{F}_1\right) - \left((1-\delta)\left(\lambda_1 \bar{F}_1 + (1-\lambda_1)\bar{F}_2\right) + \frac{\delta}{2}\right) H\left[\tilde{F}_1|Y = 1\right] \\
&- \left((1-\delta)\left(\lambda_1(1-\bar{F}_1) + (1-\lambda_1)\left(1-\bar{F}_2\right)\right) + \frac{\delta}{2}\right) H\left[\tilde{F}_1|Y_1 = 0\right]
\end{aligned}$$

$$
\begin{aligned}
&= \left(\left(1-\delta\right)\left(\lambda_1+\left(1-\lambda_1\right)\bar{F}_2\right)+\frac{1}{2}\delta\right)\bar{F}_1\log\left(\frac{\left(1-\delta\right)\left(\lambda_1+\left(1-\lambda_1\right)\bar{F}_2\right)+\frac{1}{2}\delta}{\left(1-\delta\right)\left(\lambda_1\bar{F}_1+\left(1-\lambda_1\right)\bar{F}_2\right)+\frac{1}{2}\delta}\bar{F}_1\right) \\
&+ \left(\left(1-\delta\right)\left(1-\lambda_1\right)\left(1-\bar{F}_2\right)+\frac{1}{2}\delta\right)\bar{F}_1 \\
&\cdot\log\left(\frac{\left(1-\delta\right)\left(1-\lambda_1\right)\left(1-\bar{F}_2\right)+\frac{1}{2}\delta}{\left(1-\delta\right)\left(\lambda_1\left(1-\bar{F}_1\right)+\left(1-\lambda_1\right)\left(1-\bar{F}_2\right)\right)+\frac{1}{2}\delta}\bar{F}_1\right)-\bar{F}_1\log\left(\bar{F}_1\right).
\end{aligned}
$$

Similarly, the mutual information function for the other asset payoff and its corresponding signal,

$$
\begin{aligned}
I\left(\tilde{F}_2;\tilde{Y}_2\right) &= \left(\left(1-\delta\right)\left(\lambda_2\bar{F}_1+\left(1-\lambda_2\right)\right)+\frac{1}{2}\delta\right)\bar{F}_2\log\left(\frac{\left(1-\delta\right)\left(\lambda_2\bar{F}_1+\left(1-\lambda_2\right)\right)+\frac{1}{2}\delta}{\left(1-\delta\right)\left(\lambda_2\bar{F}_1+\left(1-\lambda_2\right)\bar{F}_2\right)+\frac{1}{2}\delta}\bar{F}_2\right) \\
&+ \left(\left(1-\delta\right)\lambda_2\left(1-\bar{F}_1\right)+\frac{1}{2}\delta\right)\bar{F}_2 \\
&\cdot\log\left(\frac{\left(1-\delta\right)\lambda_2\left(1-\bar{F}_1\right)+\frac{1}{2}\delta}{\left(1-\delta\right)\left(\lambda_2\left(1-\bar{F}_1\right)+\left(1-\lambda_2\right)\left(1-\bar{F}_2\right)\right)+\frac{1}{2}\delta}\bar{F}_2\right)-\bar{F}_2\log\left(\bar{F}_2\right).
\end{aligned}
$$

B. Distribution and conditional distribution of the signal

In this appendix, I present some probabilities that have been used throughout the analysis. I characterize the distribution of the signal $\tilde{Y}=\left[\tilde{Y}_1,\tilde{Y}_2\right]'$ by using Bayes' rule repeatedly to compute the conditional probabilities. For example, the conditional probability that signal $Y_1=1$ given $Y_2=1$ is

$$
\begin{aligned}
\Pr\left(Y_1=1|Y_2=1\right) &= \left(1-\delta\right)\left(\lambda_1\bar{F}_{1|2,1}+\left(1-\lambda_1\right)\bar{F}_{2|2,1}\right)+\frac{1}{2}\delta \\
&= \left(1-\delta\right)\left(\frac{\left(\left(1-\delta\right)\left(\lambda_2+\left(1-\lambda_2\right)\bar{F}_2\right)+\frac{1}{2}\delta\right)\lambda_1\bar{F}_1}{\left(1-\delta\right)\left(\lambda_2\bar{F}_1+\left(1-\lambda_2\right)\bar{F}_2\right)+\frac{1}{2}\delta}\right. \\
&\left.+\frac{\left(\left(1-\delta\right)\left(\lambda_2\bar{F}_1+\left(1-\lambda_2\right)\right)+\frac{1}{2}\delta\right)\left(1-\lambda_1\right)\bar{F}_2}{\left(1-\delta\right)\left(\lambda_2\bar{F}_1+\left(1-\lambda_2\right)\bar{F}_2\right)+\frac{1}{2}\delta}\right)+\frac{1}{2}\delta \\
&= \frac{\left(1-\delta\right)^2\left(\lambda_2\lambda_1\bar{F}_1+\left(\left(1-\lambda_2\right)\lambda_1+\lambda_2\left(1-\lambda_1\right)\right)\bar{F}_1\bar{F}_2+\left(1-\lambda_2\right)\left(1-\lambda_1\right)\bar{F}_2\right)}{\Pr\left(Y_2=1\right)} \\
&+ \frac{\frac{1}{2}\delta\left(\Pr\left(Y_1=1\right)+\Pr\left(Y_2=1\right)-\frac{1}{2}\delta\right)}{\Pr\left(Y_2=1\right)}.
\end{aligned}
$$

Table 2.2 summarizes all the possible cases. Panel A displays the conditional probability of signal $\tilde{Y}_1$ given $\tilde{Y}_2$, and Panel B the conditional probability of signal $\tilde{Y}_2$ given $\tilde{Y}_1$.

Panel A

Y_1	Cond. on Y_2	Probability
1	1	$\frac{(1-\delta)^2\left(\lambda_2\lambda_1\bar{F}_1+((1-\lambda_2)\lambda_1+\lambda_2(1-\lambda_1))\bar{F}_1\bar{F}_2+(1-\lambda_2)(1-\lambda_1)\bar{F}_2\right)}{\Pr(Y_2=1)}+\frac{\frac{1}{2}\delta\left(\Pr(Y_1=1)+\Pr(Y_2=1)-\frac{1}{2}\delta\right)}{\Pr(Y_2=1)}$
	0	$\frac{(1-\delta)^2\left((1-\lambda_2)\bar{F}_2\lambda_1\left(1-\bar{F}_1\right)+\lambda_2\bar{F}_1(1-\lambda_1)\left(1-\bar{F}_2\right)\right)}{\Pr(Y_2=1)}+\frac{\frac{1}{2}\delta\left(\Pr(Y_1=0)+\Pr(Y_2=1)-\frac{1}{2}\delta\right)}{\Pr(Y_2=1)}$
0	1	$\frac{(1-\delta)^2\left((1-\lambda_2)\bar{F}_2\lambda_1\left(1-\bar{F}_1\right)+\lambda_2\bar{F}_1(1-\lambda_1)\left(1-\bar{F}_2\right)\right)}{\Pr(Y_2=1)}+\frac{\frac{1}{2}\delta\left(\Pr(Y_1=0)+\Pr(Y_2=1)-\frac{1}{2}\delta\right)}{\Pr(Y_2=1)}$
	0	$\frac{(1-\delta)^2\left(\lambda_2\lambda_1\left(1-\bar{F}_1\right)+((1-\lambda_2)\lambda_1+\lambda_2(1-\lambda_1))\left(1-\bar{F}_1\right)\left(1-\bar{F}_2\right)\right)}{\Pr(Y_1=0)}+\frac{(1-\delta)^2(1-\lambda_2)(1-\lambda_1)\left(1-\bar{F}_2\right)+\frac{1}{2}\delta\left(\Pr(Y_1=0)+\Pr(Y_2=0)-\frac{1}{2}\delta\right)}{\Pr(Y_1=0)}$

Panel B

Y_2	Cond. on Y_1	Probability
1	1	$\frac{(1-\delta)^2\left(\lambda_2\lambda_1\bar{F}_1+((1-\lambda_2)\lambda_1+\lambda_2(1-\lambda_1))\bar{F}_1\bar{F}_2+(1-\lambda_2)(1-\lambda_1)\bar{F}_2\right)}{\Pr(Y_1=1)}+\frac{\frac{1}{2}\delta\left(\Pr(Y_1=1)+\Pr(Y_2=1)-\frac{1}{2}\delta\right)}{\Pr(Y_1=1)}$
	0	$\frac{(1-\delta)^2\left((1-\lambda_2)\bar{F}_2\lambda_1\left(1-\bar{F}_1\right)+\lambda_2\bar{F}_1(1-\lambda_1)\left(1-\bar{F}_2\right)\right)}{\Pr(Y_1=0)}+\frac{\frac{1}{2}\delta\left(\Pr(Y_1=0)+\Pr(Y_2=1)-\frac{1}{2}\delta\right)}{\Pr(Y_1=0)}$
0	1	$\frac{(1-\delta)^2\left((1-\lambda_1)\lambda_2\left(1-\bar{F}_1\right)\bar{F}_2+(1-\lambda_2)\lambda_1\bar{F}_1\left(1-\bar{F}_2\right)\right)}{\Pr(Y_1=1)}+\frac{\frac{1}{2}\delta\left(\Pr(Y_1=1)+\Pr(Y_2=0)-\frac{1}{2}\delta\right)}{\Pr(Y_1=1)}$
	0	$\frac{(1-\delta)^2\left(\lambda_2\lambda_1\left(1-\bar{F}_1\right)+((1-\lambda_2)\lambda_1+\lambda_2(1-\lambda_1))\left(1-\bar{F}_1\right)\left(1-\bar{F}_2\right)\right)}{\Pr(Y_2=0)}+\frac{(1-\delta)^2(1-\lambda_2)(1-\lambda_1)\left(1-\bar{F}_2\right)+\frac{1}{2}\delta\left(\Pr(Y_1=0)+\Pr(Y_2=0)-\frac{1}{2}\delta\right)}{\Pr(Y_2=0)}$

Table 2.2: The conditional probabilities are computed using Bayes' rule. Panel A shows the distribution of signal $\tilde{Y}_1$ conditional on $\tilde{Y}_2$, and panel B the converse.

