Information Choice, Shock Transmission and Contagion *

Adelina Barbalau†

January 3, 2019

Abstract

During the 2007-2008 financial crisis, countries that were relatively more exposed to the crisis epicenter, the United States, were among the least affected. This counters the intuition that the impact of a shock increases with exposure to it, and raises the question of the mechanism through which the impact of shocks can decrease with exposure. I propose a model in which decision-makers learn about the risk factors they are exposed to, but have limited capacity to process information. I find that decision-makers optimally choose to learn more about the risk factors they are more exposed to, and this informational advantage mitigates the negative consequences of shocks by enabling them to take better investment decisions. Relative to an exogenous information benchmark, the endogenous information model I propose predicts that shocks to risk factors that decision-makers are relatively less exposed to are amplified. By the same token, shocks to risk factors that decision-makers are relatively more exposed to are attenuated.

Keywords: Rational Inattention, Corporate Investment, Shock Amplification
JEL Classification: G01, G15, D81, D83

*Thanks to Franklin Allen and Savitar Sundaresan for their invaluable guidance and support. Additional thanks go to Karim Abadir, Enrico Biffis, Andrea Buraschi, Patrick Bolton, Stephen Cecchetti, Gilles Chemla, Claudia Custodio, Paolo Fulghieri, Chris Hansman, Christopher Hennessy, Rustam Ibragimov, Raj Iyer, Marcin Kacperczyk, David Miles, Emiliano Pagnotta, Tarun Ramadorai, Ailsa Röell, Katrin Tinn and Ansgar Walther for discussions and comments that helped shaped the content of this paper. I also thank seminar participants at Imperial College Business School, the XXVII International Rome Conference on Money, Banking and Finance, and the Royal Economic Society PhD Meeting. All errors are my own.

†University of Alberta, Email: barbalau@ualberta.ca.

1 Introduction

During the 2007-2008 financial crisis, countries that were relatively more exposed (in terms of assets, debt, exports or trade) to the crisis epicenter, the United States, were among the least affected in real terms.[1] In financial markets, equity portfolios that were relatively more exposed to a United States-specific factor experienced a drop in returns that was lower than the one predicted by their pre-crisis exposures.[2] This evidence points to the existence of a negative correlation between the degree of exposure to a shock and the impact of that shock.

The existing literature on international financial contagion does not explain how the impact of a shock can decrease with exposure to it.[3] Specifically, theories of contagion predict a monotonically increasing relationship between the degree of exposure to a risk factor and the impact of a shock to that risk factor. My paper seeks to fill this gap in the literature. I propose an information based theory of contagion and provide a characterization of the conditions under which the impact of shocks decreases with exposure. The notion of contagion adopted in my paper is one where the transmission of shocks is unexplained by the observable measure of exposure to those shocks.[4]

I introduce a framework in which decision-makers learn about the risk factors that they are exogenously exposed to, but have limited capacity to process information. I focus on studying the real consequences of shocks to these risk factors by embedding my framework in a simple model of corporate investment. The baseline model features a representative firm which undertakes investment to maximize profits. The return on investment depends on the fundamentals of the economy in which the firm operates. Economic fundamentals, in turn, are modelled as a sum of risk factor exposures. Before investing, the firm chooses how much information to observe about the risk factors that fundamentals are exposed to. Importantly, the firm has limited resources to acquire and process information.

I find that learning about a risk factor optimally increases with exposure to it. This informational advantage gives rise to a non-monotonic relation between the degree of exposure to a risk factor and the impact of a shock to that risk factor. Specifically, the firm optimally chooses to

[1] See Rose and Spiegel (2010, 2011)

[2] See Bekaert, Ehrmann, Fratzscher, and Mehl (2014)

[3] Contagion is concerned with the transmission of shocks and can be most broadly described by the idea that shocks can spread and cause a great deal more damage than the original impact (Allen and Gale, 2009).

[4] This definition is in line with a large literature which defines contagion as a change in shock transmission mechanism that cannot be explained by "fundamentals", or co-movements that are deemed to be "excessive" (King and Wadhwani, 1990; Forbes and Rigobon, 2002; Karolyi, 2003; Pericoli and Sbracia, 2003; Jotikasthira, Lundblad, and Ramadorai, 2012; Bekaert et al., 2014).

learn more about the risks that it is relatively more exposed to. The reduction in uncertainty achieved through learning mitigates the impact of shocks to risk factors that the firm is relatively more exposed to, by enabling it to take a better informed investment decision. By the same token, the impact of shocks to risk factors that the firm is relatively less exposed to is amplified through the poorly informed investment decision of the firm. The interpretation of these prediction in terms of the motivating evidence is that countries which were relatively less exposed to the United States shock were more affected because decision-makers operating in these countries had a poorer understanding of the shock and, as a consequence, took actions that aggravated their circumstances.

My model shows that the actions of decision-makers can amplify or attenuate the direct impact of shocks and thus contribute to their transmission. The model builds on the intuition that the consequences of events depend not only on decision-makers' direct exposure to events, but also on the decision-makers' degree of understanding of and their reactions to those events. Risk perception research supports the notion that reactions to events depend importantly on the degree of understanding of those events.[5] Hence, understanding how decision-makers learn is central to understanding how their actions contribute to the transmission of shocks. In this paper, I explore how endogenous information choice affects decision-makers' responses to shocks and as a consequence the impact of those shocks.

The information based shock transmission mechanism I propose in this paper works through decision-makers' uncertainty.[6] Information choice reduces the uncertainty that the firm faces when choosing optimal investment. Consider a benchmark model in which the firm is endowed with an exogenous amount of information that is equally allocated among all risk factors. Relative to this exogenous information benchmark, uncertainty is lower when the firm can optimally choose which risks to learn about. Importantly, uncertainty is convex in exposure when the firm is exogenously endowed with information, but it is concave in exposure when the firm optimally learns about risk factor exposures. Alternatively stated, uncertainty increases with exposure to relatively important risk factors when information is given, but it decreases with exposure when information is optimally chosen.[7] In the endogenous learning model I propose, a change in exposure has two competing effects. On the one hand, higher exposure to a risk factor increases uncertainty mechanically, through an *exposure channel*. On the other hand, higher exposure to a risk factor entails a reduction

[5]See Slovic (1987); Kasperson, Renn, Slovic, Brown, Emel, Goble, Kasperson, and Ratick (1988); Renn, Burns, Kasperson, Kasperson, and Slovic (1992).

[6]Uncertainty refers to the variance of beliefs about the realization of a particular shock.

[7]The importance of a risk factor is given by the interaction between the exposure to and the volatility of the factor.

in uncertainty via learning, which operates through an *information channel*. For relatively low levels of exposure to a risk factor, the exposure channel is stronger than the information channel, but for relatively high levels of exposure to a risk factor, the information channel is stronger than the exposure channel. As a consequence, uncertainty is concave in exposure to a risk factor. Thus, my model highlights a trade-off between the cost of being highly exposed to a risk and the benefit of having a better understanding of it.

Reducing uncertainty will enable the firm to accurately incorporate the shocks affecting fundamentals into investment decisions. To the extent that the firm's investment deviates from the first-best optimum obtained under perfect information, such deviations are suboptimal. In the case of positive shocks the deviation from optimality will manifest as under-investment, while in the case of negative shocks it will manifest as over-investment. It is thus convenient to define the loss due to suboptimal investment as the squared deviation from the perfect information optimum. The loss due to suboptimal investment essentially measures the contribution of actions to the transmission of shocks. I find that the loss due to suboptimal investment decreases with exposure when information is endogenously chosen, and it increases with exposure when information is exogenous. Relative to the exogenous information benchmark, the loss due to suboptimal investment is higher if a shock hits a risk factor that the firm is relatively less exposed to. On the other hand, if a shock hits a risk factor that is relatively more important in terms of exposure, the loss due to suboptimal investment is lower under the endogenous learning model than under the benchmark.

My model explains contagion manifested as shock amplification.[8] In other words it explains how shocks can have disproportionately large effects. Importantly, it sheds light on two important dimensions of shock amplification, namely: (i) which of the shocks that an entity is exposed to are likely to amplify, and (ii) which of the entities exposed to a shock are likely to be more affected. The model predicts that the transmission of shocks is intensified as exposure to shocks decreases.[9] This prediction implies that: (i) the shocks that an entity is less exposed to are amplified, and (ii) the entities that are less exposed to a shock can be more affected.

I consider a number of extensions to the baseline model. First, I extend the baseline model to account for the degree of anticipation of shocks, and I am thus also able to explain why unanticipated

[8]Amplification refers to situations in which small shocks can have disproportionately large effects (Allen and Gale, 2004; Krishnamurthy, 2010; Benoit, Colliard, Hurlin, and Pérignon, 2017)

[9]Although my baseline model is a representative agent model, comparative statics with respect to exposure are informative about a cross-section of agents which are exposed to the same set of risk factors, but which vary in their degree of exposure to these factors.

crises are more contagious.[10] Specifically, the extended model predicts that the transmission of shocks is intensified as the degree of anticipation of shocks decreases. This prediction implies that unanticipated shocks are more likely to amplify and have more severe consequences. Second, I allow for strategic interactions and find that the amplification of shocks increases with the degree of strategic complementary in investment. Third, I relax the assumption that the risk factors affecting fundamentals are independent and find that the loss due to suboptimal investment decreases with the degree of correlation between the risks. Finally, I relax the assumption that exposures to the risk factors are exogenous and find that it is optimal for the firm to specialize in learning about one risk factor and to be relatively more exposed to that factor.

The rest of the paper proceeds as follows. Section 2 formally introduces the mechanism through which the impact of shocks can decrease with exposure. Section 3 discusses the baseline results, and Section 4 considers a number of extensions to the baseline model. Section 5 outlines some concluding remarks. All proofs and derivations can be found in the Appendix.

1.1 Related Literature

This paper is mainly related and contributes to the literature on contagion, and the literature on rational inattention.

The contagion literature is vast and fraught with disagreement over how exactly to define contagion, how to measure it or what are the channels through which it operates. Contagion concerns itself with studying the transmission of shocks and can be most broadly described by the idea that shocks can spread and cause a great deal more damage than the original impact (Allen and Gale, 2009). The main dimensions of disagreement in the literature regarding what constitutes contagion have been focused around the types of linkages through which shocks are transmitted (in the presence or absence of linkages) and the types of shocks being transmitted (systemic or idiosyncratic shocks). In contrast to the contagion literature that studies shock transmission in the absence of linkages between the crisis epicenter and the entities subsequently affected, I study the transmission of shocks that occurs in the presence of linkages but which is unexplained by the observable shock transmission mechanism. In contrast to contagion literature that studies the transmission of idiosyncratic shocks across entities, I focus on studying the transmission a systemic or common shock that occurs differentially across the entities exposed to it.

[10] See evidence in Kaminsky, Reinhart, and Vegh (2003); Rigobon and Wei (2003).

My paper can be framed in the context of the literature that studies the transmission of shocks in the presence of linkages (i.e. exposure), but disproportionate to the objective measure of these linkages.[11] These include real channels such as trade linkages (Eichengreen, Rose, and Wyplosz, 1996; Glick and Rose, 1999; Forbes, 2004), as well as financial channels such as interbank linkages (Allen and Gale, 2000; Dasgupta, 2004; Freixas, Parigi, and Rochet, 2000; Iyer and Peydro, 2011) and portfolio linkages (Yuan, 2005; Pavlova and Rigobon, 2008; Jotikasthira et al., 2012; Manconi, Massa, and Yasuda, 2012). The basic idea underlying this literature is that agents transmit shocks by directly altering the linkages or, alternatively stated, by changing their exposure to risks. My paper offers an alternative explanation by showing that information choices and decisions about learning can effectively alter their risk exposures, and hence the transmission of shocks, even when agents do not directly change their exposures.

My paper is most related and contributes to the literature on information based models of financial crises and contagion (King and Wadhwani, 1990; Calvo and Mendoza, 2000; Kodres and Pritsker, 2002; Calvo, 2004; Acharya and Yorulmazer, 2008; Allen, Babus, and Carletti, 2012). The basic idea underlying this literature is that idiosyncratic shocks that should not be transmitted across entities if observable, are in fact transmitted because of imperfect information.[12] There is no role for information choice in most of these models. In contrast, I study how endogenous information choices results in the differential transmission of a systemic shock in the cross-section of entities that are exposed to it. Mondria and Quintana-Domeque (2013) also explain financial contagion using fluctuations in attention allocation and information choice. However, they focus on a financial channel of contagion and study how international investors transmit an idiosyncratic shock by optimally reallocating their attention to the risk being shocked and away from the other risks in their portfolio. In contrast, I focus on a real channel of contagion and study how firms undertaking investment optimally allocate their attention across the risks that they are exogenously exposed to and what this attention allocation implies for the transmission of shocks.[13] Ahnert and Kakhbod (2017) propose an amplification mechanism of financial crises based on the information choice of investors. They propose a global game of regime change in which changes in the public signal affects the incentives of investors to acquire private information and thus amplify the probability of a crisis. The key mechanism in their model relies on the interplay between private and public

[11] This is to be contrasted with the theories of "pure contagion" which studies the transmission of shocks in the absence of direct linkages between the crisis epicentre and the entities being affected (Forbes, 2012).

[12] The terms entities is used generically to encompass countries, financial markets or financial institutions.

[13] Alternatively stated, rather than studying how a shock to one risk factor affects the attention allocated to other factors, I focus on studying how exposure to one risk factor affects the attention allocated to other risk factors.

information, whereas my mechanism relies on the fact that learning about one risk affect the ability to learn about other risks and there is no distinction between public and private information. In other words, I distinguish between information about different risks rather than information from different sources.

My paper is also related to the rational inattention literature popularized by Sims (2003), which builds on the idea that attention, rather than information is a scarce resource. An increasing number of rational inattention applications focus on attention allocation across many risks.[14] A more recent literature focuses on allocation across states of the world.[15] I propose a micro-founded model of attention allocation across fixed exposures to risk factors, which I then also extend to account for attention allocation across states and thus provide a unified treatment of these two frameworks. Maćkowiak and Wiederholt (2015) also explore the idea that the degree to which agents are prepared for events can exacerbate their consequences. While their model focuses on degree of anticipation of shocks to study how agents make state-contingent plans, my baseline model focuses on the degree of exposure to shocks to study how the transmission of a common shock varies across the agents exposed to it. My baseline model is extended to account for the degree of anticipation of shocks as well, so my work complements theirs in that I study attention allocation both across risky exposures (type of risk) and across states of the world (degree of anticipation of risk). The result from my paper that diversified exposures are suboptimal under endogenous learning relates to the under-diversification result noted in the portfolio allocation literature (Van Nieuwerburgh and Veldkamp, 2010). While in a portfolio context agents choose both their information about and exposure to risk factors, in my baseline model agents only choose how much information to acquire about the risks they are exogenously exposed to. My model captures situations in which agents are exposed to risk that are beyond their immediate control, with a focus on understanding how attention allocation affects the transmission of shocks.