Table 2.3 contains the values of the probabilities for all four possible outcomes. I, use Bayes' rule once again to characterize the conditional distribution of the signal vector given the asset payoffs. Table 2.4 contains the corresponding conditional probabilities.

		$\tilde{Y}_2$	
		1	0
$\tilde{Y}_1$	1	$(1-\delta)^2\left(\lambda_2\lambda_1\bar{F}_1 + ((1-\lambda_2)\lambda_1 + \lambda_2(1-\lambda_1))\bar{F}_1\bar{F}_2 + (1-\lambda_2)(1-\lambda_1)\bar{F}_2\right) + \frac{1}{2}\delta\left(\Pr(Y_1=1) + \Pr(Y_2=1) - \frac{1}{2}\delta\right)$	$(1-\delta)^2\left((1-\lambda_1)\lambda_2\left(1-\bar{F}_1\right)\bar{F}_2 + (1-\lambda_2)\lambda_1\bar{F}_1\left(1-\bar{F}_2\right)\right) + \frac{1}{2}\delta\left(\Pr(Y_1=1) + \Pr(Y_2=0) - \frac{1}{2}\delta\right)$
	0	$(1-\delta)^2\left((1-\lambda_2)\bar{F}_2\lambda_1\left(1-\bar{F}_1\right) + \lambda_2\bar{F}_1(1-\lambda_1)\left(1-\bar{F}_2\right)\right) + \frac{1}{2}\delta\left(\Pr(Y_1=0) + \Pr(Y_2=1) - \frac{1}{2}\delta\right)$	$(1-\delta)^2\left(\lambda_2\lambda_1\left(1-\bar{F}_1\right) + ((1-\lambda_2)\lambda_1 + \lambda_2(1-\lambda_1))\left(1-\bar{F}_1\right)\left(1-\bar{F}_2\right) + (1-\lambda_2)(1-\lambda_1)\left(1-\bar{F}_2\right)\right) + \frac{1}{2}\delta\left(\Pr(Y_1=0) + \Pr(Y_2=0) - \frac{1}{2}\delta\right)$

Table 2.3: Joint distribution of the signal $\tilde{Y}$

Y	Cond. on F	Probability
$(1,1)$	$(1,1)$	$\left(1-\frac{1}{2}\delta\right)^2$
	$(1,0)$	$(1-\delta)^2\lambda_2\lambda_1+\frac{1}{2}\delta(1-\delta)(\lambda_1+\lambda_2)+\left(\frac{1}{2}\delta\right)^2$
	$(0,1)$	$(1-\delta)^2(1-\lambda_2)(1-\lambda_1)+\frac{1}{2}\delta(1-\delta)(2-\lambda_1-\lambda_2)+\left(\frac{1}{2}\delta\right)^2$
	$(0,0)$	$\left(\frac{1}{2}\delta\right)^2$
$(1,0)$	$(1,1)$	$\frac{1}{2}\delta(1-\delta)+\left(\frac{1}{2}\delta\right)^2$
	$(1,0)$	$(1-\delta)^2(1-\lambda_2)\lambda_1+\frac{1}{2}\delta(1-\delta)(\lambda_1+(1-\lambda_2))+\left(\frac{1}{2}\delta\right)^2$
	$(0,1)$	$(1-\delta)^2(1-\lambda_1)\lambda_2+\frac{1}{2}\delta(1-\delta)((1-\lambda_1)+\lambda_2)+\left(\frac{1}{2}\delta\right)^2$
	$(0,0)$	$\frac{1}{2}\delta(1-\delta)+\left(\frac{1}{2}\delta\right)^2$
$(0,1)$	$(1,1)$	$\frac{1}{2}\delta(1-\delta)+\left(\frac{1}{2}\delta\right)^2$
	$(1,0)$	$(1-\delta)^2\lambda_2(1-\lambda_1)+\frac{1}{2}\delta(1-\delta)(\lambda_2+(1-\lambda_1))+\left(\frac{1}{2}\delta\right)^2$
	$(0,1)$	$(1-\delta)^2\lambda_1(1-\lambda_2)+\frac{1}{2}\delta(1-\delta)(\lambda_1+(1-\lambda_2))+\left(\frac{1}{2}\delta\right)^2$
	$(0,0)$	$\frac{1}{2}\delta(1-\delta)+\left(\frac{1}{2}\delta\right)^2$

Table 2.4: Conditional distribution of $\tilde{Y}$ given $\tilde{F}$. The results for $Y=(0,0)$ can be computed using the theorem of total probability.

Chapter 3

Learning specialists and market resilience

3.1 Introduction

The stochastic properties of asset prices are affected when the specialist has a limited capacity to learn from the market and, in a financial market where investors submit orders to specialists, they face an adverse selection problem, since they cannot observe each order's originator. In response, specialists make the market less liquid, and the asset prices differ from their fundamental value. One way to measure liquidity in a market is to determine how quickly asset prices can recover after a large order. This recovery speed is called market resilience, and there are different theories as to why a market is more or less resilient. One of these theories states that market resilience is affected by the ability of the price-setting agent, the specialist, to infer useful information from the orders she observes. These orders are called order flow, and they contain relevant information about the asset payoffs when informed agents, insiders, trade.

For example, in a market where specialists can distinguish between insiders and noise traders, a specialist uses the insiders' trades to set each period's asset price equal to the corresponding asset payoff. In such a market, there is no adverse selection problem. On the contrary, in a market with asymmetric information between traders and a specialist, the specialist cannot identify whether a large order comes from an insider's knowledge of good or bad news about the asset payoff or about a noise trader that submitted random demands.

From here on, I refer to shocks to asset payoffs as structural shocks, and other shocks to the order flow as noise trading shocks. In this paper, the specialist is allowed to allocate an exogenous learning capacity to filter out noise trading to study the effect on the corresponding pricing process.

To serve that purpose, I consider a model where specialists hold inventories of shares in the stock market and there is one specialist trading each asset. Their role is to set the asset prices as they process investors' orders. In each period, the informed investors anonymously feed new information into the market through the market orders. The existence of noise traders provides camouflage for the insider. The specialist cannot perfectly identify the insider's actions since she observes the order flow, preventing her from practicing any form of price discrimination. Here, specialists are endowed with a fixed information-processing capacity (or learning technology). They allocate this capacity to filter out noise trading and identify structural shocks affecting the market. As her learning capacity increases, the specialist reacts faster to structural shocks. In contrast, a smaller learning capacity delays the impact of structural shocks on prices.

Market resilience has two interpretations in financial markets. In this paper, I study resilience in the sense of Kyle (1985), where it is defined as the recovery speed of asset prices from an uninformative shock.[1] A specialist trades in each period, setting the asset prices to meet the expected value of the asset payoffs. For Kyle (1985), any liquidity fluctuation results from an informational asymmetry across investors in the market, as in Glosten and Milgrom (1985). As a result of this asymmetry, the specialist makes the market less liquid, and the prices are highly persistent to uninformative trades. If the specialist resolves the informational asymmetry, this persistence disappears, i.e., the market becomes more resilient.

I use a model for insider trading where specialists face a dynamic pricing problem.

[1] Alternatively, in Garbade (1982) a market is resilient if there is order replenishment, i.e., new orders arise quickly as temporary order imbalances occur. Similar definitions of resilience include the fast convergence of the price to a new steady-state after a marker order in Obizhaeva and Wang (2013), or the inability of informed traders to substantially change the market prices. Hence, resilience does not have a unique definition, and it can refer to as the capacity of either prices or trading volume to recover from external shocks.

In each period, all investors submit their orders to a specialist. After the specialist observes the order flow, which determines her information structure, sets the asset's price to minimize her expected losses. At the end of each period, the asset payoff is observable, so the history of asset payoffs to the previous period and their stochastic processes are common knowledge. The only agent with information about new shocks is the informed investor (or insider). Once the order flow is available, specialists update their beliefs about the new asset payoffs. Therefore, I assume that there is a one risk-averse specialist who sets the asset price in each market.

The model I propose allows the specialist to learn about the market orders, solving the rational inattention problem proposed by Sims (2003). In a model where there are rationally-inattentive agents, these can partially resolve the uncertainty of a variable given an information-processing constraint. Such a constraint is entropy-based and follows the information function in Shannon (1948). The learning process in rational inattention problems allows each agent to characterize an optimal information set according to their objective function. Among other applications, Peng (2005) and Peng and Xiong (2006) first introduced the rational inattention problem to financial economics in a portfolio selection problem. Later on, van Nieuwerburgh and Veldkamp (2009, 2010) explain portfolio under-diversification and the home bias puzzle as a result of rationally-inattentive investors choosing a portfolio. Therefore, in this paper specialists are allowed to allocate their information-processing capacity to reduce the noise in the order flow. By doing so, it is shown that specialists react faster when they have higher capacities, and the effect on prices of a spurious shock in the order flow dissolves quickly, i.e., the market becomes more resilient. Consequently, it is possible to characterize the stochastic behavior of asset prices as the specialist's learning capacity changes.

Maćkowiak *et al.* (2018) provide an alternative solution to the dynamic rational inattention problem in Sims (2003), where analytical solutions can be achieved for several linear quadratic problems. They show that the dynamic rational inattention problem is equivalent to solving for the Kalman gain in the equivalent state-space representation of the problem. This paper shows that when the specialist's utility function is CARA, the pricing problem is represented in linear quadratic form. I then follow Maćkowiak *et al.* (2018) to show that the specialist's pricing process follows the same structure as the stochastic process of asset payoffs. The resulting problem for the specialist is equivalent to solving the Kalman gain in a state-space representation. I represent market resilience in the impulse response functions of asset prices to structural shocks and noise trading.

The paper is structured as follows: Section 2 introduces a static model where pricing occurs through simple Bayesian updating, Section 3 presents some analytical results for ARMA(p, q) processes for the optimal pricing rule and analyzes the price dynamics using impulse response functions. Finally, Section 4 provides some concluding remarks.

3.2 Static choice

This section presents a static model for insider trading where a specialist allocates her learning capacity to identify the insider's market orders. To do so, she allocates her learning capacity to reduce the perceived noise from liquidity trades to increase the observed order flow's precision. Once the signal has been observed, the specialist sets the asset prices to maximize her expected profits.

3.2.1 Market orders

There are three types of investors in this market: a perfectly informed insider, random liquidity traders and a specialist. The insiders and liquidity traders submit their orders to a specialist. The specialist then sets the price and processes the orders. The traded asset has normally distributed payoffs, $\tilde{F} \sim N\left(\bar{F}, \sigma_F^2\right)$. Let $\tilde{f} \equiv \tilde{F} - \bar{F}$ be the deviation of the asset payoff from its mean value, such that $\tilde{f} \sim N\left(0, \sigma_F^2\right)$ contains only the structural shocks to the true payoffs. The noise traders place random liquidity demands $\tilde{z} \sim N\left(0, \sigma_z^2\right)$. The insider places his market order $\tilde{x} = x\left(\tilde{f}\right)$ for the asset in order to maximize the expected profit

$$\mathbb{E}\left[x\left(\tilde{f} - p\right)|\tilde{f}\right],$$

where $p \equiv P - \bar{F}$ is the asset price P rescaled by the mean of the asset payoff, which is set by the specialist after observing the order flow. The market order is a sufficient statistic for the asset payoff whenever the insider is perfectly informed. Moreover, in equilibrium the market order will be a linear function of the centered payoff $\tilde{f}$.

3.2.2 Asset price

There is a single risk-averse specialist who sets the price by maximizing the expected utility function. The constant absolute risk aversion, CARA, preferences are given by

$$u(\tilde{\nu}) = -e^{-\varrho\tilde{\nu}},$$

where ϱ is the risk aversion parameter. Since she is the only specialist, she absorbs the whole order flow, and her profits are determined by the profits of "making the market", i.e., $\tilde{\nu} = \tilde{\omega}\left(P - \tilde{F}\right)$.[2] Therefore, the specialist's preferences are given by

$$\mathbb{E}\left[u(\tilde{\nu})\right] = \mathbb{E}\left[\mathbb{E}\left[\tilde{\nu}|\mathcal{I}\right] - \frac{\varrho}{2}\mathbb{V}\left[\tilde{\nu}|\mathcal{I}\right]\right], \tag{3.1}$$

where the information set $\mathcal{I} = \tilde{\omega}$ is the order flow. In this static case, the information set available to the specialist is given by the aggregate orders in the market, the insider's market orders plus noise trading, the order flow $\tilde{\omega} = \tilde{x} + \tilde{z}$. Equation (3.1) follows an implicit preference for early resolution of uncertainty as in Kreps and Porteus (1978) and Epstein and Zin (1989).[3] The specialist's profits equivalently represented using centered variables, $\tilde{\nu} = \tilde{\omega}\left(p - \tilde{f}\right)$. Then, given that $\mathbb{E}\left[\tilde{\omega}\right] = 0$, the objective function in (3.1) becomes $-\frac{\varrho\tilde{\omega}^2}{2}\mathbb{E}\left[\left(\tilde{f} - p\right)^2|\tilde{\omega}\right]$. Hence, the specialist sets the prices to maximize her expected utility,

$$p(\tilde{\omega}) = \arg\min_p \mathbb{E}\left[\left(\tilde{f} - p\right)^2|\tilde{\omega}\right].$$

Then, the pricing rule $p = \mathbb{E}\left[\tilde{f}|\tilde{\omega}\right]$ solves the previous problem.

3.2.3 Learning

This subsection depicts how the specialist's prices respond to the order flow as they can filter out noise trading. The specialist allocates her capacity to learn about the true payoffs from the order flow, in this case to filter out the variance introduced by the noise

[2]See Subrahmanyam (1991).

[3]See van Nieuwerburgh and Veldkamp (2010) section 1.2 for further details of this specification of CARA expected utility.

traders. Intuitively, this can be seen as feedback provided by the specialist to the market in order to reduce the posterior variance of noise trading. This is equivalent to assuming that the specialist is endogenizing the noise-trading variance. Since there is a one-to-one correspondence between the asset payoff and the market order, the order flow can be transformed into the signal

$$\tilde{s} = \tilde{f} + \tilde{\zeta},$$

where $\tilde{\zeta}$ is a normal random variable that represents the normalization of the noise introduced by the random liquidity traders.[4]

I assume that the specialist can acquire κ units of information at a marginal cost λ. The information is measured in units of entropy-based uncertainty reduction as proposed by Shannon (1948). The mutual information determines the amount of information that a random variable contains about another variable. In our case, it gives the amount of information that the signal contains about the asset's fundamental value. This function indicates the uncertainty that can be resolved regarding the asset payoffs when the specialist observes the order flow. Let the entropy function $H\left(\tilde{f}\right) = -\mathbb{E}\left[\ln p\left(\tilde{f}\right)\right]$ quantify the information required to fully resolve the uncertainty on the variable $\tilde{f}$. After observing a signal $\tilde{s}$, the remaining unresolved uncertainty left is given by the conditional entropy function $H\left(\tilde{f}|\tilde{s}\right) = -\mathbb{E}\left[\ln p\left(\tilde{f}\right)|\tilde{s}\right]$. Moreover, the difference between the unconditional and conditional entropy gives the total resolved uncertainty achieved when the signal is observed,

$$I\left(\tilde{f};\tilde{s}\right) = H\left(\tilde{f}\right) - H\left(\tilde{f}|\tilde{s}\right).$$

The acquired units of information pose an upper bound for the mutual information $I\left(\tilde{f};\tilde{s}\right) \leq \kappa$. This upper limit is referred to as the capacity constraint and, in this problem, it binds. Furthermore, when $\tilde{f}$ and $\tilde{s}$ are normally distributed, the capacity constraint becomes

$$\kappa = \frac{1}{2}\ln\left(\frac{\sigma_F^2}{\sigma_{F|s}^2}\right),$$

where $\sigma_{F|s}^2 = \frac{\sigma_F^2\sigma_\epsilon^2}{\sigma_F^2+\sigma_\epsilon^2}$ is the conditional variance of $\tilde{f}$ after signal $\tilde{s}$ is observed.

[4]Due to the linearity of the market order, the transformed signal $\tilde{s}$ is such that $\tilde{f} = x^{-1}\left(\tilde{\omega} - \tilde{z}\right)$ is also linear. Therefore, the signal is obtained from $\tilde{f} = \tilde{s} - \tilde{\zeta}$.

Recall that Bayesian updating under normality causes a linear dependence of the conditional expectation $\mathbb{E}\left[\tilde{f}|\tilde{s}\right]$ on $\tilde{f}$. Hence, the price is given by

$$\mathbb{E}\left[\tilde{f}|\tilde{s}\right] = (1-\delta)\underbrace{\bar{f}}_{=0} + \delta\tilde{s},$$

where the weight $\delta \equiv \left(1 - \frac{\sigma^2_{F|s}}{\sigma^2_F}\right) \in [0,1]$ represents the attention level assigned to the signal. Note that the conditional expectation represents a weighted average between the unconditional expected value of the payoff and the signal. Therefore, the price chosen by the specialist is

$$\tilde{p} = \delta\tilde{f} + \tilde{u}, \tag{3.2}$$

where δ is the attention value and $\tilde{u}$ is a zero-mean Gaussian noise generated by liquidity traders that cannot be filtered out. A perfectly informed specialist sets the price $\tilde{p} = \tilde{f}$. On the contrary, a fully uninformed specialist sets a price that does not reflect the true behavior of the asset payoff.

The specialist acquires κ units of information at a unit cost λ in order to reduce the uncertainty resulting from noise trading. The objective function for the attention allocation problem is

$$\min_{\sigma^2_{F|s}} \left\{ \mathbb{E}\left[\left(\tilde{f} - \mathbb{E}\left[\tilde{f}|\tilde{s}\right]\right)^2 \middle| \tilde{s}\right] + \lambda\kappa \right\}$$
$$= \min_{\sigma^2_{F|s}} \sigma^2_{F|s} + \frac{\lambda}{2}\ln\frac{\sigma^2_F}{\sigma^2_{F|s}}.$$

Alternatively, the previous maximization problem can be written as

$$\max_{\delta\in[0,1]} \left\{ (1-\delta)\,\sigma^2_F - \frac{\lambda}{2}\ln\frac{1}{1-\delta} \right\}.$$

Therefore, the optimal attention level is given by

$$\delta = \max\left\{0, 1 - \frac{\lambda}{2\sigma^2_F}\right\}.$$

Note that if the cost of information is high enough, the specialist pays no attention to the uncertainty reduction and the price depends only on the unfiltered order flow.[5]

[5] An uninformed specialist chooses her information structure that minimizes the loss function,

$$\mathbb{E}\left[\left(\tilde{F} - \mathbb{E}\left[\tilde{F}|\tilde{\omega}\right]\right)^2\right].$$

Which is equivalent to the problem stated above.