[14]Rational inattention applications to setups in which agents learn about many risks include asset pricing and portfolio choice (Mondria, 2010; Van Nieuwerburgh and Veldkamp, 2010; Kacperczyk, Van Nieuwerburgh, and Veldkamp, 2016), monetary policy (Woodford, 2001, 2009; Paciello and Wiederholt, 2013; Alvarez, Lippi, and Paciello, 2015), consumption dynamics (Luo, 2008; Tutino, 2013), price setting (Mackowiak and Wiederholt, 2009; Stevens, 2015; Matějka, 2015).

[15]Rational inattention applications to setups in which agents learn about future states of the world has been explored in a static setting by Maćkowiak and Wiederholt (2015) and in a dynamic setting by Sundaresan (2018); Nimark and Sundaresan (2018); Ilut and Valchev (2017).

2 Model

This section formally introduces the mechanism through which the impact of shocks can decrease with exposure. It outlines a baseline model of learning about exogenous risk factor exposures when the capacity to process information is limited. The basic result is that it is optimal to learn more about higher risk exposures and this information advantage can mitigate the negative impact or consequences of shocks.

2.1 Structure of the Economy

I illustrate the mechanism in the context of a simple canonical model of investment. There is a risk-neutral, representative firm in the economy. The firm undertakes investment with an aim to maximize expected profits. Realized profits are given by

$$\pi = \lambda\theta - C(\lambda) \tag{1}$$

where λ is the chosen level of investment, $C(\lambda)$ is the cost of investment and θ is the exogenous gross return to investment. The return on investment is thus parametrized by an unknown exogenous state variable θ. Following Angeletos and Pavan (2004) and in line with the motivation, I interpret the random variable θ as the underlying fundamentals in the economy, but it can also be thought of as exogenous productivity or production technology. My preferred interpretation captures the idea that real investment returns are affected by the state of the economy in which the firm operates. The cost function $C(\lambda)$ is increasing and convex in investment, and assumed to take the form $\frac{\lambda^2}{2}$.

Economic fundamentals θ are a function of exogenous exposures to independent risk factors. More specifically, economic fundamentals are modelled as an exhaustive sum of independent risks

$$\theta = \alpha f_1 + (1-\alpha) f_2 \tag{2}$$

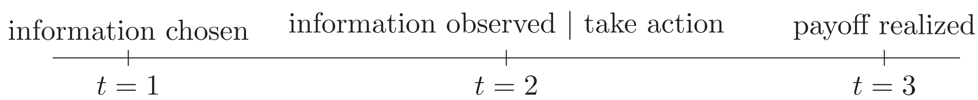
where f_1 and f_2 are exogenous risk factors which affect fundamentals in proportion to exogenous exposures or factor loadings α and $1-\alpha$ respectively. The risk factors f_i are given by

$$f_i = \mu_i + \epsilon_i, \ i = 1, 2$$

where μ_i are constants and ϵ_i are independently distributed normal random variables $\epsilon_i \sim \mathcal{N}(0, \sigma_i^2)$, which will be further referred to as shocks.

This simple factor structure accommodates a number of interpretations. The interpretation preferred here is that they are country-specific risks. This captures the intuition that aggregate economic outcomes in a country, such as GDP, are affected both by events occurring within the country i.e. domestic risks, and, to the extent that it engages in economic relations with other countries, by events occurring within those other countries i.e. foreign risks. Alternatively stated, fundamentals in a country are determined by risks that are specific to that country and risks that are specific to other countries that the country has links with. The extent to which these risks affect economic fundamentals is captured by the exposure parameters α and $1-\alpha$. For instance, the first factor f_1 can be thought of as capturing domestic risks and the second factor f_2 can be thought of as capturing foreign risks; a relatively high exposure parameter $\alpha > 0.5$ would thus describe a relatively closed economy for which domestic risks are more important, while a relatively low parameter $\alpha < 0.5$ would describe a relatively open economy that is more exposed to foreign risks.

Fundamentals are realized but unknown when the firm chooses its investment and this introduces uncertainty about the optimal level of investment. To reduce this uncertainty, the firm chooses how much information to observe about the risk factors affecting fundamentals, before investing. Reducing the uncertainty about the unobserved fundamentals will enable the firm to reduce the loss due to suboptimal investment and thus achieve a higher payoff and utility. The sequence of events is illustrated below.



The model can thus be broken down into three periods 1, 2 and 3. In the first period the representative firm chooses its information. In the second period, the firm observes the chosen information and optimally decides on a level of investment. In the third period the payoff of the investment is realized and utility is consumed. The firm's objective function is to maximize date-1 utility given by

$$U_1 \equiv E_1[E_2[\pi]] \tag{3}$$

where $E_i[\cdot]$, $V_i[\cdot]$ and $U_i[\cdot]$ denote the expected value, variance and expected utility, respectively, conditional on the information available at time i.

2.2 Information Structure

The firm takes the structure of risk factor exposures as given and decides how much to reduce uncertainty about each risk through learning. However, the firm has limited capacity to process information, meaning that its choice of how much information to observe about each risk factor is subject to a constraint on the total amount of information that it can observe.

The firm devotes limited information processing resources to learn about the factor-specific shocks affecting fundamentals. It is endowed with the prior beliefs that $\epsilon_i \sim \mathcal{N}(0, \sigma_i^2)$ and acquires noisy signals about each shock

$$s_i \quad = \quad \epsilon_i + \epsilon_{s_i}, \quad i = 1, 2 \tag{4}$$

where the signal noise is normally distributed $\epsilon_{s_i} \sim \mathcal{N}(0, \sigma_{s_i}^2)$ and uncorrelated with the other signals. The firm combines the prior beliefs with the acquired signals and forms posterior beliefs according to Bayes' law. Let $\hat{\theta}$ and $\hat{\sigma}^2$ denote the posterior mean and variance of fundamentals, respectively, conditional on the information available at time 2

$$\hat{\theta} \equiv E[\theta | s_1, s_2] \quad = \quad \alpha \left(\mu_1 + \frac{\sigma_{s_1}^{-2}}{\sigma_1^{-2} + \sigma_{s_1}^{-2}} s_1 \right) + (1 - \alpha) \left(\mu_2 + \frac{\sigma_{s_2}^{-2}}{\sigma_2^{-2} + \sigma_{s_2}^{-2}} s_2 \right) \tag{5}$$

$$\hat{\sigma}^2 \equiv V[\theta | s_1, s_2] \quad = \quad \alpha^2 \frac{1}{\sigma_1^{-2} + \sigma_{s_1}^{-2}} + (1 - \alpha)^2 \frac{1}{\sigma_2^{-2} + \sigma_{s_2}^{-2}}. \tag{6}$$

Denoting the factor-specific posterior variance by $\hat{\sigma}_i^2 \equiv \left(\sigma_i^{-2} + \sigma_{s_i}^{-2} \right)^{-1}$, $i = 1, 2$, the conditional mean and variance of fundamentals can be re-written as

$$\hat{\theta} \equiv E[\theta | s_1, s_2] \quad = \quad \alpha \left(\mu_1 + \frac{\hat{\sigma}_1^2}{\sigma_{s_1}^2} s_1 \right) + (1 - \alpha) \left(\mu_2 + \frac{\hat{\sigma}_2^2}{\sigma_{s_2}^2} s_2 \right) \tag{7}$$

$$\hat{\sigma}^2 \equiv V[\theta | s_1, s_2] \quad = \quad \alpha^2 \hat{\sigma}_1^2 + (1 - \alpha)^2 \hat{\sigma}_2^2. \tag{8}$$

The firm has limited resources or capacity to process information about the risk factors that fundamentals are exposed to. Let K denote the total capacity to process information and let k_i the amount of capacity devoted to learning about risk factor $i = 1, 2$. The information processing constraint can be generically formulated as

$$k_1 + k_2 \leq K. \tag{9}$$

Essentially, this condition tells us that the firm's choice of how much information to observe about each risk factor is subject to a constraint on the total amount of information it can observe. It also implies that for a given total capacity, learning more about one risk reduces the resources that can be devoted to learning about the other risk.

The capacity devoted to learning about a risk factor k_i, henceforth referred to as factor-specific information processing capacity, is essentially a measure of the reduction in uncertainty that can be achieved through learning. I employ the rational inattention framework proposed by Sims (2003), and model this reduction in uncertainty using tools from information theory, namely entropy and mutual information. Entropy is the standard measure of information in information theory, and it measures the amount of uncertainty in a distribution. Mutual information is the difference between the entropy of an unconditional and a conditional distribution, and it measures the amount of uncertainty resolved by conditioning on information. The factor-specific information processing capacity k_i is defined as the difference between the entropy of prior and posterior beliefs. The higher the factor-specific information processing capacity k_i, the higher the uncertainty resolved by a signal and the more informative or precise the signal is said to be. Given the assumption of normally distributed priors and signals, the factor-specific information processing capacity k_i is given by

$$k_i \equiv \frac{1}{2} \ln \frac{\sigma_i^2}{\hat{\sigma}_i^2}, \quad i = 1, 2 \tag{10}$$

where $\hat{\sigma}_i^2$ is the factor-specific posterior variance.

The entropy-based learning technology essentially imposes a bound on the product of posterior precisions[16]. This has two implications that make this information processing technology suitable for the setup considered here. First, it has a form of increasing returns to learning built into it, which means that it is less costly to learn about factors that are already well understood[17]. This captures the intuition that learning about risk factor exposures that are fixed, stable or sticky, and which take a long time to build or terminate, such as the cross-country real or financial links considered in the motivating example, is a process of refined learning. Secondly, the entropy-based learning technology accounts for the fact that the number and composition of risks affecting fundamentals are relevant for learning, which means that it is costly to learn about all the risk factors[18]. It

[16] This follows from combining (9) and (10) and re-writing it as $\prod_{i=1}^{2} \sigma_i^2 \hat{\sigma}_i^{-2} \leq e^{2K}$. Thus, more information capacity implies a higher product of (weighted) posterior precisions $\hat{\sigma}_i^{-2}$.

[17] This is due to the fact that an increase in signal precision when prior precision is high increases the product by less (since posterior precision is the sum of prior precision and signal precision).

[18] This comes from the fact that the marginal cost of an increase in precision about one risk factor is proportional

captures the intuition that learning about one risk factor will affect the ability to learn about the other factors and it is less costly to specialize in learning about one rather than all the risk factors.

In addition to the capacity constraint, the agent also faces a no-forgetting constraint which rules out the possibility of forgetting information about one risk in order to obtain more information about another one, without violating the capacity constraint.

$$k_i \geq 0, \quad i = 1, 2 \tag{11}$$

This is essentially a condition that the precision of each signal must be non-negative and it captures the intuition that learning about a risk factor should not increase uncertainty i.e. posterior variance should not exceed prior variance.

2.3 Solving the model

Given a level of capacity K, a solution to the model is: a choice of factor-specific capacity k_i to maximize date-1 utility (3), subject to the capacity constraint (9), the no-forgetting constraint (11) and rational expectations about the date-2 (conditional) investment; posterior beliefs which are formed according to Bayes' law (5) and (6), given a signal about the risk factors; a choice of investment that maximizes expected utility, given the signal realization.

The model is solved using backward induction. First, given an arbitrary information choice, the firm decides the optimal investment. Then, given the optimal investment for each information choice, the firm decides the optimal information choice.

3 Results

In this section, I derive the equilibrium allocation of information capacity across risk factors. Then I discuss the implications of these information choices in terms of the uncertainty faced by the firm when investing. Finally, I discuss the implications in terms of the chosen level of investment and the loss due to suboptimal investment.

to the precision about the other risk factors (since the constraint applies to the product of factor-specific posterior precisions).

3.1 Optimal Information Choice

The date-2 problem consists of the firm choosing an investment level to maximize expected profits (1), while taking information choice as given

$$\max_{\lambda} U_2 \equiv E_2[\pi] = \lambda E_2[\theta] - \frac{\lambda^2}{2} \tag{12}$$

where E_2 denotes the expected value conditional on the information available at date 2.

The first order condition with respect to λ yields best investment response $\lambda = E_2[\theta] = \hat{\theta}$. Thus, for any given information choice, the optimal investment level is the expected level of economic fundamentals. Substituting this optimal investment choice into the objective (12) delivers the indirect date-2 utility of having any posterior beliefs and investing optimally

$$U_2 = \frac{(E_2[\theta])^2}{2} = \frac{\hat{\theta}^2}{2}. \tag{13}$$

The date-1 problem consists of choosing the optimal level of capacity devoted to learning about each risk factor to maximize the expected value of (13), and subject to the capacity constraint (9) and the no-forgetting constraint (11). The posterior mean belief $\hat{\theta}$ is unknown at date-1. It is a normally distributed random variable, $\hat{\theta} \sim \mathcal{N}(\theta, \sigma^2 - \hat{\sigma}^2)$, so date-1 utility is given by

$$U_1 \equiv E_1[U_2] = \frac{E_1[\hat{\theta}^2]}{2} = \frac{E_1[\hat{\theta}]^2 + V_1[\hat{\theta}]}{2} = \frac{\theta^2 + \sigma^2 - \hat{\sigma}^2}{2} \tag{14}$$

Since date-1 utility is decreasing in posterior uncertainty $\hat{\sigma}^2$ and all other terms are exogenous variables, maximizing equation (14) is equivalent to

$$\max_{k_1, k_2} \quad -\hat{\sigma}^2 = \alpha^2 \hat{\sigma}_1^2 + (1-\alpha)^2 \hat{\sigma}_2^2$$

$$\text{s.t. } \hat{\sigma}_i^2 = \sigma_i^2 e^{-2k_i} \quad \text{and} \quad \sum_{i=1}^{2} k_i \leq K \quad \text{and} \quad 0 \leq k_i, \quad i = 1, 2.$$

The unique solution to this problem delivers the following optimal factor-specific capacity allocation

$$k_1 = \begin{cases} 0 & \text{if} \quad \frac{\alpha\sigma_1}{(1-\alpha)\sigma_2} < e^{-K} \\ \frac{1}{2}\left(K + \ln \frac{\alpha\sigma_1}{(1-\alpha)\sigma_2}\right) & \text{if} \quad e^{-K} \leq \frac{\alpha\sigma_1}{(1-\alpha)\sigma_2} \leq e^{K} \\ K & \text{if} \quad \frac{\alpha\sigma_1}{(1-\alpha)\sigma_2} > e^{K} \end{cases} \tag{15}$$

and $k_2 = K - k_1$.

At the interior optimum, the optimal level of capacity devoted to learning about a risk factor increases with factor-specific exposure, factor-specific prior uncertainty and total information processing capacity. In other words, the firm will optimally choose to learn more about the risk factors that fundamentals are more exposed to and which are ex-ante more uncertain. Note that corner solutions are possible. For a given level of capacity, no (all) capacity is allocated to a risk factor if factor-specific exposure and prior uncertainty are low (high) enough relative to the other factor.