However, for lower values of the unit cost λ, the specialist acquires information about noise trading, which makes δ higher. This implies that the market becomes more liquid as the uncertainty about market orders is reduced.

3.3 Dynamic model

Consider now a model where the specialist repeatedly trades in the market. This setup extends the static model presented in the previous section to a market where asset payoffs follow different stochastic processes. There are two main differences between the model proposed in this section and the static model above. First, as the specialists repeatedly trade in the market, their information set is composed of the observed order flow and the history of asset payoffs, which are common knowledge. Second, endogenous information acquisition in a dynamic environment is not modeled hereby. Instead, the results for different exogenous capacity levels are compared.

In the dynamic version of this model, the same types of investors interact. In each period, a perfectly informed insider places his market order $\tilde{x}_t = x\left(\tilde{f}_t\right)$ to the specialist. There is a different insider trading each period, so the market order is a time invariant function of the asset payoff. Note that the market order does not depend on the history of $\tilde{f}_t$, because the insider does not have any uncertainty to resolve in period t. Noise traders place independent random liquidity demands $\tilde{z}_t \sim N\left(0, \sigma_z^2\right)$ in each period. The specialist observes the order flow as well as all past realizations of the asset payoff and chooses to learn about noise trading to increase the precision of the order flow.

I follow Maćkowiak *et al.* (2018) to characterize a pricing process when asset payoffs follow any ARMA(p, q) process. Then, I compare the results with Sims' (2006) approach for AR(1) processes. In section 3.2, I use the optimal pricing processes to simulate the effects on price dynamics of structural and noisy shocks through impulse response functions. For the simulation and IRF's, I use the DRIPs toolbox by Afrouzi and Yang (2021).

3.3.1 Asset prices with ARMA(p,q) payoffs

This section provides a broad analytical framework for the specialist's optimal pricing function given a processing capacity. As a result, I find the order of the optimal stochastic process for the prices.

The specialist chooses the stochastic properties of the signals in period $t=0$. Note that the information available to the specialist in each period is not only the current order flow, but all the history up to this period. Therefore, she should construct the signals for each period, since she chooses from the information set in period zero. Even if the specialist now faces a dynamic problem, she still exhibits the same instantaneous preferences as in the static model. Thus, the loss function for the specialist becomes,

$$\mathbb{E}\left[\sum_{t=1}^{\infty}\beta^t\left(\tilde{f}_t-\tilde{p}_t\right)^2\bigg|\mathcal{I}_t\right]$$
$$=\frac{\beta}{1-\beta}\mathbb{E}\left[\left(\tilde{f}_t-\tilde{p}_t\right)^2\bigg|\mathcal{I}_t\right],\ \beta\in(0,1),$$

which is a monotone transformation of the problem in the previous section. Hence, the pricing function $\tilde{p}_t$ is time invariant, and equivalent to the original dynamic rational inattention model proposed by Sims (2003). The pricing that minimizes the loss function is $\tilde{p}_t=\mathbb{E}\left[\tilde{f}_t|\mathcal{I}_t\right]$.

Consider now that $\tilde{f}_t$ follows an ARMA(p,q) process

$$\tilde{f}_t=\sum_{s=1}^{p}\phi_s\tilde{f}_{t-s}+\sum_{j=0}^{q}\theta_j\tilde{\varepsilon}_{t-j},$$

where $\tilde{\varepsilon}_t\sim N\left(0,\sigma_{\varepsilon}^2\right)$ and has zero autocorrelation. The optimal signal is one dimensional and composed of linear combinations of the p lags of the asset payoffs $\tilde{f}_t$ and the q lags of the exogenous shocks $\tilde{\varepsilon}_t$.[6] The intuition behind the optimality of this structure for the signal is that the inclusion of additional variables cannot increase the objective function but it can reduce the information flow. Similarly, if there is more than one signal that contains different linear combinations of the history of $\tilde{f}_t$ and its shocks $\tilde{\varepsilon}_t$, there is always a one dimensional signal that contains the same information with higher precision.

[6]See Maćkowiak *et al.* (2018) for the proof.

Also recall that any ARMA(p,q) process has a VAR(1) representation,

$$\tilde{X}_{t+1} = \Phi \tilde{X}_t + \tilde{\nu}_{t+1},$$

where $\tilde{X}_t$ is the following vector

$$\tilde{X}_t = \begin{cases} \tilde{f}_t & \text{if } p = 0, q = 0 \\ \left(\tilde{f}_t, \dots, \tilde{f}_{t-p+1}\right)' & \text{if } p \geq 1, q = 0 \\ \left(\tilde{f}_t, \tilde{\varepsilon}_t, \dots, \tilde{\varepsilon}_{t-q+1}\right)' & \text{if } p = 0, q \geq 1 \\ \left(\tilde{f}_t, \dots, \tilde{f}_{t-p+1}, \tilde{\varepsilon}_t, \dots, \tilde{\varepsilon}_{t-q+1}\right)' & \text{if } p \geq 1, q \geq 1 \end{cases},$$

with covariance matrix $\Sigma_{X,t}$, and $\tilde{\nu}_t$ is a vector containing white noise elements with covariance matrix Σ_ν, and Φ is a squared matrix of coefficients. Moreover, the optimal signal should have a state space representation of the form

$$\tilde{X}_{t+1} = \Phi \tilde{X}_t + \tilde{\nu}_{t+1},$$
$$\tilde{S}_t = \eta' \tilde{X}_t + \tilde{\psi}_t,$$

where η is the $p+q$ vector of optimal weights on the signal and $\tilde{\psi}_t$ corresponds to the noise introduced each period by noise trading, and which is uncorrelated across periods and takes the form of Gaussian white noise with variance σ_ψ^2. The specialist chooses the weights η such that the information flow constraint is satisfied.

The specialist's objective can be interpreted as the choice of an information structure that reduces her uncertainty. This problem is equivalent to the selection of the weights η in the Kalman filter equations

$$\Sigma_{X,t+1|t} = \Phi \Sigma_{X,t|t} \Phi' + \Sigma_\nu,$$
$$\Sigma_{X,t|t} = \Sigma_{X,t|t-1} - \Sigma_{X,t|t-1} \eta \left(\eta' \Sigma_{X,t|t-1} \eta + \sigma_\psi^2\right)^{-1} \eta \Sigma_{X,t|t-1}.$$

By stationarity of $\tilde{F}_t$, $\Sigma_{X,1} \equiv \lim_{t\to\infty} \Sigma_{X,t|t-1}$ and $\Sigma_{X,0} \equiv \lim_{t\to\infty} \Sigma_{X,t|t}$ are given by

$$\Sigma_{X,1} = \Phi \Sigma_{X,0} \Phi' + \Sigma_\nu,$$
$$\Sigma_{X,0} = \Sigma_{X,1} - \Sigma_{X,1} \eta \left(\eta' \Sigma_{X,1} \eta + \sigma_\psi^2\right)^{-1} \eta' \Sigma_{X,1}.$$

Note that the specialist's objective function corresponds to the first element in the matrix $\Sigma_{X,0}$ because it coincides with the mean squared error for each period.

The capacity constraint is the upper limit of the information flow,

$$\lim_{T\to\infty} \frac{1}{T} I\left(F_0, \tilde{F}_1, \ldots, \tilde{F}_T; \tilde{S}_1, \ldots, \tilde{S}_T\right) \leq \kappa,$$

and is reduced to

$$\frac{1}{2}\ln\left(\frac{\eta'\Sigma_1\eta}{\sigma_\psi^2} + 1\right) \leq \kappa,$$

given the previous Kalman filter equations. That is, the information flow depends only on the signal to noise ratio. The ratio is constant in time due to the stationarity of $\tilde{f}_t$.

The objective function for the specialist can be rewritten as

$$\min_{\eta\,\in\mathbb{R}^{p+q},\sigma_\psi^2} \left\{(1,0,\ldots,0)\,\Sigma_{X,0}\,(1,0,\ldots,0)'\right\}$$

such that

$$\frac{1}{2}\ln\left(\frac{\eta'\Sigma_1\eta}{\sigma_\psi^2} + 1\right) \leq \kappa.$$

Since the information flow constraint is always binding,

$$\frac{\eta'\Sigma_1\eta}{\sigma_\psi^2} = e^{2\kappa} - 1.$$

For a given vector of weights, the variance of ψ is

$$\sigma_\psi^2 = \frac{\eta'\Sigma_1\eta}{e^{2\kappa} - 1}.$$

As a result, the matrix

$$\Sigma_{X,0} = \Sigma_{X,1} - \frac{1 - e^{-2\kappa}}{\eta'\Sigma_{X,1}\eta}\Sigma_{X,1}\eta\eta'\Sigma_{X,1}.$$

Additionally, one extra identification assumption is required to solve the problem. The vector of weights η can be normalized by one element to find a unique solution. For example, to set the first element to be one.

Therefore, the specialist solves the following problem

$$\min_{\eta\,\in\mathbb{R}^{p+q}} \left\{(1,0,\ldots,0)\,\Sigma_{X,0}\,(1,0,\ldots,0)'\right\},$$

where

$$\Sigma_{X,1} = \Phi\Sigma_{X,0}\Phi' + \Sigma_\nu,$$

$$\Sigma_{X,0} = \Sigma_{X,1} - \frac{1 - e^{-2\kappa}}{\eta'\Sigma_{X,1}\eta}\Sigma_{X,1}\eta\eta'\Sigma_{X,1}.$$

Then, the optimal pricing rule given a capacity κ is

$$\tilde{p}_t = \mathbb{E}\left[\tilde{f}_t|\tilde{S}_t\right].$$

Note that the optimal signal chosen by the specialist considers only the current period order flow and replicates the order of the generating process for the asset payoffs. In general, if $\tilde{f}_t$ follows an ARMA(p, q) process, the optimal signal is

$$\tilde{S}_t = \sum_{s=1}^{p} \eta_s \tilde{f}_{t-s+1} + \sum_{j=1}^{q} \eta_{p+j}\tilde{\varepsilon}_{t-j+1} + \tilde{\psi}_t.$$

As the capacity increases, the vector of optimal weights η converges to $(1, 0, \ldots, 0)$.[7] This result implies that a specialist with infinite information flow processing capacity can filter out all the liquidity traders' noise. Then, as she filters out the noise trading, the specialist does not need to learn from the asset payoffs process, since she learns perfectly about the asset payoffs from the insider's market orders. Alternatively, if she observes the asset payoffs perfectly, the problem becomes static. The asset price is set to match the asset payoff perfectly, $\tilde{p}_t = \tilde{f}_t$. The particular cases of different generating processes are considered below.