The higher the capacity to process information K the smaller the range of (exposure and uncertainty) parameter values for which corner solutions are obtained, in line with the intuition that less capacity constrained agents are able to attend to more sources of risk. Note that for any limited information processing capacity, the firm will stop learning about one of the risk factor exposures. In other words, for any finite level of capacity $\forall\ K < \infty$, there exists a level of exposure $0 < \alpha < 1$ for which the conditions in (15) hold and corner solutions are obtained.

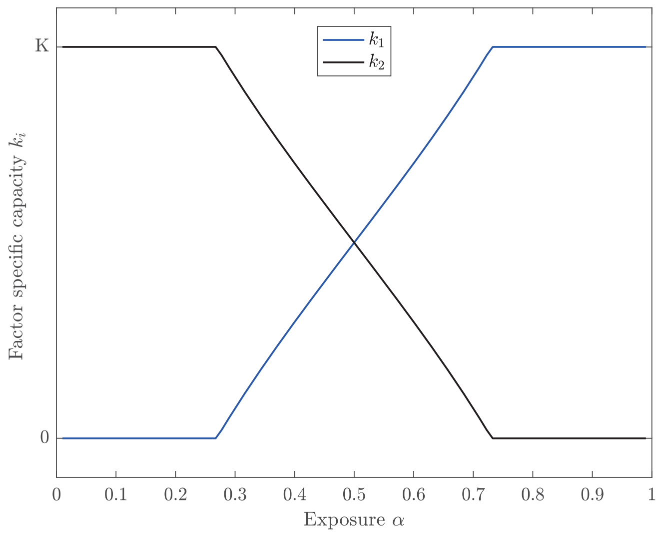


Figure 1. The parameter values are $\sigma_1 = \sigma_2 = 1$ and $K = 1$.

Figure 1 plots the optimal level of information processing capacity allocated to the two factors against exposure to factor 1. If exposure to factor 1 is very low, then the firm optimally chooses to pay no attention to it and instead devotes all information processing resources to factor 2. The economic intuition is that when exposure to factor 1 is very low, the marginal benefit of learning about factor 1 is lower than the benefit of learning about factor 2, whose exposure is relatively higher. Consequently, the firm would like to forget information about factor 1 in order to obtain more information about factor 2 but the no-forgetting constraint prevents it from doing so and the zero corner solution is obtained. A similar reasoning is applied when exposure to factor 1 is very high, leading to the full capacity corner solution. At the interior optimum, factor-specific capacity increases with factor-specific exposure and the firm optimally learns more about the risk factor fundamentals are relatively more exposed to.

3.2 Implications for Uncertainty

Information choice is a tool to reduce the uncertainty that the firm faces when deciding on the level of investment. Relative to an equal capacity, exogenous information benchmark, uncertainty is lower when the firm can optimally choose what risk factors to learn about. Importantly, uncertainty is convex in exposure when the firm is exogenously endowed with information, but it is concave in exposure when the firm optimally learns about risk factor exposures.

Given the optimal information processing capacity allocated to each factor in (15), uncertainty can be backed out using results (10) and (8), and it is given by

$$\hat{\sigma}^2 = \begin{cases} \alpha^2\sigma_1^2 + (1-\alpha)^2\sigma_2^2 e^{-2K} & \text{if} \quad \frac{\alpha\sigma_1}{(1-\alpha)\sigma_2} < e^{-K} \\ 2\alpha(1-\alpha)\sigma_1\sigma_2 e^{-K} & \text{if} \quad e^{-K} \le \frac{\alpha\sigma_1}{(1-\alpha)\sigma_2} \le e^{K} \\ \alpha^2\sigma_1^2 e^{-2K} + (1-\alpha)^2\sigma_2^2 & \text{if} \quad \frac{\alpha\sigma_1}{(1-\alpha)\sigma_2} > e^{K} \end{cases} \tag{16}$$

The uncertainty expressions for the corner solutions reflect the intuition that when there is no learning about a factor its factor-specific posterior uncertainty is equal to its prior uncertainty ($\hat{\sigma}_i^2 = \sigma_i^2$), while the uncertainty about the other factor is reduced in proportion to the total capacity ($\hat{\sigma}_i^2 = \sigma_i^2 e^{-2K}$). At the interior optimum, uncertainty increases with the attention grabbing attributes of the two risk factors (exposure and prior uncertainty) and decreases with total capacity.

In order to assess the implications of learning for uncertainty, and thus conditional investment, a suitable benchmark is needed for comparison. I consider as benchmark a model in which the firm is endowed with an exogenous amount of information which is equally allocated among the risk factors. This model will be further referred to as the exogenous information benchmark. The firm in this model has the same priors as the endogenous learning firm in my model but it is now endowed with signals with exogenous noise $\tilde{\epsilon}_{s_i} \sim \mathcal{N}(0, \tilde{\sigma}^2_{s_i})$, $i = 1, 2$. Posterior beliefs are formed according to Bayes' rule, such that benchmark uncertainty is

$$\tilde{\sigma}^2_B = \alpha^2 \tilde{\sigma}^2_1 + (1-\alpha)^2 \tilde{\sigma}^2_2. \tag{17}$$

where $\tilde{\sigma}^2_i \equiv \left(\sigma_i^{-2} + \tilde{\sigma}_{s_i}^{-2}\right)^{-1}$, $i = 1, 2$.

Figure 2 depicts the relationship between the degree of exposure to a risk factor and the uncertainty implied by the endogenous information model $\hat{\sigma}^2$ (blue line) and the exogenous information benchmark model $\tilde{\sigma}^2$ (red line). This exercise is informative of the uncertainty faced by firms operating in economies whose fundamentals share the same factor structure but vary in the degree of exposure to a risk factor (in this case the exposure to factor 1, measured by α). To enable meaningful comparisons, the total capacity in the endogeous learning model is set equal to the capacity implied by exogenous signal precisions in the benchmark model when the capacity constraint is binding i.e. $\frac{1}{2} \ln \frac{\sigma_1^2}{\tilde{\sigma}_1^2} + \frac{1}{2} \ln \frac{\sigma_2^2}{\tilde{\sigma}_2^2} = K$. This ensures that the two models are otherwise identical except for the ability to optimally reallocate information processing resources. Note that Figure 2 illustrates a symmetric equilibrium whereby the factors are ex-ante equally risky and the exogenous signals are equally informative, hence the symmetry around and intersection of the two lines at the $\alpha = 0.5$ exposure midpoint.

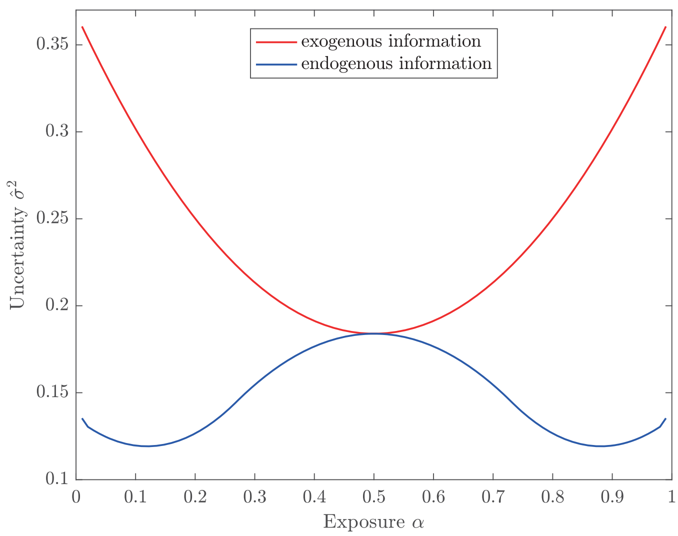


Figure 2. The parameter values are $\sigma_1 = \sigma_2 = 1$, $\tilde{\sigma}_{s_1} = \tilde{\sigma}_{s_2} = 0.75$ and the total capacity implied by these parameters is $K = \frac{1}{2} \ln \sigma_1^2 \left(\sigma_1^{-2} + \tilde{\sigma}_{s_1}^{-2}\right) + \frac{1}{2} \ln \sigma_1^2 \left(\sigma_2^{-2} + \tilde{\sigma}_{s_2}^{-2}\right) = 1$.

In terms of levels, note that relative to the equal capacity, exogenous information benchmark, uncertainty is lower when the firm can optimally allocate information resources across risk factor exposures. This is because learning effectively reduces the uncertainty about individual factors. The reduction in uncertainty achieved through learning operates through what will be further referred to as the *information channel*. Recall that the endogenous learning firm optimally learns more about the risk factor fundamentals are relatively more exposed to. The regions at the left and right of the $\alpha = 0.5$ midpoint depict situations of relatively higher exposure to a factor whose effective or learning-adjusted uncertainty is lower under the endogenous learning model than under the benchmark. Consequently, for any given level of exposure in these regions, overall uncertainty will be higher under the exogenous information model than under the endogenous information model. Appendix A.1 provides an analytical proof of this intuition.

In terms of dynamics, note that both models illustrates a non-monotonic relationship between overall uncertainty and the degree of exposure to a risk factor. However, whereas in the benchmark model uncertainty is convex in exposure, in the endogenous learning model uncertainty is concave in exposure when the firm optimally learns about both risk factors i.e. at the interior optimum. Alternatively stated, uncertainty increases with exposure to relatively important risk factors when information is exogenously given, but it decreases with exposure when information about the two risks is endogenously chosen. Thus, one implication of the endogenous learning model is that in the cross-section of entities exposed to a relatively important risk factor, an entity that is more exposed to that risk will have a better understanding of it and will thus face a lower uncertainty than an entity which is less exposed to it.

PROPOSITION 1. *Provided that a risk is relatively important, uncertainty decreases with exposure when the firm optimally learns i.e.* $\frac{\partial \hat{\sigma}^2}{\partial \alpha} < 0$ *if* $\alpha > 0.5$ *and* $k \in (0, K)$.

Proof. See Appendix A.1

Proposition 1 sheds light on why the shock transmission patterns observed during the financial crisis of 2007-2008 were different relative to those observed during past contagious crises such as the Mexican crisis of 1994 and the Asian crisis of 1997.[19] It is because in the last crisis the shock originated in a country that plays a central role in the global economy: the United States is both a country that is relatively important for most other countries ($\alpha > 0.5$), as well as a country that

[19] During these past crises the epicentre country in which the shock originated was more severely affected than the other countries that were subsequently affected by the shock. More generally, the impact of the original shock increased with exposure to it during the Mexican and Asian crises, but it decreased with exposure during the last financial crisis of 2007-2008.

foreign decision-makers are likely to learn about ($k \in (0, K)$). Thus, variation in the extent to which other countries were exposed to the United States, translates into variation in uncertainty which follows the dynamics illustrated at the right of the $\alpha = 0.5$ exposure midpoint in Figure 2. In other words, my model suggests that countries which were relatively more exposed to the United States shock faced a relatively lower level of uncertainty regarding the implications of the shock. This provides a potential explanation for why during the financial crisis of 2007-2008, unlike during previous crises, the transmission of the shock decreased with exposure to the shock.

The benchmark model illustrates a diversification effect whereby high exposure to a risk factor implies high overall uncertainty, and the lowest level of uncertainty is achieved when exposure to the two factors is equal. To understand the forces that are at play let us focus on the right part of Figure 2, where exposure to factor 1 is relatively higher i.e. $\alpha > 0.5$. Given two equally risky factors, which is the case in light of the symmetric equilibrium considered, overall uncertainty is driven by the factor that is important in terms of exposure. Consequently, increasing exposure to factor 1 beyond the 0.5 midpoint results in an increase in overall uncertainty. This works through what will be further referred to as the *exposure channel.* The idea behind it is that the overall risk of a bundle of factors which are equally risky will be driven by those that are important in terms of exposure; in other words, for fixed risk overall dynamics are dictated by exposure.

The endogenous learning model, on the other hand, is convex in exposure when the firm optimally learns about one factor only, and it is concave in exposure when the firm learns about both factors.[20] In the corners, when the firm devotes all capacity to learning about one factor, information is essentially exogenous, so uncertainty dynamics will be the same as in the benchmark and will operate through the exposure channel. At the interior optimum though, uncertainty initially increases with exposure, but once the factor becomes important in terms of exposure, uncertainty will start to decrease. To understand the mechanism behind this, let us start from the situation in which exposure to the two factors is equal i.e. $\alpha = 0.5$. As exposure to factor 1 is increased beyond this 0.5 midpoint, the uncertainty about factor 1 is effectively reduced through learning. The reduction in uncertainty entailed by learning is stronger than the increase in exposure, so increasing exposure to factor 1 results in a decrease in overall uncertainty. In other words, the information channel dominates the exposure channel. However, the benefits of learning are limited and increasing exposure beyond a certain point will undo the reduction in uncertainty achieved

[20]Note that uncertainty in the endogenous information model is convex on the same interval of exposure parameters over which the factor-specific capacity allocation illustrated in Figure 1 is extreme i.e. either zero or maximum.

through learning, resulting in an increase in overall uncertainty. In other words, the exposure channel will overturn the information channel if exposure is high enough and the firm only learns about one factor. This is reflected in the turning point in the uncertainty, which occurs when the learning-adjusted risk exposures of the two factors are equal.[21]

In sum, the benchmark illustrates a classic diversification effect which operates through the exposure channel. In the endogenous learning model, learning substitutes for diversification in reducing risk, resulting in a specialization effect. In a portfolio allocation context, Van Nieuwerburgh and Veldkamp (2010) also make the point that diversification is not optimal if portfolio choice is preceded by information choice. While they focus on the implications of information choice in terms of portfolio holdings when agents choose both their information about and exposure to risks, I focus on the role of information in terms of the impact of shocks when agents choose their information about risks they are exogenously exposed to. Section 4.4 relaxes the assumption of exogenous exposures to risk factors and provides a characterization of the optimal exposure points.

3.3 Implications for Investment

Reducing uncertainty will enable the firm to take an investment decision that is more aligned with underlying fundamentals and thus reduce the loss due to suboptimal investment.[22] Insofar as the firm' investment deviates from the first-best optimum obtained under perfect information, it contributes to the transmission of shocks. The basic result is that the transmission of shocks decreases with exposure when the firm optimally learns about the shocks affecting fundamentals. Consequently, relative to the exogenous learning benchmark, the impact of shocks that fundamentals are relatively less exposed to is amplified, while the impact of shocks that fundamentals are relatively more exposed to is attenuated.

Recall that the firm optimally chooses a level of investment that is equal to the expected level of fundamentals conditional on the information available at the intermediate date 2 i.e. $\lambda = E_2[\theta] = \hat{\theta}$. Thus, the investment decision is essentially a response to information about the realization of shocks affecting fundamentals. Figure 3 depicts the relationship between the degree of exposure to a shock and the response or sensitivity of investment to that shock that is implied by the exogenous

[21] At the first corner (when learning about factor 2) uncertainty decreases if $\alpha\sigma_1^2 < (1-\alpha)\sigma_2^2 e^{-2K}$, which essentially reads that increasing exposure to a factor whose learning-adjusted risk is relatively lower decreases overall uncertainty. At the second corner (when learning about factor 1), uncertainty starts to increase when $\alpha\sigma_1^2 e^{-2K} > (1-\alpha)\sigma_2^2$ i.e. when learning-adjusted risk is higher.