3.3.1.1 AR(1)

Consider the example in the previous subsection where the asset payoffs follow an AR(1) process, $\tilde{f}_t = \phi_1\tilde{f}_{t-1} + \theta_0\tilde{\varepsilon}_t$. Then, the optimal signal that the specialist can extract is $\tilde{S}_t = \tilde{f}_t + \tilde{\psi}_t$. The corresponding values for the state space representation are, $\tilde{X}_t = \tilde{f}_t$, $\Phi = \phi_1$, $\Sigma_\nu = \theta_0^2$, $\Sigma_{X,1} = \phi_1^2\Sigma_{X,0} + \theta_0^2$ and $\Sigma_{X,0} = e^{-2\kappa}\Sigma_{X,1}$, all scalars. The resulting value for the objective function is

$$\Sigma_{X,0} = \frac{\theta_0^2}{e^{2\kappa} - \phi_1^2}.$$

The precision of the optimal signal is given by

$$\sigma_\psi^{-2} = \left(\frac{e^{2\kappa} - 1}{e^{2\kappa}}\right)\left(\frac{e^{2\kappa} - \phi_1^2}{\theta_0^2}\right).$$

The behavior of the pricing function with respect to the original process is shown in Figure 3.1. The figure displays the simulated values of an AR(1) process for asset payoffs

[7]See Maćkowiak *et al.* (2018) Proposition 6 for proof.

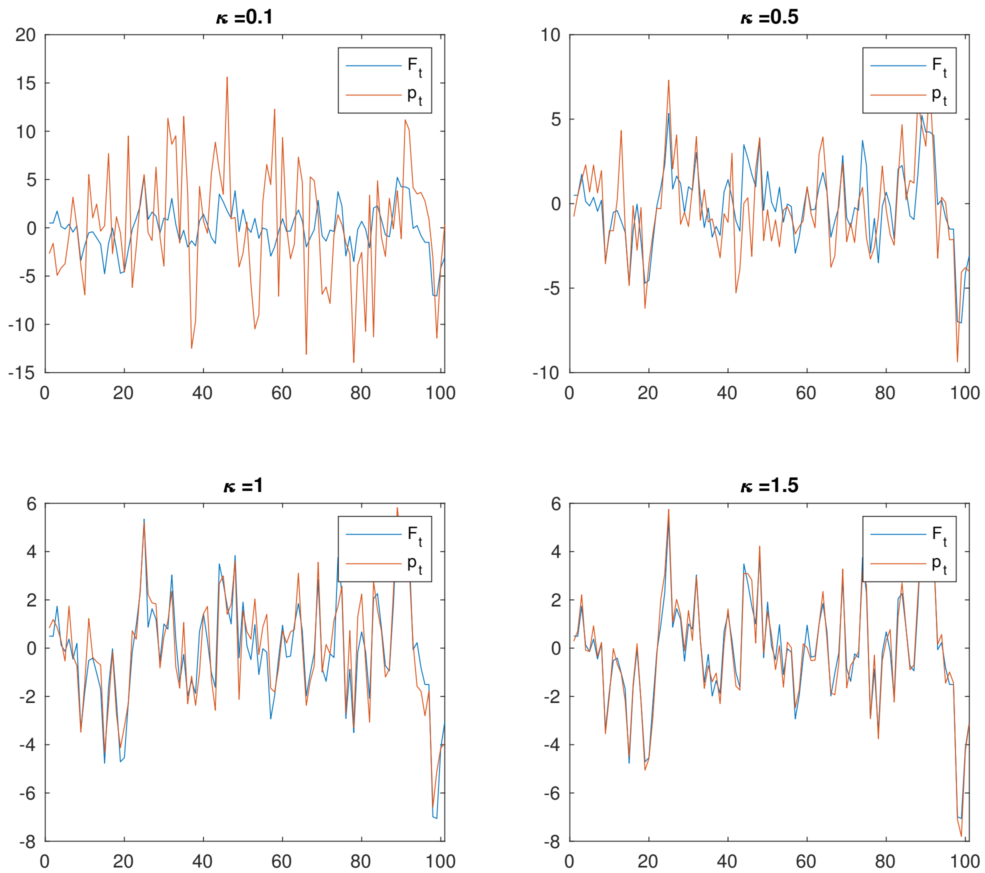

Figure 3.1: Asset payoffs and prices for different levels of information flow

and plots the optimal pricing rule as the information processing capacity increases. The following section presents the price dynamics from an impulse-response function perspective. Note that if the asset payoffs follow an AR(1) process, the complexity of the problem does not increase with respect to a static case in the previous section. The specialist only uses what she learns from the order flow in each period plus the history of asset payoffs to construct the pricing process.

Alternative approach

In this alternative approach, I replicate the example in Sims (2003), who characterizes a solution for an AR(1) problem. Now, we will determine the pricing process when the asset payoffs follow an AR(1) process. This method considers the invertibility property between stationary AR(1) processes and MA(∞) processes. Sims' method gives equivalent results for the specialist's pricing choice. Sims' method is useful for any MA process, but fails to produce analytical results for other asset payoffs processes.

First, consider a dynamic setup where the asset payoffs follow a moving average process

$$\tilde{f}_t = \sum_{s=0}^{t} a_s \tilde{\varepsilon}_s, \tag{3.3}$$

where $\tilde{\varepsilon}_s \sim N(0,1)$ The process has a first order autoregressive representation,

$$\tilde{f}_t = \rho \tilde{f}_{t-1} + a_0 \tilde{\varepsilon}_t,$$

where $a_s = \rho^s a_0$. There is a perfectly informed investor who places a profit maximizing market order each period

$$\tilde{x}_t = x\left(\tilde{f}_t\right).$$

Since the investor is perfectly informed, the agent faces a static problem each period. The market order is a time invariant function of the asset payoffs and, as above, the market order is a sufficient statistic about the asset payoff in period t. Additionally, there are noise traders who place iid random liquidity orders each period, $\tilde{z}_t \sim N(0, \sigma_z^2)$. The order flow for each period is $\tilde{\omega}_t = \tilde{x}_t + \tilde{z}_t$.

The specialist wants to set a price that mimics the true process of the asset payoffs, the objective function being

$$\mathbb{E}\left[-\sum_{t=1}^{\infty} \beta^t \left(\tilde{f}_t - \tilde{p}_t\right)^2\right].$$

She has access to the order flow for each period, plus the history of realizations of the asset payoffs up to period $t-1$. Since the insider's response is time invariant, the order flow can be transformed into a signal of the form

$$\tilde{S}_t = \tilde{f}_t + \gamma \tilde{\upsilon}_t.$$

The signal delivers the asset payoff plus the noise introduced by noise trading represented by $\tilde{\upsilon}_t$.

Now, consider the information that the price process contains about the asset payoff process. As in the static case, this mutual information is computed through the uncertainty reduction. Let $\tilde{f}^T = \left(\tilde{f}_1, \ldots, \tilde{f}_T\right)$ and $\tilde{p}^T = (\tilde{p}_1, \ldots, \tilde{p}_T)$ the vectors of the first

T elements of the process for the asset payoffs and asset prices respectively. Then, the mutual information is defined as

$$I\left(\tilde{f}^T;\tilde{p}^T\right) = H\left(\tilde{f}^T\right) - H\left(\tilde{f}^T|\tilde{p}^T\right).$$

The information flow between two stochastic processes is the average mutual information per period of the two processes $\left\{\tilde{f}_t\right\}_{t=0}^{\infty}$ and $\{\tilde{p}_t\}_{t=0}^{\infty}$. It is defined as

$$\mathbb{I}\left(\left\{\tilde{f}_t\right\}_{t=0}^{\infty};\{\tilde{p}_t\}_{t=0}^{\infty}\right) = \lim_{T\to\infty}\frac{1}{T}I\left(\tilde{f}^T;\tilde{p}^T\right),$$

the definition holds for both univariate and vector processes.[8] An imperfectly informed specialist faces a constraint that sets an upper bound κ to the information flow between the two variables.