[22] Mathematically a lower level of uncertainty means higher response to signals that are informative about the underlying shocks affecting fundamentals.

information benchmark (red line), the endogenous information model (blue line) as well as a full information model (black line). Note that the shock sensitivity increases with exposure under all the three models considered but the rate of increase is different. Under the full information model shocks can be perfectly observed so there is a one to one mapping between exposure to the shock and the sensitivity to it. This is the optimal or first-best response; to the extent that responses deviate or are not aligned with it they are said to be suboptimal and to act as a shock transmission mechanism. Under exogenous information, the response departs from the optimal one as exposure increases because shocks are observed with a precision that is fixed so the exposure effect dominates. Under the endogenous info model three situations can be observed: for sufficiently low exposure, the firm does not acquire any information and there is no response to shocks; shock sensitivity is zero because as the firm is essentially unaware of the underlying shock realization. Second, as exposure increases and the firm starts learning about factor 1, its response to shocks will become increasingly more aligned with the optimal one because the precision of information about shocks affecting fundamentals optimally increases with exposure. Finally, when exposure is sufficiently high that the firm only learns about factor 1, the response starts to decrease again with exposure because information precision, albeit set at the maximum, is essentially exogenous.

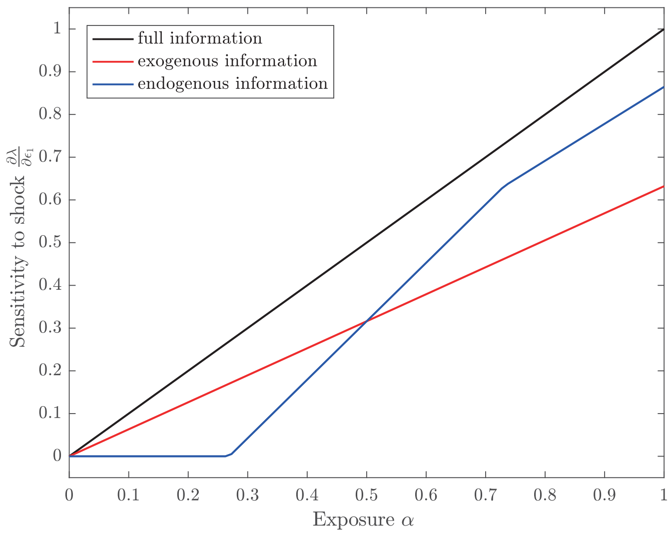


Figure 3. The parameter values are $\sigma_1 = \sigma_2 = 1$, $\tilde{\sigma}_{s_1} = \tilde{\sigma}_{s_2} = 0.75$ and $K = 1$.

Thus, information frictions introduce a loss or inefficiency relative to the full information benchmark. When shocks can be perfectly observed they affect investment in proportion to exposure. However, when shocks cannot be perfectly observed investment decisions are relatively less responsive to shocks; the firm fails to fully incorporate underlying shocks into decision-making. In case of positive shocks, the deviation from the full information optimum is negative and represents a situation of underinvestment. The firm increases investment by less than it optimally should, resulting in missed or lost business opportunities. In case of negative shocks the deviation is positive and decreases with exposure. This is a situation of overinvestment, whereby the firm reduces investment by less than it optimally should, resulting in wasted resources through excess capacity. In order to ease analysis and abstract from the nature of shocks, I define the loss due to suboptimal action as $L \equiv (\hat{\theta} - \theta)^2$. This symmetric loss function captures a more general situation in which shocks are changes in circumstances that the firm needs to adapt to rather than good or bad events i.e. introduction of standards, technological disruption, terms of trade changes.

Figure 4 plots the loss due to suboptimal action against exposure to factor 1. The example considers a one standard deviation shock to factor 1, abstracts from factor 2 shocks as well as from information shocks. Under the exogenous information benchmark, loss increases monotonically with exposure to the shock (red line). Given that signal noise is fixed, this result works through the exposure channel and is due to increasing exposure to a risk that is constant. Under the endogenous learning model, three scenarios can be observed as exposure increases (blue line). First, when exposure is sufficiently low such that the firm does not learn about factor 1, the loss due to suboptimal investment is increasing in exposure. The rate of increase is higher relative to the exogenous information benchmark because the firm chooses to observe no information (as opposed to fixed information) about factor 1. Second, as exposure increases and the firm starts learning about factor 1, the loss due to suboptimal action decreases with exposure because the firm is able to incorporate shocks more accurately into decision-making. Third, when the firm learns only about factor 1, the loss due to suboptimal action starts to increase again because information precision is essentially exogenous and as a consequence dynamics resemble the benchmark.

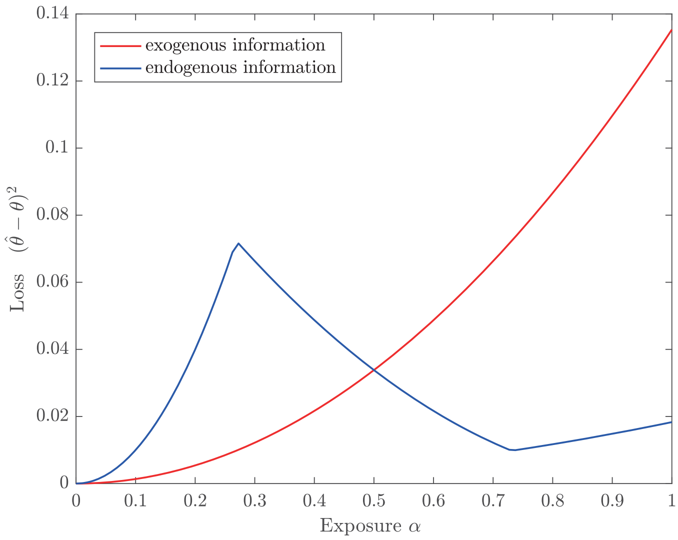


Figure 4. This example considers a one standard deviation shock to factor 1 i.e. $\epsilon_1 = \sigma_1$, abstracts from factor 2 shocks i.e. $\epsilon_2 = 0$ and information shocks i.e. $\epsilon_{s_1} = \epsilon_{s_2} = 0$. The parameter values are $\mu_1 = \mu_2 = 1$, $\sigma_1 = \sigma_2 = 1$ and $\tilde{\sigma}_{s_1} = \tilde{\sigma}_{s_2} = 0.75$ and the total capacity implied by these parameters is $K = \frac{1}{2}\ln\sigma_1^2\left(\sigma_1^{-2} + \tilde{\sigma}_{s_1}^{-2}\right) + \frac{1}{2}\ln\sigma_1^2\left(\sigma_2^{-2} + \tilde{\sigma}_{s_2}^{-2}\right) = 1$.

Worth emphasizing is the fact that the predictions of the two models are starkly different when an interior solution is obtained for endogenous information choice i.e. when the firm learns about both risks. Whereas the loss due to suboptimal action is increasing with exposure when the firm is exogenously endowed with information about the risk factors, it is decreasing with exposure when the firm optimally chooses what risks to learn about. Consequently, relative to the exogenous learning benchmark, shocks that fundamentals are relatively more exposed to are attenuated (lower loss) while shocks that fundamentals are relatively less exposed to are amplified (higher loss). The magnitude of amplification decreases with the firm's capacity to learn. In other words, amplification is more severe for more capacity constrained firms that are only able to learn about one risk factor.

If the shock to factor 1 is interpreted as a shock to United States, Figure 4 implies that a country which is less exposed to the United States (and which is situated towards the left end of the x-axis) will incur a higher loss due to suboptimal action compared to a country which is more exposed to the United States (which is situated towards the right end of the x-axis). It also implies that a country which is less exposed to the United States will incur a higher loss due to suboptimal action compared to the United States itself (which is likely to be situated towards the right end of the x-axis). These predictions are in line with evidence from the GFC that countries other than the epicentre have been more severely affected than United States itself, and more generally countries that were less exposed to the United States were more affected.

The results depicted in Figure 4 can be understood by examining how investment decisions act as a shock transmission mechanism analytically. To that end, I define the shock transmission mechanism as the change in the loss function due to a shock i.e. $\frac{\partial L}{\partial \epsilon_i}$. The shock transmission mechanism essentially measures the extent to which shocks translate into losses. The interaction between the magnitude of the shock and the strength of the shock transmission mechanism determines the impact of a shock. Consequently, statements about the impact of a shock amount to statements about the negative consequences of a shock or the losses induced by it. The stronger the shock transmission mechanism, the higher the loss due to suboptimal investment and the higher the impact of the shock is said to be.

PROPOSITION 2. *The transmission of shocks decreases with exposure when the firm optimally learns i.e.* $\frac{\partial^2 L}{\partial \epsilon_1 \partial \alpha} < 0$ *if* $k \in (0, K)$.

Proof. See Appendix A.2.

Proposition 2 pins down the mechanism through which the impact of a shock can decrease with exposure to it. When the firm optimally learns about the risk factors affecting fundamentals, the loss due to suboptimal investment that is induced by a shock decreases with exposure. In other words, the contribution of actions to the transmission of shocks decreases with exposure. This is because the reduction in uncertainty that is achieved through learning increases with exposure, and as a consequence the deviation of the firm's investment from the perfect information optimum decreases. Thus, the informational benefit mitigates the direct impact of shocks to risk factors that fundamentals are relatively more exposed to as it enables the firm to take better informed decisions and minimize the loss due to suboptimal investment. By the same token, the impact of shocks to risk factors that fundamentals are relatively less exposed to is amplified through the firm's poorly informed investment decision which imply a higher loss due to suboptimal investment.

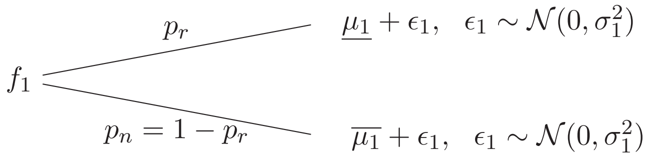
4 Extensions

In this section, I consider the following extensions to the baseline model studied so far. First, I account for the degree of anticipation of shocks. Second, I allow for strategic complementarity in investment. Third, I relax the assumption that the risk factors affecting fundamentals are independent. Finally, I relax the assumption that exposures to the risk factors are exogenous.

4.1 Extension: Shock Anticipation

The baseline model considers the case in which the firm allocates an exogenously given information processing capacity K across risk factor exposures. This section endogenizes the capacity available to the firm in a certain state of nature, by linking it to the degree of anticipation of that state. The basic result is that the impact of unanticipated shocks is amplified because the firm optimally allocates less information processing capacity to learning about low probability states of nature.

One of the risk factors affecting fundamentals is assumed to be in one of two possible states of nature: a low-probability, low-mean so-called bad state of nature - interpreted as rare times, and a high-probability, high-mean so-called good state of nature - interpreted as normal times. Let $p_r > 0$ denote the probability that factor 1 is in the bad state of nature. If it is ex-ante unlikely for this state to occur, then p_r will be small, and a rare event or crisis is said to occur if the state is revealed to be bad. This setup can thus be graphed as



where $\underline{\mu_1} < \overline{\mu_1}$ and $p_r < p_n$. To isolate the effect of the degree of anticipation alone, I assume that only the mean level of the risk factors is different in the two states while the priors associated with the shocks affecting the risk factors in the two states are the same.[23]

In this setup, there are two dimensions of information choice: how much information to acquire about a state of the world, and how to allocate that information across risk factors in each state of the world.[24] Let $\mathcal{K}$ denote the *total capacity* or total amount of information resources available to the firm , K_s the *state capacity* or amount of information resources dedicated to state of nature $s \in \{n, r\}$, and k_{is} denote the amount of information processing capacity dedicated to factor i in state $s \in \{n, r\}$. The capacity constraint (9) governing the allocation of capacity across risks in any state of nature $s \in \{n, r\}$ can now be formulated as

$$k_{1s} + k_{2s} \leq K_s$$

[23] In any state of nature, the capacity allocation across risk factors is mean independent but increases with the prior uncertainty surrounding that factor i.e. volatility. Accounting for the intuition that crises episodes are characterized by high-volatility would weaken the ensuing result but the main result still holds. Appendix A.3 deals with this case.

[24] This can be thought of as capturing situations in which decision-makers prepare for different contingencies.

and the capacity constraint governing the allocation of total capacity across states of nature is

$$K_n + K_r \leq \mathcal{K}.$$

The model is solved following the same steps as in the baseline model in Section 3, except that now the date-1 problem consists of two steps. As before, the firm first allocates information resources across risk factor exposures given the optimal investment level and an arbitrary state capacity. The second step is to allocate total information resources across states of nature given the optimal investment and optimal capacity allocation across exposures in any state. Let $U_1(K_s)$ denote the date-1 utility of investing optimally at the second date and optimally allocating the available state capacity K_s across risk factors at the first date

$$U_1(K_s) = \begin{cases} -\alpha^2\sigma_1^2 - (1-\alpha)^2\sigma_2^2 e^{-2K_s} & \text{if } k_1 = 0 \\ -2\alpha(1-\alpha)\sigma_1\sigma_2 e^{-K_s} & \text{if } k_1 = \frac{1}{2}\left(K_s + \ln\frac{\alpha\sigma_1}{(1-\alpha)\sigma_2}\right) \\ -\alpha^2\sigma_1^2 e^{-2K_s} - (1-\alpha)^2\sigma_2^2 & \text{if } k_1 = K_s \end{cases} \tag{18}$$

The date-1 problem for the allocation of total capacity across states is

$$\max_{K_n, K_r} \; p_n U_1(K_n) + p_r U_1(K_r) \tag{19}$$

$$\text{s.t. } \; K_n + K_r \leq \mathcal{K} \; \text{ and } \; K_s \geq 0, \; s \in \{n, r\} \tag{20}$$

The basic result is that the optimal level of information processing capacity dedicated to a state of nature increases with the probability of occurrence of the state. Appendix A.3 provides a full characterization of the equilibria. At the interior optimum, when the firm learns about both factors, the optimal level of capacity dedicated to the rare state is given by

$$K_r = \begin{cases} 0 & \text{if } \frac{p_r}{p_n} < e^{-\mathcal{K}} \\ \frac{1}{2}\left(\mathcal{K} + \ln\frac{p_r}{p_n}\right) & \text{if } \frac{p_r}{p_n} \geq e^{-\mathcal{K}} \end{cases} \tag{21}$$

Note that a corner solution is possible. More specifically, if the probability of the state of nature is low enough, no capacity is allocated to the state. Otherwise, the information-processing capacity allocated to a state of nature increases with its degree of anticipation, as well as with the total capacity $\mathcal{K}$ available.

The implication of this capacity allocation in terms of state contingent investment schedules and the loss due to suboptimal investment is that the impact of shocks decreases with their degree of anticipation because the loss due to suboptimal action decreases with anticipation. The firm optimally devotes little attention to low-probability events. Thus, the contribution of investment decisions to the transmission of shocks is higher the lower their ex-ante probability of occurrence. In other words, the deviation of the firm's investment from the perfect information optimum will be higher the lower the probability of occurrence of a shock, which essentially intensifies its transmission. The amplification of unanticipated shocks will thus be higher.

PROPOSITION 3. *The transmission of shocks decreases with their degree of anticipation when the firm optimally learns i.e.* $\frac{\partial^2 L}{\partial \epsilon_1 \partial p_s} < 0$ *if* $k_{1s} \in (0, K_s)$.

Proof. See Appendix A.3.

Proposition A.3 essentially says that the contagious transmission of crises is higher the lower their probability of occurrence. This prediction is in line with evidence that documents a negative relation between the degree of anticipation of crises and the occurrence of contagion (Kaminsky et al., 2003; Rigobon and Wei, 2003; Didier, Mauro, and Schmukler, 2008; Mondria and Quintana-Domeque, 2013). Additionally, it is in line with the debates on whether the highly unexpected nature of the Lehman shock might have amplified its transmission. My model predicts that contagion is more likely to occur following unexpected crises because decision-makers optimally prepare less for unexpected states of the nature.

Figure 5 plots the loss due to suboptimal investment against exposure to a one standard deviation shock to factor 1, for varying degrees of anticipation of the shock. The loss due to suboptimal investment is larger when the shock occurs with a small probability (solid lines), relative to the case in which the shock occurs with a higher probability (dashed lines). The loss due to suboptimal investment that is induced by a shock decreases with the degree of anticipation of that shock because the firm optimally learns less about less anticipated states of nature. Thus, endogenizing the information processing capacity available in a state of nature has the effect of amplifying the impact of shocks occurring in unexpected states; decision-makers are unprepared for unexpected shocks and this amplifies their consequences.

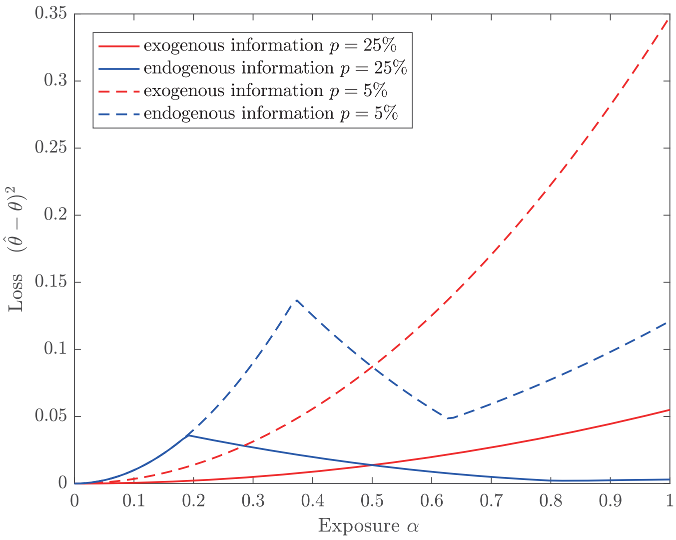


Figure 5. This example considers a one standard deviation shock to factor 1 i.e. $\epsilon_1 = \sigma_1$, abstracts from factor 2 shocks i.e. $\epsilon_2 = 0$ and information shocks i.e. $\epsilon_{s_1} = \epsilon_{s_2} = 0$. The parameter values are $\sigma_1 = \sigma_2 = 1$; when $p = 25\%$ $\tilde{\sigma}_{s_1} = \tilde{\sigma}_{s_2} = 0.55$ and $K = 1.45$, and when $p = 5\%$ $\tilde{\sigma}_{s_1} = \tilde{\sigma}_{s_2} = 1.20$ and $K = 0.53$.

4.2 Extension: Investment Complementarities

This subsection extends the baseline model to account for strategic interactions and their implications for equilibrium information choice and the loss due to suboptimal investment. Relative to the baseline model, strategic complementarity in investment reduces the incentive to learn about risks fundamentals have relatively low exposure to. Thus, the transmission of shocks that fundamentals are relatively less exposed to is intensified when firms seek to coordinate their investment decisions, which implies that the loss due to suboptimal investment is higher relative to the baseline.

There is a continuum of firms indexed by j. Each firm chooses level of investment λ_j to maximize expected profits. The profit function for firm j is given by

$$\pi_j = R\lambda_j - \frac{1}{2}\lambda_j^2 \tag{22}$$

The return on investment R is a function of the unknown fundamentals in the economy θ and the average investment in the population $\bar{\lambda} = \int_j \lambda_j$.

$$R = (1 - r)\theta + r\bar{\lambda}$$

where r is a constant governing the type of strategic interactions between firms. Real investment environments have typically been treated as being characterized by strategic complementarities. In such environments agents want to do what others do. This is modelled by imposing that $r > 0$, which implies that optimal individual responses λ_j increase in the actions of others $\bar{\lambda}$. If $r = 0$ individual actions are independent of the average action in the population and the baseline model is obtained.

As in the baseline model, the solution strategy is to work backwards. At date-2, each firm undertakes a level of investment to maximize the expected profit π_j in (22), while taking information choice as given

$$\max_{\lambda_j} U_{2j} \equiv E_{2j}[R\lambda_j - \frac{1}{2}\lambda_j^2]$$

The first-order condition yields

$$\lambda_j = E_{2j}[R] = (1-r)E_{2j}[\theta] + rE_{2j}[\bar{\lambda}] \tag{23}$$

I consider equilibria in which the mean investment in the population is a linear function of the shocks affecting fundamentals

$$\bar{\lambda} = \psi + \phi_1\epsilon_1 + \phi_2\epsilon_2 \tag{24}$$

where ψ, ϕ_1 and ϕ_2 are constants determined in equilibrium. Recalling that $E_{2j}[\theta] = \alpha(\mu_1 + E_{2j}[\epsilon_1]) + (1-\alpha)(\mu_2 + E_{2j}[\epsilon_2])$ and substituting conjecture (24) into the first order condition (23) yields

$$\lambda_j = r\psi + (1-r)(\alpha\mu_1 + (1-\alpha)\mu_2) + [\alpha(1-r) + r\phi_1]E_2[\epsilon_1] + [(1-\alpha)(1-r) + r\phi_2]E_2[\epsilon_2]$$

Calculating first the conditional expectation of the shocks $E_{ij}[\epsilon_i] = (1-\gamma_i)s_{ij}$, and then the mean action in the population yields

$$\bar{\lambda} = r\psi + (1-r)(\alpha\mu_1 + (1-\alpha)\mu_2) + [\alpha(1-r) + r\phi_1](1-\gamma_1)\epsilon_1 + [(1-\alpha)(1-r) + r\phi_2](1-\gamma_2)\epsilon_2$$

Matching coefficients verifies the conjecture (24) that the average investment level is linear in the shocks and that the linear weights are $\psi = r\psi + (1-r)(\alpha\mu_1 + (1-\alpha)\mu_2)$, $\phi_1 = [\alpha(1-r) + r\phi_1](1-\gamma_1)$

and $\phi_2 = [(1-\alpha)(1-r) + r\phi_2](1-\gamma_2)$. Collecting the unknown coefficients yields

$$\psi = \alpha\mu_1 + (1-\alpha)\mu_2 \tag{25}$$
$$\phi_1 = \frac{\alpha(1-r)(1-\gamma_1)}{1-r(1-\gamma_1)} \tag{26}$$
$$\phi_2 = \frac{(1-\alpha)(1-r)(1-\gamma_2)}{1-r(1-\gamma_2)} \tag{27}$$

The date-1 problem consists of choosing the optimal level of capacity devoted to learning about each risk factor to maximize expected utility implied by the investment rule (24) and the equilibrium coefficients (25)-(27), subject to the capacity constraint (9) and the no-forgetting constraint (11)

$$\max_{k_1,k_2} \; U_{1j} = E_{1j}[U_{2j}] = \frac{1}{2}\Big[\psi^2 + \frac{\alpha^2(1-r)^2\sigma_1^2(1-\gamma_1)}{(1-r(1-\gamma_1))^2} + \frac{(1-\alpha)^2(1-r)^2\sigma_2^2(1-\gamma_2)}{(1-r(1-\gamma_2))^2}\Big]$$

$$\text{s.t.} \quad \gamma_i = e^{-2k_i}, \; \sum k_i \le K, \; 0 \le k_i, \; i = 1, 2$$

Numerical results indicate that relative to the baseline model ($r = 0$), the firm is more likely to learn about a single risk rather than both risks when investment actions are strategic complements ($r > 0$). In other words, as the degree of strategic complementarities increases, corner solutions occur more easily. In fact, if the degree of strategic complementarity r is high enough the parameter region for which an interior solution is obtained collapses to a single point.[25] The implication is that for high levels of strategic complementarity, a small change in the exposure parameter can have a large effect on the equilibrium allocation of attention.

At the interior optimum, when the firm optimally learns about both risks, the optimum level of capacity allocated to factor 1 decreases with the degree of complementarity r if exposure to factor 1 is relatively low ($\alpha < 0.5$) but it increases if exposure to factor 1 is relatively high ($\alpha > 0.5$). This is because the firm anticipates that learning optimally increases with exposure in the population, and as a consequence it also chooses to learn less about low-exposure risk factor and more about high-exposure risk factors. In other words, strategic complementarity in investment reduces the incentive to learn about risk factors that fundamentals have relatively low exposure to, which implies that the transmission of shocks to these risk factors will be amplified compared to the baseline model. In other words, the contagious transmission of shocks is intensified as the degree of strategic complementarity increases.

[25]For strategic complementary parameters beyond this point multiple equilibria exist.

The implications in terms of shock impact are illustrated in Figure 6, which plots the loss due to suboptimal investment against exposure to a one standard deviation shock to factor 1, for varying levels of strategic complementarity. The loss due to suboptimal investment is larger in environments characterized by a higher level of strategic complementarity (dashed lines), relative to the case in which there are no strategic interactions (solid lines). This is due to the fact that the firm's incentive to hedge against shocks through learning is weakened by its desire to coordinate its investment decison with the average investment in the population. Since the firm anticipates other firms will optimally choose to learn less about the risk factors that fundamentals have relatively little exposure to, its incentive to learn about these low-exposure risk factors decreases as the degree of strategic complementarity increases. As a consequence, the loss due to suboptimal action that is induced by a shock is amplified in the presence of strategic complementarities.

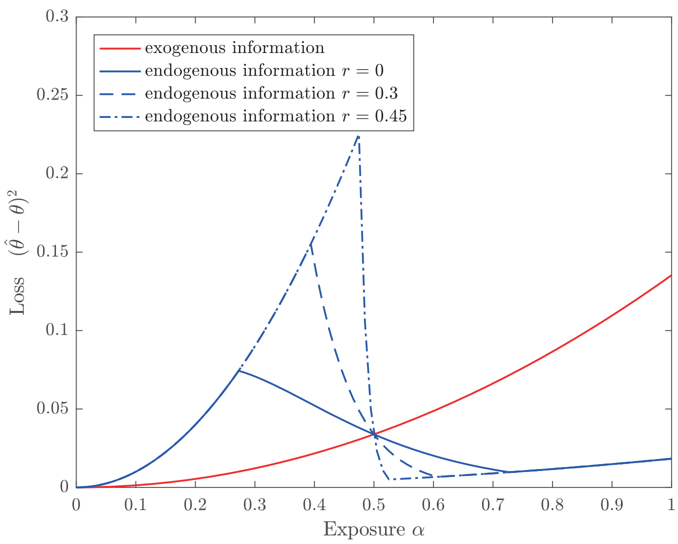


Figure 6. This example considers a one standard deviation shock to factor 1 i.e. $\epsilon_1 = \sigma_1$, abstracts from factor 2 shocks i.e. $\epsilon_2 = 0$ and information shocks i.e. $\epsilon_{s_1} = \epsilon_{s_2} = 0$. The other parameter values are $\sigma_1 = \sigma_2 = 1$, $K = 1$ and $\tilde{\sigma}_{s_1} = \tilde{\sigma}_{s_2} = 0.75$.

4.3 Extension: Correlated Risks

The baseline model considers the case in which the risk factors affecting fundamentals are independent. In this section I allow for the risk factors to be correlated. I find that the baseline main result remains unchanged and the loss due to suboptimal investment increases with the degree of correlation between the two risk factors.

To deal with the case of correlated risks it is useful to use matrix notation. As in the baseline model, fundamentals are a sum of risk factors. Let f be a $N \times 1$ vector of risk factors and A be a $N \times 1$ vector of exposures to these factors. Fundamentals can be expressed as

$$\theta = A'f$$

where the factors are ex-ante known to be correlated i.e. the prior variance-covariance matrix of the risk factors f is non-diagonal. Assuming that the prior variance-covariance matrix of the factors f is non-diagonal is equivalent to assuming the following linear structure for the risk factors[26]

$$f = \mu + \Gamma\epsilon, \quad \epsilon \sim \mathcal{N}(0, \Sigma) \tag{28}$$

where μ is a $N \times 1$ vector of constants measuring the mean level of each risk factor, ϵ is a $N \times 1$ vector of independent shocks with diagonal variance-covariance matrix Σ, and Γ is an $N \times N$ matrix of loadings which measures the extent to which the independent shocks affect the risk factors. The i^{th} row of the matrix Γ, denoted Γ_i, gives the loadings of the i^{th} risk factor f_i on the independent shocks in the vector ϵ. Thus, each risk factor f_i is expressed as the sum of a factor-specific mean μ_i and the independent random variables or shocks contained in the vector ϵ, which affect it in proportion to the loadings Γ_i.

This factor structure essentially allows for correlations between the risk factors through shared exposure to underlying independent shocks. It accounts for the existence of underlying forces that might be driving more than one of the risk factors affecting fundamentals. I conceptualize these independent shocks as factor-specific shocks. For instance, I interpret shock 1 ϵ_1 as being specific to factor 1 f_1, and shock 2 ϵ_2 as being specific to factor 2 f_2. Correlation is introduced by allowing

[26] The variance-covariance matrix of the risk factors that is implied by (28) is $\Gamma\Sigma\Gamma'$. Note that an alternative solution method is to assume that the prior variance-covariance matrix of the shocks is non-diagonal, say Ω, and then use eigen-decomposition to re-write is as $\Omega = \Gamma\Sigma\Gamma'$; learning would then be about the principal components with diagonal variance-covariance matrix Σ.

factor 1 to load on the shock specific to factor 2, and vice versa. Interpreted in the context of the motivating example, this setup accounts for the reality that domestic and foreign risks are likely to be correlated. The linear factor modelling approach adopted above is essentially equivalent to principal component analysis, which provides a way to decompose correlated risks into independent risks. In the portfolio literature it is common to use principal components analysis to decompose sets of correlated asset returns into independent underlying risk factors such as business-cycle risk, industry-specific risk, and firm-specific risk (Ross, 1976). Similarly, correlated domestic and foreign risks can be decomposed into an exhaustive set of independent underlying risk factors, which can be interpreted as pure country-specific risks.