Following the same process as in the static model, the pricing process that an imperfectly informed specialist sets is

$$\tilde{p}_t = \mathbb{E}\left[\tilde{f}_t|\mathcal{I}_t\right],$$

where the information set in each period is composed of any initial information plus all signals up to this period t,

$$\mathcal{I}_t = \mathcal{I}_0 \cup \{S_1,\ldots,S_t\}.$$

Since all variables are Gaussian the pricing process takes the form

$$\tilde{p}_t = \sum_{s=0}^{t}\alpha_s\tilde{\varepsilon}_{t-s} + \sum_{s=0}^{t}c_s\tilde{\upsilon}_{t-s}.$$

The informationally-constrained specialist's problem, consists of the selection of weights α_s and c_s, to minimize the objective function.[9] If there is no informational constraint, the specialist cannot perfectly filter the asset payoffs from the signals, and therefore she sets $\alpha_s = a_s$ and $c_s = 0$ for all s. However, the existence of the constraint on the information

[8] If the two processes were Gaussian with zero autocorrelation, the information flow becomes

$$\mathbb{I}\left(\left\{\tilde{F}_t\right\};\{\tilde{p}_t\}\right) = \frac{1}{2}\ln\left(\frac{\sigma_F^2\sigma_p^2}{\sigma_F^2\sigma_p^2 - Cov\left(\tilde{F}_t,\tilde{p}_t\right)^2}\right).$$

[9] Sims (2003) provides a numerical approach to this problem by changing the information flow from the time to the frequency domain. Thus, the information processing constraint becomes

$$\mathbb{I} = -\frac{1}{2}\int_{-\pi}^{\pi}\ln\left(1-\frac{1}{1+|\hat{c}|/|\hat{\alpha}|}\right),$$

where $\hat{c}$ and $\hat{\alpha}$ correspond to the Fourier transforms of the corresponding parameters.

flow, sets an upper bound on the specialist's objective function. The coefficients that solve this problem are

$$\alpha_s = \left[\rho^s - \frac{1}{e^{2\kappa}}\left(\frac{\rho}{e^{2\kappa}}\right)^s\right] a_0$$
$$c_s = \left(\frac{e^{2\kappa}-1}{e^{2\kappa}\left(e^{2\kappa}-\rho^2\right)}\right)^{\frac{1}{2}} a_0.$$

The pricing process set by a perfectly informed specialist results from taking the limit $\kappa \to \infty$ in the equations for the coefficients. As a result, the pricing process tracks perfectly the asset payoffs, $\alpha_s = a_s = \rho^s a_0$. This result provides some insight into how the processing capacity affects the response of prices to fundamental shocks. There is an alternative approach to solve this problem, which is equivalent. The specialist chooses an optimal weight γ on the signals, such that the information constraint binds. Then, she updates her beliefs according to such a signal and sets a price.

3.3.1.2 AR(p)

In general, when the asset payoffs follow an AR(p) process $\tilde{f}_t = \sum_{s=1}^{p} \phi_s \tilde{f}_{t-p} + \theta_0 \tilde{\varepsilon}_t$.

$$\Sigma_{X,1} = \Phi\left[\Sigma_{X,1} - \frac{1-e^{-2\kappa}}{\eta'\Sigma_{X,1}\eta}\Sigma_{X,1}\eta\eta'\Sigma_{X,1}\right]\Phi' + \Sigma_\nu,$$

where

$$\Sigma_\nu = \begin{bmatrix} \theta_0^2 & 0 & \cdots & 0 \\ 0 & 0 & & 0 \\ \vdots & & \ddots & \\ 0 & 0 & \cdots & 0 \end{bmatrix}.$$

See Figure 3.2 for matrix $\Phi\Sigma_{X,1}\eta\eta'\Sigma_{X,1}\Phi$.

The optimal signal is given by $S_t' = \vartheta\tilde{f}_t + (1-\vartheta)\left(\phi(L)\tilde{f}_t\right) + \varrho\psi_t$, where ϑ is the Kalman gain. It is a function of the capacity $\kappa$and parameters $\phi_1, \ldots, \phi_p$ and θ_0. Such a gain is an increasing function of κ. A larger learning capacity gives greater weight to the contemporary information making the pricing process more precise.

Suppose for example that the asset payoffs follow an AR(2) process, $\tilde{f}_t = \phi_1\tilde{f}_{t-1} + \phi_2\tilde{f}_{t-2} + \theta_0\tilde{\varepsilon}_t$. The optimal signal is given by $\tilde{S}_t = \tilde{f}_t + \eta_2\tilde{f}_{t-1} + \tilde{\psi}_t$. Note that, in

$$\begin{bmatrix} \sum_{i=1}^{p}\sum_{s=1}^{p}\phi_i\phi_s\left(\sum_{j=1}^{p}\eta_j\sigma_{1,ij}\right)\left(\sum_{j=1}^{p}\eta_j\sigma_{1,sj}\right) & \left(\sum_{i=1}^{p}\eta_i\sigma_{1,1i}\right)\left(\sum_{s=1}^{p}\phi_s\left(\sum_{i=1}^{p}\eta_i\sigma_{1,si}\right)\right) & \cdots & \left(\sum_{i=1}^{p}\eta_i\sigma_{1,(p-1)i}\right)\left(\sum_{s=1}^{p}\phi_s\left(\sum_{i=1}^{p}\eta_i\sigma_{1,si}\right)\right) \\ \left(\sum_{i=1}^{p}\eta_i\sigma_{1,1i}\right)\left(\sum_{s=1}^{p}\phi_s\left(\sum_{i=1}^{p}\eta_i\sigma_{1,si}\right)\right) & \left(\sum_{i=1}^{p}\eta_i\sigma_{1,1i}\right)^2 & \cdots & \left(\sum_{i=1}^{p}\eta_i\sigma_{1,1i}\right)\left(\sum_{i=1}^{p}\eta_i\sigma_{1,(p-1)i}\right) \\ \left(\sum_{i=1}^{p}\eta_i\sigma_{1,2i}\right)\left(\sum_{s=1}^{p}\phi_s\left(\sum_{i=1}^{p}\eta_i\sigma_{1,si}\right)\right) & \left(\sum_{i=1}^{p}\eta_i\sigma_{1,1i}\right)\left(\sum_{i=1}^{p}\eta_i\sigma_{1,2i}\right) & \cdots & \left(\sum_{i=1}^{p}\eta_i\sigma_{1,2i}\right)\left(\sum_{i=1}^{p}\eta_i\sigma_{1,(p-1)i}\right) \\ \vdots & \vdots & \ddots & \vdots \\ \left(\sum_{i=1}^{p}\eta_i\sigma_{1,(p-1)i}\right)\left(\sum_{s=1}^{p}\phi_s\left(\sum_{i=1}^{p}\eta_i\sigma_{1,si}\right)\right) & \left(\sum_{i=1}^{p}\eta_i\sigma_{1,1i}\right)\left(\sum_{i=1}^{p}\eta_i\sigma_{1,(p-1)i}\right) & \cdots & \left(\sum_{i=1}^{p}\eta_i\sigma_{1,(p-1)i}\right)^2 \end{bmatrix}$$

Figure 3.2: Matrix $\Phi\Sigma_{X,1}\eta\eta'\Sigma_{X,1}\Phi$.

this case, the specialist choses to learn from the order flow plus additional information that the history of the asset payoffs provide. The structure of the optimal signal allows the specialist to use the information about the process that $\tilde{f}_t$ follows by only including a number of lags in the signal, equal to the order of the autoregressive process. This structure for the payoffs has the state space representation mentioned above where $\tilde{X}_t = \left(\tilde{f}_t, \tilde{f}_{t-1}\right)'$, the matrix

$$\Sigma_{X,1} = \Phi \left[\Sigma_{X,1} - \frac{1 - e^{-2\kappa}}{\eta' \Sigma_{X,1} \eta} \Sigma_{X,1} \eta \eta' \Sigma_{X,1} \right] \Phi' + \Sigma_\nu$$

and the matrix

$$\Sigma_\nu = \begin{bmatrix} \theta_0^2 & 0 \\ 0 & 0 \end{bmatrix},$$

the coefficient matrix

$$\Phi = \begin{bmatrix} \phi_1 & \phi_2 \\ 1 & 0 \end{bmatrix}.$$

and the variance of the noise term

$$\sigma_\psi^2 = \frac{\eta' \Sigma_{X,1} \eta}{e^{2\kappa} - 1},$$

where $\eta = (1, \eta_2)'$. Furthermore, note that due to the Bayesian updating process embedded here, the first element of the matrix $\Sigma_{X,0}$ is the element in position (2,2) in matrix $\Sigma_{X,1}$.

3.3.2 Price behavior and market resilience

As we have said, the ability of asset prices to return to their steady-state value after a large order is commonly referred to as resilience. A large order in this model may come from either an insider or a noise trader. Insiders place their orders according to their information about each period's innovation, while noise trading can be large with some nonzero probability. Naturally, asset prices should only react to the insider's orders if the specialist is perfectly informed. An imperfectly informed specialist should generate prices that recover faster from noise trading shocks than from insider shocks. This section presents the dynamic behavior of asset prices in response to new information through impulse response functions (IRFs).

In period t, the insider places an order according to the observed information, i.e., the realization of $\tilde{\varepsilon}_t$. For any asset payoff series $\tilde{f}_t$ that follows a stationary ARMA(p, q) process, the IRF illustrates the effect of a structural shock $\tilde{\varepsilon}_t$ on $\tilde{f}_t$. However, such a response is not directly reflected in price dynamics. The specialist's information set contains two ex-ante indistinguishable sources of disturbances, structural shocks $\tilde{\varepsilon}_t$, and noise trading, $\tilde{\psi}_t$. A specialist without information processing capacity, $\kappa = 0$, is not able to distinguish the source of the shock, so both shocks generate the same price dynamics. Thus, the response of prices to shocks either on $\tilde{\varepsilon}_t$ or $\tilde{\psi}_t$ is the same.

As the specialist starts to increase her capacity to process the information in the order flow, the price dynamics responds differently to structural and noise trading shocks. Two different IRFs describe the price response to a structural shock $\tilde{\varepsilon}_t$ and a noise trading shock on $\tilde{\psi}_t$. In this model, market resilience is measured by the persistence of noise trading shocks on the asset prices. Consequently, the response of asset prices to structural shocks mimics the asset payoffs' dynamics in a highly resilient market. Later in this section, the price dynamics under different stochastic processes for asset payoffs are analyzed. Figure 3.3 illustrates the price response when the asset payoffs follow an AR(1) process. The solid black dotted line depicts the IRF of the true asset payoff process, the black circled line is the response of prices to structural shocks $\tilde{\varepsilon}_t$, and the gray starred line is the IRF of the pricing process with respect to a noise trading shock $\tilde{\psi}_t$. Market resilience is represented by the quick recovery of the prices to noise trading, i.e., the third IRF drops quickly. Figure 3.3 shows the IRFs for different capacity levels, market resilience increases as the learning capacity expands. Moreover, the response of prices to structural shocks $\tilde{\varepsilon}_t$ approaches the asset payoff's response as the capacity grows.