The firm aims to reduce uncertainty about these underlying shocks through learning. Thus, signals will be about the independent shocks contained in the vector ϵ. I assume that learning about independent shocks is done independently. In other words, the firm acquires independent noisy signals about each of the independent shocks contained in the vector ϵ, and thus receives a $N \times 1$ vector of independent signals

$$s = \epsilon + \epsilon_s, \quad \epsilon_s \sim \mathcal{N}(0, \Sigma_s)$$

where the variance-covariance matrix of the $N \times 1$ vector of signal noise Σ_s is diagonal, thus capturing the assumption that signal about independent risks are independent.

Applying Bayes's rule on the transformed variable $\Gamma^{-1}f$, and then pre-multiplying this solution by Γ, I obtain that posterior beliefs about the correlated risk factors have mean $E[f|s] = \mu + \Gamma(I - \hat{\Sigma}\Sigma^{-1})s$ and variance $V[f|s] = \Gamma\hat{\Sigma}\Gamma'$, where $\hat{\Sigma} \equiv (\Sigma^{-1} + \Sigma_s^{-1})^{-1}$ denotes the posterior variance-covariance matrix of the independent shocks ϵ.[27] Consequently, the posterior mean and variance of fundamentals, respectively, conditional on the information available at time 2 are

$$\begin{aligned} \hat{\theta} &\equiv E[\theta|s] = A'E[f|s] = A'(\mu + \Gamma(I - \hat{\Sigma}\Sigma^{-1})s) \\ \hat{\sigma}^2 &\equiv V[\theta|s] = A'V[f|s]A = A'\Gamma\hat{\Sigma}\Gamma'A \end{aligned}$$

The solution strategy follows the same steps as in the baseline model. The date-2 problem is unchanged and yields solution $\lambda = E[\theta|s] = \hat{\theta}$. Similarly, the date-1 utility decreases with

[27] Transforming the variable $f^* = \Gamma^{-1}f = \Gamma^{-1}\mu + \epsilon$, allows applying standard Bayesian rules for updating normally distributed variables and yields posterior mean $E[f^*|s] = \Gamma^{-1}\mu + V[f^*|s]\Sigma_s^{-1}s$ and posterior variance $V[f^*|s] = V[\epsilon|s] = (\Sigma^{-1} + \Sigma_s^{-1})^{-1} \equiv \hat{\Sigma}$.

uncertainty regarding fundamentals $V[\theta|s]$, so the date-1 problem is to minimize

$$V[\theta|s] = A'\Gamma\hat{\Sigma}\Gamma' A$$

subject to the information processing constraint

$$\frac{1}{2}\ln\frac{|\Sigma|}{|\hat{\Sigma}|} \leq K \tag{29}$$

and the restriction that the matrix Σ_s is positive semi-definite i.e. the no-forgetting constraint. Note that since the variance-covariance matrices that enter the determinants in (29) are diagonal, the information processing constraint can be re-written as a sum. Furthermore, define the information-processing capacity devoted to learning about each of the underlying independent shocks as $k_i \equiv \frac{1}{2}\ln\frac{\Sigma_{ii}}{\hat{\Sigma}_{ii}}$. The information-processing constraint thus reduces to $\sum_i k_i \leq K$.

The matrix Γ is essentially a measure of the correlation structure between the risk factors. The rows of the loadings matrix give the loadings of each factor on all the shocks and the columns give the loadings of all the factors on each shock . The i^{th} row of the matrix Γ, denoted Γ_i, gives the loadings of the i^{th} risk factor on the independent shocks in the vector ϵ. Denote by Γ_j the j^{th} column of the matrix Γ, which gives the loadings of all the risk factors on the j^{th} shock. Define exposure to shock j as

$$E_j \equiv A'\Gamma_j = \sum_{i=1}^{n} \alpha_i \Gamma_{ij}$$

This measure of exposure captures the intuition that when the risk factors are correlated, the degree to which an underlying shock affects fundamentals will depend on the interaction between the observable exposure to the risk factors (captured by A) as well as on the loading of the risk factor on the underlying common shocks (captured by Γ). Thus, the effective exposure to a shock depends on the interaction between the observed exposure and the underlying correlation structure.

For ease of exposition, in what follows I consider and solve for the case in which $n = 2$. In this case the relevant effective exposure parameters are $E_j = \alpha_1\Gamma_{1j} + \alpha_2\Gamma_{2j},\ j = 1, 2$. The date-1 problem can be expressed as

$$\min_{k_1,k_2}\ V[\theta|s] = A'\Gamma\hat{\Sigma}\Gamma' A = \hat{\Sigma}_{11}E_1^2 + \hat{\Sigma}_{22}E_2^2$$

$$\text{s.t.}\quad \hat{\Sigma}_{ii} = \Sigma_{ii}e^{-2k_i},\ \ \sum k_i \leq K,\ \ 0 \leq k_i,\ \ i = 1, 2$$

The optimal capacity allocated to the factor-specific shocks is

$$k_1 = \begin{cases} 0 & \text{if} \quad \sqrt{\frac{\Sigma_{11} E_1^2}{\Sigma_{22} E_2^2}} < e^{-K} \\ \frac{1}{2}\left(K + \ln\sqrt{\frac{\Sigma_{11} E_1^2}{\Sigma_{22} E_2^2}}\right) & \text{if} \quad e^{-K} \leq \sqrt{\frac{\Sigma_{11} E_1^2}{\Sigma_{22} E_2^2}} \leq e^{K} \\ K & \text{if} \quad \sqrt{\frac{\Sigma_{11} E_1^2}{\Sigma_{22} E_2^2}} > e^{K} \end{cases} \tag{30}$$

and $k_2 = K - k_1$. Note that this solution for the attention allocated to the underlying shocks is similar in spirit to the attention allocated the independent risk factors that was discussed in Section 3.1. However, the implications in terms of magnitude of losses are different relative to baseline model.

Figure 7 plots the loss due to suboptimal action against exposure to factor 1, for varying degrees of correlation between the two risk factors. The example considers a one standard deviation shock to factor 1 ϵ_1, abstracts from factor 2 shocks ϵ_2 as well as from information shocks ϵ_s. The main result on the non-monotonic relationship between exposure and the loss due to suboptimal investment remains unchanged. However, the figure shows that relative to the zero correlation baseline, the loss due to suboptimal investment increases with the degree of correlation between the two factors (Appendix A.4 provides an algebraic derivation).

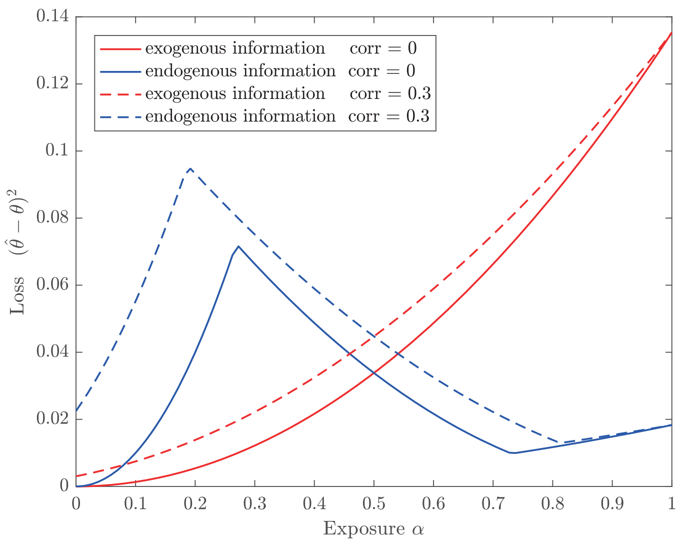


Figure 7. This example considers a one standard deviation factor 1 specific shock i.e. $\epsilon_1 = \Sigma_{11}$, abstracts from factor 2 shocks i.e. $\epsilon_2 = 0$ and information shocks i.e. $\epsilon_{s_1} = \epsilon_{s_2} = 0$. The parameter values are $K = 1$, $\tilde{\sigma}_{s_1} = \tilde{\sigma}_{s_2} = 0.75$, $\Sigma_{11} = \Sigma_{22} = 1$. Both risk factors are assumed to load equally on the underlying independent shocks: $\Gamma_{11} = \Gamma_{22} = 1$ and $\Gamma_{12} = \Gamma_{21} = 0.15$. The implied correlation between the two risk factors is 0.3

A higher degree of correlation between the two risk factors has two implications. On the one hand, correlation introduces learning complementarity benefits, as the firm can use information about one factor-specific shock to reduce uncertainty about both risk factors. On the other hand correlation also increases the effective exposure to shocks because now a shock specific to factor 1 will affects fundamentals not only through exposure to factor 1, but also through exposure to factor 2. The effect of complementary in learning is to shift the loss turning point along the x-axis. In the specific case illustrated in Figure 7, the loss function is shifted to the left relative to the zero-correlations baseline because the firm starts learning about the shock specific to factor 1 at a lower level of observable exposure (the observable exposure plotted on the x-axis is lower than the effective exposure which drives learning choices (30)). The effect of increased effective exposure to shocks is to shift the loss function upwards along the y-axis. The loss is higher relative to the zero-correlations baseline because effective exposure is higher than the observable exposure that is plotted on the x-axis. This result highlights that the apparently unexplained transmission of shocks i.e. transmission of shocks that is not explained by observable measure of exposure to shock can also occur because of underlying correlations between the risk factors driving fundamentals.

4.4 Extension: Endogenous Exposure

The baseline model considered the case in which exposures to risk factors are exogenous. In this section, I allow for exposure to be endogenous quantities determined in equilibrium. The basic result is that it is optimal for the firm to specialize in learning about one risk factor and to be more relatively exposed to that factor.

The solution strategy follows the same steps as in the baseline model in Section 3, except that at the first date, in addition to choosing the amount of information processing resources devoted to each risk factor, the firm also chooses the exposure to these factors. In particular, given the optimal action and the optimal factor-specific information processing capacity for any given exposure, the firm then chooses the optimal level of exposure by solving

$$\max_{\alpha} U_1 = \begin{cases} -\alpha^2\sigma_1^2 - (1-\alpha)^2\sigma_2^2 e^{-2K} & \text{if } k_1 = 0 \\ -2\alpha(1-\alpha)\sigma_1\sigma_2 e^{-K} & \text{if } k_1 = \frac{1}{2}\left(K + \ln\frac{\alpha\sigma_1}{(1-\alpha)\sigma_2}\right) \\ -\alpha^2\sigma_1^2 e^{-2K} - (1-\alpha)^2\sigma_2^2 & \text{if } k_1 = K \end{cases} \tag{31}$$

PROPOSITION 4. *It is optimal to be relatively more exposed to the risk factor that the firm learns about.*

$$\alpha^* = \begin{cases} \frac{\sigma_2^2}{\sigma_2^2+\sigma_1^2 e^{2K}} & if \quad k_1 = 0 \\ \frac{\sigma_2^2}{\sigma_2^2+\sigma_1^2 e^{-2K}} & if \quad k_1 = K \end{cases} \tag{32}$$

If the risk factors are ex-ante equally volatile, then the firm is indifferent between the two exposure allocations. If the risk factors are not ex-ante equally volatile, then it is optimal to be relatively more exposed to the factor that is ex-ante less volatile.

Proof. See Appendix A.5 for the full characterization of the equilibria.

Therefore, the firm prefers to be relatively more exposed to one factor, namely the factor it learns about. Indifference between exposure allocations arises if the risk factors are ex-ante equally volatile. Otherwise, it is optimal to learn about and to be relatively more exposed to the risk factor that is ex-ante less volatile. Thus, it is ex-ante optimal for the firm to specialize in learning about one risk factor and to be relatively more exposed to it. Note that for any finite level of capacity full exposure to a risk factor is never optimal and the optimal exposure is an interior solution. These ex-ante optimal exposure allocations will, however, expose the firm to the risk of incurring a higher loss due to suboptimal investment in the event that the risk factor that it is relatively less exposed to (and about which it does not learn) is hit by a shock.

The analytical expressions (32) reveal that the optimal level of exposure to factor 1 decreases with factor 1 uncertainty and increases with factor 2 uncertainty. Furthermore, optimal factor 1 exposure increases with total capacity K if the firm chooses to learn about factor 1 and it decreases with capacity K if the firm chooses to learn about factor 2.[28] Note that for any limited precision and ex-ante uncertain risk factor, unit exposure to one factor is not optimal i.e. for any $K < \infty$ and $\sigma_i^2 > 0,\ i = 1, 2 \Rightarrow 0 < \alpha^* < 1$.

5 Concluding Remarks

The financial crisis of 2007-2008 has highlighted the existence of a remarkable and poorly understood type of contagion whereby countries that were relatively less exposed to the crisis epicentre, the United States, were among the most severely affected. In other words, this crisis has

[28]This is because less capacity constrained agents are able to learn more and the levels of exposure for which specialized learning occurs are more extreme

shown that the impact of a shock can decrease with exposure to it. In this paper, I study how endogenous information choice affects decision-makers' reactions to shocks as a consequence the impact of those shocks. By linking information choice and learning behavior with exposure, the model I propose explains the puzzling observation that the impact of a shocks can decrease with exposure to it. The key mechanism in my model is that learning increases with exposure, such that the cost of being highly exposed to a shock is mitigated by the benefit of knowing it better. My model contributes to understanding observed cross-sectional and time-series patterns of contagion. In particular, my model explains how countries that are more exposed to a crisis can be less affected and why contagion is more likely to occur following unexpected crises.

References

Acharya, Viral V, and Tanju Yorulmazer, 2008, Information contagion and bank herding, *Journal of Money, Credit and Banking* 40, 215–231.

Ahnert, Toni, and Ali Kakhbod, 2017, Information choice and amplification of financial crises, *The Review of Financial Studies* 30, 2130–2178.

Allen, Franklin, Ana Babus, and Elena Carletti, 2012, Asset commonality, debt maturity and systemic risk, *Journal of Financial Economics* 104, 519–534.

Allen, Franklin, and Douglas Gale, 2000, Financial contagion, *Journal of Political Economy* 108, 1–33.

Allen, Franklin, and Douglas Gale, 2004, Financial fragility, liquidity, and asset prices, *Journal of the European Economic Association* 2, 1015–1048.

Allen, Franklin, and Douglas Gale, 2009, *Understanding financial crises* (Oxford University Press).

Alvarez, Fernando E, Francesco Lippi, and Luigi Paciello, 2015, Monetary shocks in models with inattentive producers, *The Review of Economic Studies* 83, 421–459.

Angeletos, George-Marios, and Alessandro Pavan, 2004, Transparency of information and coordination in economies with investment complementarities, *American Economic Review* 94, 91–98.

Bekaert, Geert, Michael Ehrmann, Marcel Fratzscher, and Arnaud Mehl, 2014, The global crisis and equity market contagion, *The Journal of Finance* 69, 2597–2649.

Benoit, Sylvain, Jean-Edouard Colliard, Christophe Hurlin, and Christophe Pérignon, 2017, Where the risks lie: A survey on systemic risk, *Review of Finance* 21, 109–152.

Calvo, Guillermo A, 2004, *Contagion in emerging markets: when Wall Street is a carrier* (Springer).

Calvo, Guillermo A, and Enrique G Mendoza, 2000, Rational contagion and the globalization of securities markets, *Journal of International Economics* 51, 79–113.

Dasgupta, Amil, 2004, Financial contagion through capital connections: A model of the origin and spread of bank panics, *Journal of the European Economic Association* 2, 1049–1084.

Didier, Tatiana, Paolo Mauro, and Sergio L Schmukler, 2008, Vanishing financial contagion?, *Journal of Policy Modeling* 30, 775–791.

Eichengreen, Barry, Andrew K Rose, and Charles Wyplosz, 1996, Contagious currency crises, Technical report, National Bureau of Economic Research.

Forbes, Kristin, 2012, The big C: identifying and mitigating contagion, in *Proceedings-Economic Policy Symposium-Jackson Hole*, 23–87, Federal Reserve Bank of Kansas City.

Forbes, Kristin J, 2004, The asian flu and russian virus: the international transmission of crises in firm-level data, *Journal of International Economics* 63, 59–92.

Forbes, Kristin J, and Roberto Rigobon, 2002, No contagion, only interdependence: measuring stock market comovements, *The Journal of Finance* 57, 2223–2261.

Freixas, Xavier, Bruno M Parigi, and Jean-Charles Rochet, 2000, Systemic risk, interbank relations, and liquidity provision by the central bank, *Journal of Money, Credit and Banking* 611–638.

Glick, Reuven, and Andrew K Rose, 1999, Contagion and trade: Why are currency crises regional?, *Journal of International Money and Finance* 18, 603–617.

Ilut, Cosmin, and Rosen Valchev, 2017, Economic agents as imperfect problem solvers, *Working Paper* .

Iyer, Rajkamal, and Jose-Luis Peydro, 2011, Interbank contagion at work: Evidence from a natural experiment, *The Review of Financial Studies* 24, 1337–1377.

Jotikasthira, Chotibhak, Christian Lundblad, and Tarun Ramadorai, 2012, Asset fire sales and purchases and the international transmission of funding shocks, *The Journal of Finance* 67, 2015–2050.

Kacperczyk, Marcin, Stijn Van Nieuwerburgh, and Laura Veldkamp, 2016, A rational theory of mutual funds' attention allocation, *Econometrica* 84, 571–626.

Kaminsky, Graciela L, Carmen M Reinhart, and Carlos A Vegh, 2003, The unholy trinity of financial contagion, *Journal of Economic Perspectives* 17, 51–74.

Karolyi, G Andrew, 2003, Does international financial contagion really exist?, *International Finance* 6, 179–199.

Kasperson, Roger E, Ortwin Renn, Paul Slovic, Halina S Brown, Jacque Emel, Robert Goble, Jeanne X Kasperson, and Samuel Ratick, 1988, The social amplification of risk: A conceptual framework, *Risk Analysis* 8, 177–187.

King, Mervyn A., and Sushil Wadhwani, 1990, Transmission of volatility between stock markets, *Review of Financial Studies* 3, 5–33.

Kodres, Laura E, and Matthew Pritsker, 2002, A rational expectations model of financial contagion, *The Journal of Finance* 57, 769–799.

Krishnamurthy, Arvind, 2010, Amplification mechanisms in liquidity crises, *American Economic Journal: Macroeconomics* 2, 1–30.

Luo, Yulei, 2008, Consumption dynamics under information processing constraints, *Review of Economic Dynamics* 11, 366–385.

Mackowiak, Bartosz, and Mirko Wiederholt, 2009, Optimal sticky prices under rational inattention, *American Economic Review* 99, 769–803.

Maćkowiak, Bartosz, and Mirko Wiederholt, 2015, Inattention to rare events, Technical report, ECB Working Paper.

Manconi, Alberto, Massimo Massa, and Ayako Yasuda, 2012, The role of institutional investors in propagating the crisis of 2007–2008, *Journal of Financial Economics* 104, 491–518.

Matějka, Filip, 2015, Rationally inattentive seller: Sales and discrete pricing, *The Review of Economic Studies* 83, 1125–1155.

Mondria, Jordi, 2010, Portfolio choice, attention allocation, and price comovement, *Journal of Economic Theory* 145, 1837.

Mondria, Jordi, and Climent Quintana-Domeque, 2013, Financial contagion and attention allocation, *The Economic Journal* 123, 429–454.

Nimark, Kristoffer, and Savitar Sundaresan, 2018, Inattention and belief polarization, *Working Paper* .

Paciello, Luigi, and Mirko Wiederholt, 2013, Exogenous information, endogenous information, and optimal monetary policy, *Review of Economic Studies* 81, 356–388.

Pavlova, Anna, and Roberto Rigobon, 2008, The role of portfolio constraints in the international propagation of shocks, *The Review of Economic Studies* 75, 1215–1256.

Pericoli, Marcello, and Massimo Sbracia, 2003, A primer on financial contagion, *Journal of Economic Surveys* 17, 571–608.

Renn, Ortwin, William J Burns, Jeanne X Kasperson, Roger E Kasperson, and Paul Slovic, 1992, The social amplification of risk: Theoretical foundations and empirical applications, *Journal of Social Issues* 48, 137–160.

Rigobon, Roberto, and Shang-Jin Wei, 2003, News, contagion, and anticipation understanding the differences between the Asian and the Latin American crises, *LAEBAWorking Paper Series* .

Rose, Andrew K, and Mark M Spiegel, 2010, Cross-country causes and consequences of the 2008 crisis: International Linkages and American Exposure, *Pacific Economic Review* 15, 340–363.

Rose, Andrew K, and Mark M Spiegel, 2011, Cross-country causes and consequences of the crisis: An update, *European Economic Review* 55, 309–324.

Sims, Christopher A, 2003, Implications of rational inattention, *Journal of Monetary Economics* 50, 665–690.

Slovic, Paul, 1987, Perception of risk, *Science* 236, 280–285.

Stevens, Luminita, 2015, Coarse pricing policies, Technical report, Federal Reserve Bank of Minneapolis.

Sundaresan, Savitar, 2018, Emergency preparedness: Rare events and the persistence of uncertainty, *Working Paper* .

Tutino, Antonella, 2013, Rationally inattentive consumption choices, *Review of Economic Dynamics* 16, 421–439.

Van Nieuwerburgh, Stijn, and Laura Veldkamp, 2010, Information acquisition and under-diversification, *The Review of Economic Studies* 77, 779–805.

Woodford, Michael, 2001, Imperfect common knowledge and the effects of monetary policy, Technical report, National Bureau of Economic Research.

Woodford, Michael, 2009, Information-constrained state-dependent pricing, *Journal of Monetary Economics* 56, S100–S124.

Yuan, Kathy, 2005, Asymmetric price movements and borrowing constraints: A rational expectations equilibrium model of crises, contagion, and confusion, *The Journal of Finance* 60, 379–411.

A Appendix

A.1 Appendix: Uncertainty

Recalling the relation between the information processing capacity allocated to a risk factor and reduction in uncertainty achieved by it implied by the entropy constraint in (10), factor-specific posterior uncertainty can be determined to be

$$\hat{\sigma}_1^2 = \begin{cases} \sigma_1^2 & \text{if} \quad \frac{\alpha\sigma_1}{(1-\alpha)\sigma_2} < e^{-K} \\ \frac{1-\alpha}{\alpha}\sigma_1\sigma_2 e^{-K} & \text{if} \quad e^{-K} \leq \frac{\alpha\sigma_1}{(1-\alpha)\sigma_2} \leq e^{K} \\ \sigma_1^2 e^{-2K} & \text{if} \quad \frac{\alpha\sigma_1}{(1-\alpha)\sigma_2} > e^{K} \end{cases} \tag{33}$$

and $\hat{\sigma}_2^2 = \sigma_1^2\sigma_2^2\hat{\sigma}_1^{-2}e^{-2K}$. Note that factor-specific risk is effectively reduced through learning. Relative to a full information benchmark (i.e. $\hat{\sigma}_1^2 = \sigma_1^2$), risk is lower if exposure is relatively high and/or prior uncertainty is relatively high (i.e. $\hat{\sigma}_1^2 < \sigma_1^2$ if $\alpha \geq 0.5$ and/or $\sigma_1 \geq \sigma_1$); this can also be interpreted as risks being under-estimated.

Proof. Proposition (1)

Uncertainty under the exogenous information benchmark is given by

$$\tilde{\sigma}_B^2 = \alpha^2\tilde{\sigma}_1^2 + (1-\alpha)^2\tilde{\sigma}_2^2 \tag{34}$$

Under the endogenous information model uncertainty $\hat{\sigma}^2 = \alpha^2\hat{\sigma}_1^2 + (1-\alpha)^2\hat{\sigma}_2^2$ is given by

$$\hat{\sigma}^2 = \begin{cases} \alpha^2\sigma_1^2 + (1-\alpha)^2\sigma_2^2 e^{-2K} & \text{if} \quad \frac{\alpha\sigma_1}{(1-\alpha)\sigma_2} < e^{-K} \\ 2\alpha(1-\alpha)\sigma_1\sigma_2 e^{-K} & \text{if} \quad e^{-K} \leq \frac{\alpha\sigma_1}{(1-\alpha)\sigma_2} \leq e^{K} \\ \alpha^2\sigma_1^2 e^{-2K} + (1-\alpha)^2\sigma_2^2 & \text{if} \quad \frac{\alpha\sigma_1}{(1-\alpha)\sigma_2} > e^{K} \end{cases} \tag{35}$$

Contrast comparative statics with respect to exposure under the two models (at the interior optimum for endogenous information choice)

$$\frac{\partial\tilde{\sigma}_B^2}{\partial\alpha} = 2\alpha\tilde{\sigma}_1^2 - 2(1-\alpha)\tilde{\sigma}_2^2 > 0 \quad \text{if} \quad \alpha > 0.5 \tag{36}$$

$$\frac{\partial\hat{\sigma}^2}{\partial\alpha} = 2(1-2\alpha)\sigma_1\sigma_2 e^{-K} < 0 \quad \text{if} \quad \alpha > 0.5 \tag{37}$$

□

Analytically, we can contrast the results under the learning model with the exogenous information benchmark results by setting equal the information processing capacity in the two models (plug K when the capacity constraint is binding i.e. $\frac{1}{2}\ln\frac{\sigma_1^2\sigma_2^2}{\tilde{\sigma}_1^2\tilde{\sigma}_2^2} = K \Rightarrow e^{-K} = \frac{\tilde{\sigma}_1\tilde{\sigma}_2}{\sigma_1\sigma_2}$ into result (35)). This is informative of the factor-specific risk reduction entailed by learning, or the learning-adjusted risk of a factor and its contribution to overall uncertainty.

$$\hat{\sigma}^2 = \begin{cases} \alpha^2\sigma_1^2 + (1-\alpha)^2\tilde{\sigma}_2^2\frac{\tilde{\sigma}_1^2}{\sigma_1^2} & \text{if } \quad \frac{\alpha\sigma_1}{(1-\alpha)\sigma_2} < \frac{\tilde{\sigma}_1\tilde{\sigma}_2}{\sigma_1\sigma_2} \\ \alpha^2\tilde{\sigma}_1^2\frac{1-\alpha}{\alpha}\frac{\tilde{\sigma}_2}{\tilde{\sigma}_1} + (1-\alpha)^2\tilde{\sigma}_2^2\frac{\alpha}{1-\alpha}\frac{\tilde{\sigma}_1}{\tilde{\sigma}_2} & \text{if } \quad \frac{\tilde{\sigma}_1\tilde{\sigma}_2}{\sigma_1\sigma_2} \le \frac{\alpha\sigma_1}{(1-\alpha)\sigma_2} \le \frac{\sigma_1\sigma_2}{\tilde{\sigma}_1\tilde{\sigma}_2} \\ \alpha^2\tilde{\sigma}_1^2\frac{\tilde{\sigma}_2^2}{\sigma_2^2} + (1-\alpha)^2\sigma_2^2 & \text{if } \quad \frac{\alpha\sigma_1}{(1-\alpha)\sigma_2} > \frac{\sigma_1\sigma_2}{\tilde{\sigma}_1\tilde{\sigma}_2} \end{cases} \tag{38}$$

These analytical results confirm the message conveyed in Figure 2 that relative to the equal capacity, exogenous information benchmark, uncertainty is lower when agents are allowed to optimally allocate their information resources across risk factor exposures. It can be shown that the corner solutions are always smalller than the benchmark solutions. For the first benchmark solution: $\alpha^2\tilde{\sigma}_1^2 + (1-\alpha)^2\tilde{\sigma}_2^2 > \alpha^2\sigma_1^2 + (1-\alpha)^2\tilde{\sigma}_2^2\frac{\tilde{\sigma}_1^2}{\sigma_1^2} \Leftrightarrow \tilde{\sigma}_2^2 > \frac{\alpha^2}{(1-\alpha)^2}\sigma_1^2$ which is true in light of the condition for obtaining the corner solution, which can be written as $\frac{\tilde{\sigma}_1}{\sigma_1}\tilde{\sigma}_2 > \frac{\alpha}{1-\alpha}\sigma_1$. Since $\tilde{\sigma}_2 \ge \frac{\tilde{\sigma}_1}{\sigma_1}\tilde{\sigma}_2\left(> \frac{\alpha}{1-\alpha}\sigma_1\right) \Rightarrow \tilde{\sigma}_2^2 > \frac{\alpha^2}{(1-\alpha)^2}\sigma_1^2$. Similarly, for the second corner solution we have that $\alpha_1^2\tilde{\sigma}_1^2 + \alpha_2^2\tilde{\sigma}_2^2 > \alpha^2\tilde{\sigma}_1^2\frac{\tilde{\sigma}_2^2}{\sigma_2^2} + \alpha_2^2\sigma_2^2 \Leftrightarrow \frac{\alpha}{1-\alpha}\frac{1}{\sigma_2} > \frac{1}{\tilde{\sigma}_1}$. This follows from the corner solution condition which can be re-written as $\frac{\alpha}{1-\alpha}\frac{1}{\sigma_2} \ge \frac{\sigma_2}{\tilde{\sigma}_1\tilde{\sigma}_2}$ and the fact that $\frac{1}{\tilde{\sigma}_1}\frac{\sigma_2}{\tilde{\sigma}_2} > \frac{1}{\tilde{\sigma}_1}$. For the interior solution we have that $\alpha^2\tilde{\sigma}_1^2 + (1-\alpha)^2\tilde{\sigma}_2^2 \ge 2\alpha_1\alpha_2\tilde{\sigma}_1\tilde{\sigma}_2$, which holds because $(\alpha_1\tilde{\sigma}_1 - \alpha_2\tilde{\sigma}_2)^2 \ge 0$.