Figure 3.4 presents the price response functions to shocks when the asset payoff follows an AR(2) process with high persistence. The first plot in the figure shows how, for low capacity levels, the order flow's informativeness is scarce. However, as time passes, the specialist can still learn from the history of the process, and the IRF of prices to structural shocks approaches the asset payoffs IRF. In contrast, the response to noise trading disappears as the second IRF approaches the true payoffs, i.e., the specialist learns the truth from history after some periods. Just as before, as the information processing

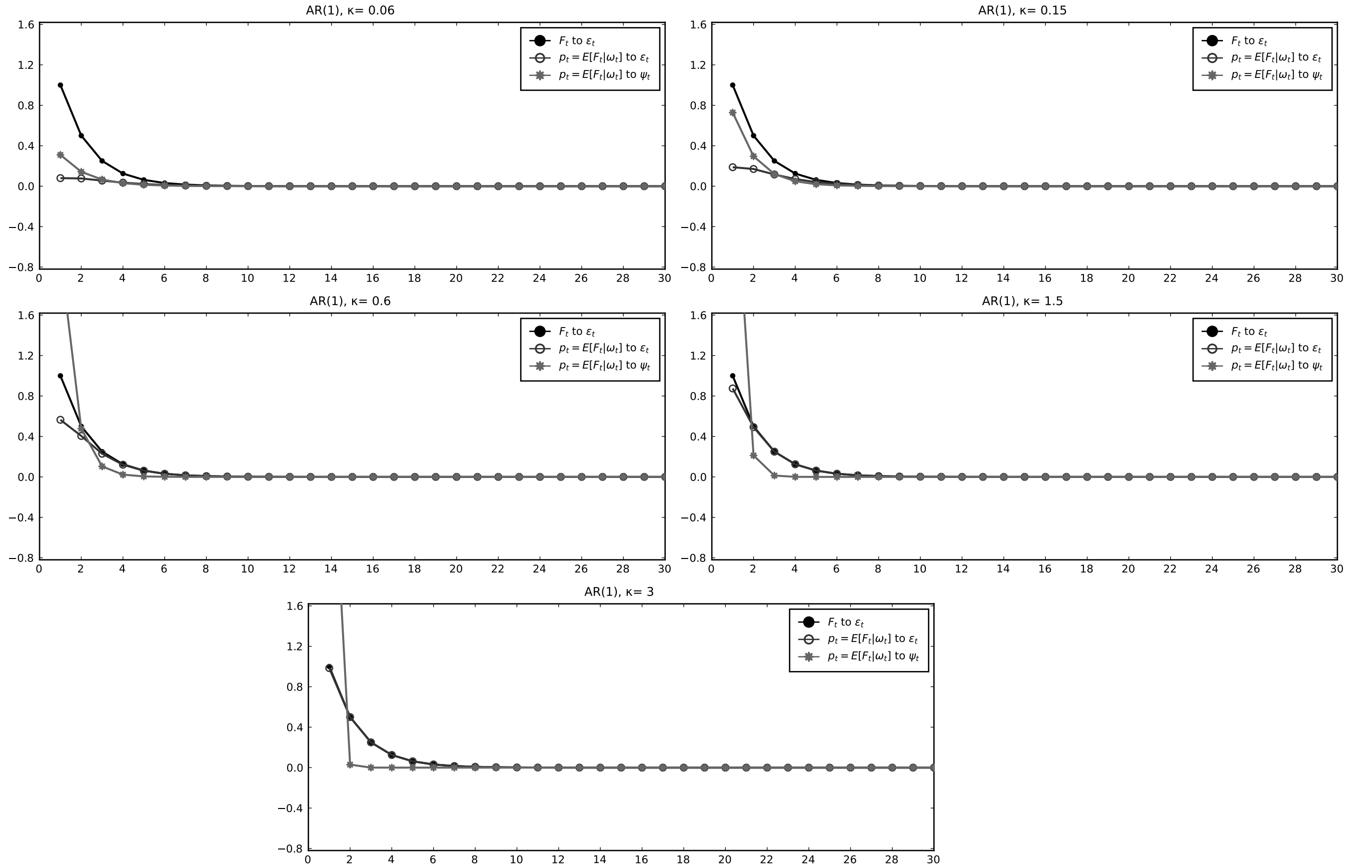

Figure 3.3: Impulse response function when the payoffs follow an AR(1) process.

capacity increases, prices respond faster to shocks and converge to the true asset payoffs process.

In Figures 3.5 and 3.6, it is assumed that asset payoffs follow an AR(3) process. Figure 3.5 displays the transition to a new steady-state after the shock. The contemporary response of prices to the shock is almost null for low processing capacity levels, and the prices slowly and smoothly converge to the new steady state. As the processing capacity increases, the contemporary response is now observable, and the market resilience increases.

When the asset payoffs follow a stationary process with imaginary roots, the prices overshoot before they return to the steady-state value. Figure 3.6 depicts an example of a process with imaginary roots. Such an overshooting effect disappears as the learning capacity increases whenever there is a noise trading shock.

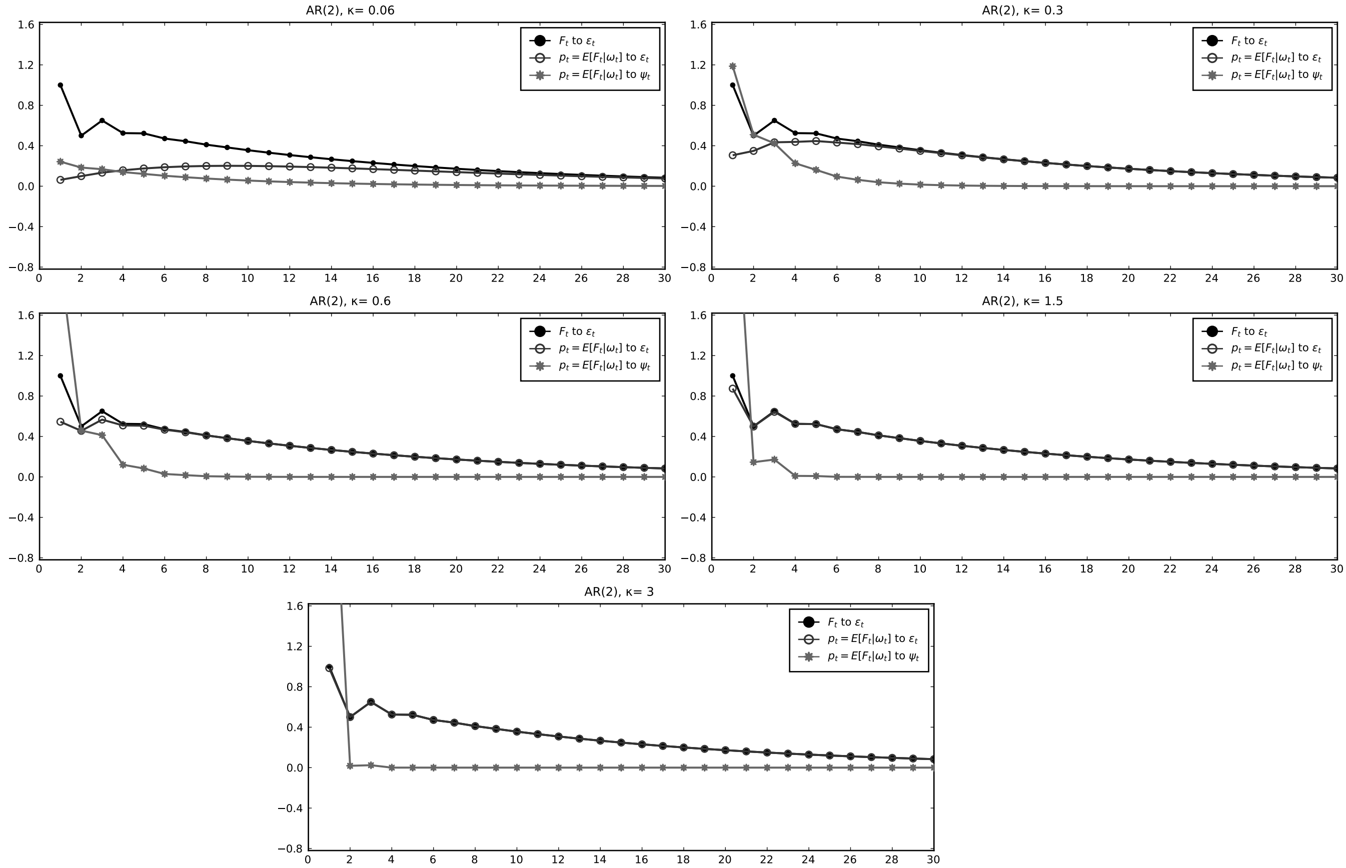

Figure 3.4: Impulse response function when payoffs follow an AR(2) process.

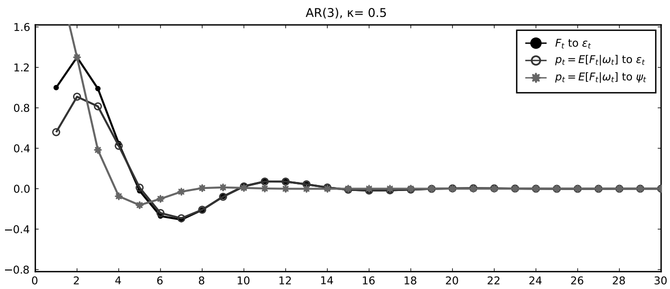

Figure 3.6: Impulse response when asset payoffs follow an AR(3) process with imaginary roots.

Finally, Figure 3.7 illustrates two additional possible examples of processes for asset payoffs. The first plot contains the IRFs for MA(1) asset payoffs. Due to the behavior of MA(q) processes, the IRF collapses to zero after $q$periods ahead in time. This case is not typical asset payoff behavior. In the second plot, the asset payoffs follow an ARMA(2,1) process. Compared to the previous examples, for a given learning capacity of $\kappa = 0.5$, the market resilience is lower, i.e., it takes more periods to dissolve noise trading shocks, as the specialist who learns about processes of greater complexity levels has to allocate

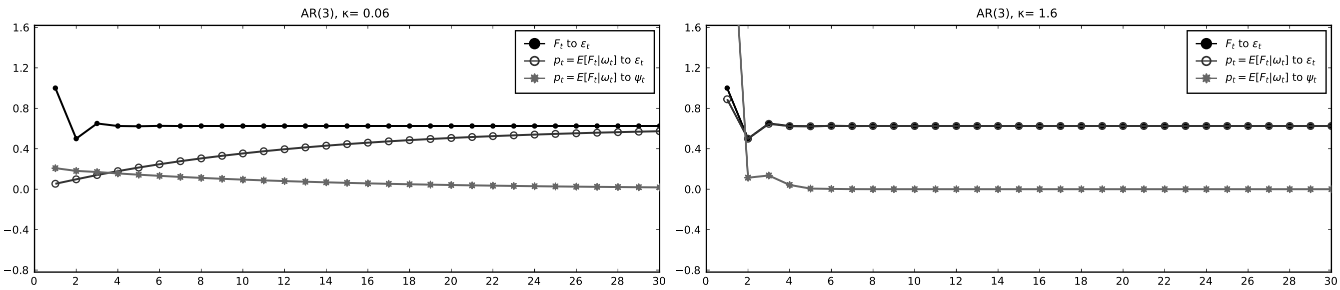

Figure 3.5: Impulse response function when asset payoffs follow an AR(3) process.

more capacity to identify noise trading.

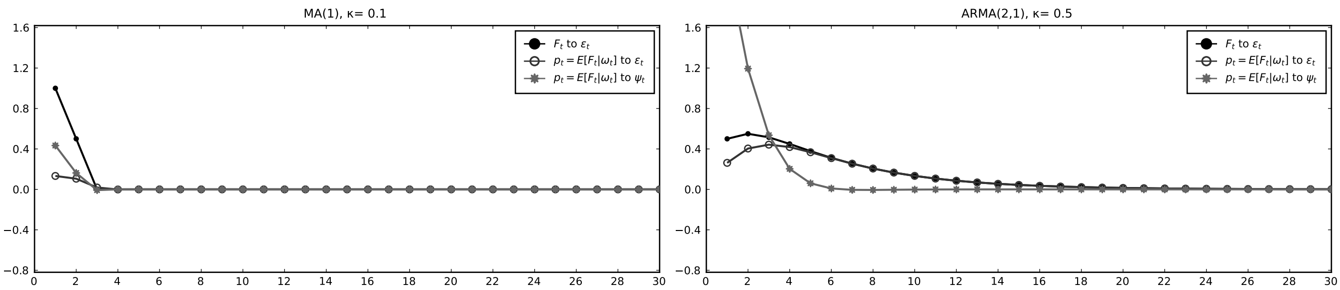

Figure 3.7: Impulse response function when payoffs follow an MA(1) and ARMA(2,1) process.

3.4 Concluding remarks

I have characterized the stochastic properties of the asset prices in a market where an insider trades along with noise traders and a specialist sets the asset prices. I have also illustrated the specialist's reaction to shocks that impact the order flow as her capacity to learn about the market changes. When a specialist observes a non-anticipated behavior of the order flow, she allocates her learning capacity to distinguish whether it comes from a structural shock -actual change in asset payoffs- or a noise trading shock. However, her learning capacity limits this process, and she allocates her attention to learning from the history of asset payoffs rather than the noisy information in the order flow. As her learning capacity increases, the specialist allocates her capacity to reduce the order flow noise. By doing so, she can distinguish between structural shocks and noise trading.

The specialist's optimal pricing process allows us to decompose the reaction of prices to the different types of shocks. As the learning capacity increases, noise trading shocks are quickly identified. Furthermore, with a higher learning capacity, asset prices only show persistence to structural shocks, following the same stochastic behavior as the asset payoffs. In contrast, the specialist has a delayed reaction to low processing capacities since she can only learn from the history of asset payoffs. As a result, noise trading shocks exhibit a high degree of persistence in terms of asset price.

This price behavior enables a direct interpretation of the ability of prices to recover after a shock or market resilience. To explain the recovery speed of asset prices after a shock to the order flow, I decompose the price reaction using an impulse response function for each possible shock. Therefore, a market is more resilient as the second IRF to a noise trading shock drops quickly. Equivalently, a market becomes more resilient as the IRF of the asset price to a structural shock approaches the IRF of the asset payoff to a structural shock.

I illustrate for different stationary ARMA processes how market resilience increases rapidly as the specialist's learning capacity increases. Note that market resilience is nothing but the result of the specialist's ability to resolve the adverse selection problem that insiders pose to her. Further extensions of this analysis can include insiders that trade repeatedly in the market and introducing several insiders, which allows for cross-sectional competition between investors.

Bibliography

Admati, A. R. (1985). A noisy rational expectations equilibrium for multi-asset securities markets. *Econometrica*, **53** (3), 629–657.

Afrouzi, H. and Yang, C. (2021). *Dynamic Rational Inattention and the Phillips Curve*. 8840, CESifo.

Blackwell, D. (1951). Comparison of experiments. In *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley, Calif.: University of California Press, pp. 93–102.

— (1953). Equivalent comparisons of experiments. *The Annals of Mathematical Statistics*, **24** (2), 265–272.

Caballé, J. and Krishnan, M. (1994). Imperfect competition in a multi-security market with risk neutrality. *Econometrica*, **62** (4), 695–704.

de Oliveira, H. (2013). *Axiomatic Foundations for Entropic Costs of Attention*. Mimeo, Northwestern University.

— (2018). Blackwell's informativeness theorem using diagrams. *Games and Economic Behavior*, **109** (C), 126–131.

Epstein, L. G. and Zin, S. E. (1989). Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework. *Econometrica*, **57** (4), 937–969.

Garbade, K. D. (1982). *Securities Markets*. McGraw Hill Series in finance.

Glosten, L. R. and Milgrom, P. (1985). Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics*, **14** (1), 71–100.

Jung, J., ho Kim, J., Matejka, F., Christopher and Sims, A. (2015). Discrete actions in information-constrained decision problems.

Kreps, D. M. and Porteus, E. L. (1978). Temporal resolution of uncertainty and dynamic choice theory. *Econometrica*, **46** (1), 185–200.

Krishnan, M. (1992). An equivalence between the kyle (1985) and the glosten–milgrom (1985) models. *Economics Letters*, **40** (3), 333–338.

Kyle, A. S. (1985). Continuous auctions and insider trading. *Econometrica*, **53** (6), 1315–1335.

Maćkowiak, B., Matejka, F. and Wiederholt, M. (2018). Dynamic rational inattention: Analytical results. *Journal of Economic Theory*, **176** (C), 650–692.

Marschak, J. and Miyasawa, K. (1968). Economic comparability of information systems. *International Economic Review*, **9** (2), 137–174.

Matějka, F. and McKay, A. (2015). Rational inattention to discrete choices: A new foundation for the multinomial logit model. *American Economic Review*, **105** (1), 272–98.

Mondria, J. (2010). Portfolio choice, attention allocation, and price comovement. *Journal of Economic Theory*, **145** (5), 1837–1864.

Obizhaeva, A. A. and Wang, J. (2013). Optimal trading strategy and supply/demand dynamics. *Journal of Financial Markets*, **16** (1), 1–32.

Peng, L. (2005). Learning with information capacity constraints. *Journal of Financial and Quantitative Analysis*, **40** (2), 307–329.

— and Xiong, W. (2006). Investor attention, overconfidence and category learning. *Journal of Financial Economics*, **80** (3), 563–602.

Rényi, A. (1961). On measures of information and entropy. *Proceedings of the fourth Berkeley Symposium on Mathematics, Statistics and Probability*, pp. 547–561.

Shannon, C. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, **27**, 379–423, 623–656.

Sims, C. A. (1998). Stickiness. *Carnegie-Rochester Conference Series on Public Policy*, **49** (1), 317–356.

— (2003). Implications of rational inattention. *Journal of Monetary Economics*, **50** (3), 665–690, swiss National Bank/Study Center Gerzensee Conference on Monetary Policy under Incomplete Information.

Subrahmanyam, A. (1991). Risk aversion, market liquidity, and price efficiency. *The Review of Financial Studies*, **4** (3), 417–441.

van Nieuwerburgh, S. and Veldkamp, L. (2009). Information immobility and the home bias puzzle. *Journal of Finance*, **64** (3), 1187–1215.

— and — (2010). Information acquisition and under-diversification. *The Review of Economic Studies*, **77** (2), 779–805.

Vitale, P. (2012). Risk-averse insider trading in multi-asset sequential auction markets. *Economics Letters*, **117** (3), 673–675.

Wiederholt, M. (2010). rational inattention. In S. N. Durlauf and L. E. Blume (eds.), *The New Palgrave Dictionary of Economics*, Basingstoke: Palgrave Macmillan.