Importantly, they shed further light into the mechanism behind the observed dynamics. Relative to the benchmark, when exposure to a factor is very low and there is no learning about it (corner solution), the effective or learning-adjusted risk of that factor is higher ($\sigma_1^2 \ge \tilde{\sigma}_1^2$), while the effective risk of the other factor is lower ($\tilde{\sigma}_1^2\frac{\tilde{\sigma}_1^2}{\sigma_1^2} \le \tilde{\sigma}_1^2$). At the interior optimum, factor-specific effective risk is higher if factor-specific exposure is relatively low (all else equal $\tilde{\sigma}_1^2\frac{1-\alpha}{\alpha} \ge \tilde{\sigma}_1^2$ if $\alpha < 1-\alpha$) and it is lower if factor-specific exposure is relatively high (all else equal $\tilde{\sigma}_1^2\frac{1-\alpha}{\alpha} \le \tilde{\sigma}_1^2$ if $\alpha > 1-\alpha$).

A.2 Appendix: Shock Transmission Mechanism

Proof. Proposition (2)

Define $\gamma_i = \frac{\hat{\sigma}_i^2}{\sigma_i^2}$ so that we can re-write the conditional mean value of fundamentals as

$$\hat{\theta} \equiv E[\theta|s_1, s_2] \quad = \quad \alpha\left[\mu_1 + (1-\gamma_1)s_1\right] + (1-\alpha)\left[\mu_2 + (1-\gamma_2)s_2\right] \tag{39}$$

where

$$\gamma_1 = \begin{cases} 1 & \text{if } \frac{\alpha\sigma_1}{(1-\alpha)\sigma_2} < e^{-K} \\ \frac{1-\alpha}{\alpha}\frac{\sigma_2}{\sigma_1}e^{-K} & \text{if } e^{-K} \leq \frac{\alpha\sigma_1}{(1-\alpha)\sigma_2} \leq e^{K} \\ e^{-2K} & \text{if } \frac{\alpha\sigma_1}{(1-\alpha)\sigma_2} > e^{K} \end{cases} \quad (40)$$

$$\gamma_2 = \gamma_1^{-1}e^{-2K} \quad (41)$$

Recalling that $\theta = \alpha(\mu_1 + \epsilon_1) + (1-\alpha)(\mu_2 + \epsilon_2)$ and $s_i = \epsilon_i + \epsilon_{s_i}$, the loss due to suboptimal action is given by

$$L \equiv (\hat{\theta} - \theta)^2 = \Big(\alpha\left[(1-\gamma_1)\epsilon_{s_1} - \gamma_1\epsilon_1\right] + (1-\alpha)\left[(1-\gamma_2)\epsilon_{s_2} - \gamma_2\epsilon_2\right]\Big)^2 \quad (42)$$

such that the interpretation of the parameters $\gamma_i,\ i = 1, 2$ is that of the weight assigned to the shocks affecting fundamentals. The loss function under the exogenous information benchmark takes the same form except that the weight coefficients are different $\tilde{\gamma}_i \equiv \frac{\tilde{\sigma}_i^2}{\sigma_i^2},\ i = 1, 2$

Introduce the exogenous information, equal capacity benchmark

$$L_B = (\hat{\theta}_B - \theta)^2 = (\alpha\left[(1-\tilde{\gamma}_1)\epsilon_{s_1} - \tilde{\gamma}_1\epsilon_1\right] + (1-\alpha)\left[(1-\tilde{\gamma}_2)\epsilon_{s_2} - \tilde{\gamma}_2\epsilon_2\right])^2 \quad (43)$$

where $\tilde{\gamma}_1 \equiv \frac{\tilde{\sigma}_1^2}{\sigma_1^2}$, $\tilde{\gamma}_2 \equiv \frac{\tilde{\sigma}_2^2}{\sigma_2^2}$, $\gamma_1 \equiv \frac{\hat{\sigma}_1^2}{\sigma_1^2} = \frac{1-\alpha}{\alpha}\sigma_1\sigma_2 e^{-K}$, $\gamma_2 \equiv \frac{\hat{\sigma}_2^2}{\sigma_2^2} = \frac{\alpha}{1-\alpha}\sigma_1\sigma_2 e^{-K}$.

Define the shock transmission mechanism as the loss induced by a shock

$$\frac{\partial L}{\partial \epsilon_1} = -2\alpha\gamma_1\Big(\alpha\left[(1-\gamma_1)\epsilon_{s_1} - \gamma_1\epsilon_1\right] + (1-\alpha)\left[(1-\gamma_2)\epsilon_{s_2} - \gamma_2\epsilon_2\right]\Big) \quad (44)$$

$$\frac{\partial L_B}{\partial \epsilon_1} = -2\alpha\tilde{\gamma}_1\Big(\alpha\left[(1-\tilde{\gamma}_1)\epsilon_{s_1} - \tilde{\gamma}_1\epsilon_1\right] + (1-\alpha)\left[(1-\tilde{\gamma}_2)\epsilon_{s_2} - \tilde{\gamma}_2\epsilon_2\right]\Big) \quad (45)$$

We are interested in how the transmission mechanism varies with exposure α. Let us recall that the weight coefficients in the endogenous learning model are functions of exposure $\gamma_1 \equiv \frac{\hat{\sigma}_1^2}{\sigma_1^2} = \frac{1-\alpha}{\alpha}\sigma_1\sigma_2 e^{-K}$, $\gamma_2 \equiv \frac{\hat{\sigma}_2^2}{\sigma_2^2} = \frac{\alpha}{1-\alpha}\sigma_1\sigma_2 e^{-K}$. Abstracting from factor 2 effects by setting $\epsilon_2 = \epsilon_{s_2} = 0$, and using the result that $\frac{\partial\gamma_1}{\partial\alpha} = -\frac{1}{\alpha(1-\alpha)}\gamma_1$, we have that

$$\frac{\partial^2 L}{\partial\epsilon_1\partial\alpha} = 2\alpha\gamma_1\left[\left(\frac{2\alpha-1}{1-\alpha} - \frac{2\alpha}{1-\alpha}\gamma_1\right)\epsilon_{s_1} - \frac{2\alpha}{1-\alpha}\gamma_1\epsilon_1\right] \quad (46)$$

$$\frac{\partial^2 L_B}{\partial\epsilon_1\partial\alpha} = -4\alpha\tilde{\gamma}_1\left[(1-\tilde{\gamma}_1)\epsilon_{s_1} - \tilde{\gamma}_1\epsilon_1\right] \quad (47)$$

In the limiting case in which the signal noise is zero $\epsilon_{s_1} = 0$ (the signal is perfectly informative), it is clear that $\frac{\partial^2 L}{\partial \epsilon_1 \partial \alpha} < 0$ (the transmission mechanism decreases with exposure) and $\frac{\partial^2 L_B}{\partial \epsilon_1 \partial \alpha} > 0$ (the transmission mechanism increases with exposure). These results hold more generally if the signal is informative i.e. the signal noise is sufficiently small $\epsilon_{s_1} < \frac{\gamma_1}{1-\gamma_1}\epsilon_1$.

□

A.3 Appendix: Shock Anticipation

Proof. Proposition (3)

The date-1 utility function (18) is a continuous piecewise function that is increasing in state capacity. It is useful to distinguish between three types of equilibria: (i) the equilibrium capacity allocation across factors has the property $k_1 = 0$, (ii) the equilibrium capacity allocation across factors has the property $k_1 = K$, and (iii) the equilibrium capacity allocation across factors has the property $k_1 = \frac{1}{2}\left(K + \ln \frac{\alpha\sigma_1}{(1-\alpha)\sigma_2}\right)$. Substituting the constraint (20) into the objective function (19) and differentiating with respect to K_r yields first-order condition in the three set of equilibria

$$\begin{cases} 2p_r(1-\alpha)^2\sigma_2^2 e^{-2K_r} - 2p_n(1-\alpha)^2\sigma_2^2 e^{-2(\mathcal{K}-K_r)} & \text{if } \ k_1 = 0 \\ 2p_r\alpha(1-\alpha)\sigma_1\sigma_2 e^{-K_r} - 2p_n\alpha(1-\alpha)\sigma_1\sigma_2 e^{-(\mathcal{K}-K_r)} & \text{if } \ k_1 = \frac{1}{2}\left(K + \ln \frac{\alpha\sigma_1}{(1-\alpha)\sigma_2}\right) \\ 2p_r\alpha^2\sigma_1^2 e^{-2K_r} - 2p_n\alpha^2\sigma_1^2 e^{-2(\mathcal{K}-K_r)} & \text{if } \ k_1 = K \end{cases} \tag{48}$$

Solving for K_r yields

$$K_r = \begin{cases} \frac{1}{4}\left(2\mathcal{K} + \ln \frac{p_r}{p_n}\right) & \text{if } \ k_1 = 0 \text{ or } k_1 = K \\ \frac{1}{2}\left(\mathcal{K} + \ln \frac{p_r}{p_n}\right) & \text{if } \ k_1 = \frac{1}{2}\left(K + \ln \frac{\alpha\sigma_1}{(1-\alpha)\sigma_2}\right) \end{cases} \tag{49}$$

Imposing the no-forgetting constraint $0 \leq K_r$ and noting that $K_r \leq \mathcal{K}$ is always satisfied because $\ln \frac{p_r}{p_n} < 0$ when $p_r < p_n$, the optimal capacity allocation across states when the agent can only learn about one state (when a corner solution is obtained for capacity allocation across factors) is given by

$$K_r = \begin{cases} 0 & \text{if } \ \frac{p_r}{p_n} < e^{-2\mathcal{K}} \\ \frac{1}{4}\left(2\mathcal{K} + \ln \frac{p_r}{p_n}\right) & \text{if } \ \frac{p_r}{p_n} \geq e^{-2\mathcal{K}} \end{cases} \tag{50}$$

$$K_n = \mathcal{K} - K_r \tag{51}$$

and the optimal information processing capacity allocated to the rare state of nature when the agent can only learn about one state (when an interior solution is obtained for capacity allocation across factors) is given by

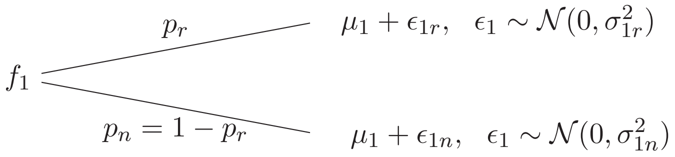
$$K_r = \begin{cases} 0 & \text{if } \frac{p_r}{p_n} < e^{-\mathcal{K}} \\ \frac{1}{2}\left(\mathcal{K} + \ln\frac{p_r}{p_n}\right) & \text{if } \frac{p_r}{p_n} \geq e^{-\mathcal{K}} \end{cases} \tag{52}$$

$$K_n = \mathcal{K} - K_r \tag{53}$$

Note that if the agent only learns about one as opposed to both factors, it is more likely that the agent will dedicate information processing resources to the rare state (since $e^{-2\mathcal{K}} < e^{-\mathcal{K}}$) but the overall capacity allocated to the state is smaller (since $\frac{1}{4}\left(2\mathcal{K} + \ln\frac{p_r}{p_n}\right) < \frac{1}{2}\left(\mathcal{K} + \ln\frac{p_r}{p_n}\right)$).

□

If the shock variance is different in the two states



then the capacity allocation is

$$K_r = \begin{cases} 0 & \text{if } \frac{p_r\sigma_{1r}}{p_n\sigma_{1n}} < e^{-\mathcal{K}} \\ \frac{1}{2}\left(\mathcal{K} + \ln\frac{p_r\sigma_{1r}}{p_n\sigma_{1n}}\right) & \text{if } e^{-\mathcal{K}} \leq \frac{p_r\sigma_{1r}}{p_n\sigma_{1n}} \leq e^{\mathcal{K}} \\ \mathcal{K} & \text{if } \frac{p_r\sigma_{1r}}{p_n\sigma_{1n}} > e^{\mathcal{K}} \end{cases}$$

A.4 Appendix: Correlated Risks

The shock-specific posterior uncertainty implied by the capacity allocation (30) is given by

$$\hat{\Sigma}_{11} = \begin{cases} \Sigma_{11} & \text{if } \sqrt{\frac{\Sigma_{11}E_1^2}{\Sigma_{22}E_2^2}} < e^{-K} \\ \sqrt{\frac{\Sigma_{11}\Sigma_{22}E_2^2}{E_1^2}}e^{-K} & \text{if } e^{-K} \leq \sqrt{\frac{\Sigma_{11}E_1^2}{\Sigma_{22}E_2^2}} \leq e^{K} \\ \Sigma_{11}e^{-2K} & \text{if } \sqrt{\frac{\Sigma_{11}E_1^2}{\Sigma_{22}E_2^2}} > e^{K} \end{cases} \tag{54}$$

and $\hat{\Sigma}_{22} = \Sigma_{11}\Sigma_{22}e^{-2K}\hat{\Sigma}_{11}^{-1}$

Consequently, the uncertainty about fundamentals is given by

$$V[\theta|s] = \hat{\Sigma}_{11}E_1^2 + \hat{\Sigma}_{22}E_2^2 = \begin{cases} \Sigma_{11}E_1^2 + \Sigma_{22}E_2^2 e^{-2K} & \text{if } \sqrt{\frac{\Sigma_{11}E_1^2}{\Sigma_{22}E_2^2}} < e^{-K} \\ 2\sqrt{\Sigma_{11}\Sigma_{22}E_1^2E_2^2}e^{-K} & \text{if } e^{-K} \leq \sqrt{\frac{\Sigma_{11}E_1^2}{\Sigma_{22}E_2^2}} \leq e^{K} \\ \Sigma_{11}E_1^2e^{-2K} + \Sigma_{22}E_2^2 & \text{if } \sqrt{\frac{\Sigma_{11}E_1^2}{\Sigma_{22}E_2^2}} > e^{K} \end{cases}$$

The loss due to suboptimal investment is

$$\begin{aligned} L \equiv (E[\theta|s] - \theta)^2 &= (A'E[f|s] - A'f)^2 = [A'(\mu + \Gamma(I - \hat{\Sigma}\Sigma^{-1})s - A'(\mu + \Gamma\epsilon)]^2 \\ &= [A'(\mu + \Gamma(I - \hat{\Sigma}\Sigma^{-1})(\epsilon + \epsilon_s) - A'(\mu + \Gamma\epsilon)]^2 \\ &= [A'\Gamma(I - \hat{\Sigma}\Sigma^{-1})\epsilon_s + A'\Gamma(I - \hat{\Sigma}\Sigma^{-1})\epsilon - A'\Gamma\epsilon]^2 \\ &= [A'\Gamma(I - \hat{\Sigma}\Sigma^{-1})\epsilon_s - A'\Gamma\hat{\Sigma}\Sigma^{-1}\epsilon]^2 \end{aligned}$$

Abstracting from information shocks i.e. $\epsilon_s = 0$ we have that

$$L = (A'\Gamma\hat{\Sigma}\Sigma^{-1}\epsilon)^2$$

A.5 Appendix: Endogenous Exposure

The objective function (31) is a continuous piecewise function, which is concave in exposure when a corner solution is obtained for capacity allocation, and convex in exposure when an interior solution is obtained. Hence, an interior solution is obtained for optimal exposure if a corner solution is obtained for information choice, and a corner solution is obtained for optimal exposure if an interior solution is obtained for information choice. The first-order condition is

$$\frac{\partial U_1}{\partial \alpha} = \begin{cases} -2\alpha\sigma_1^2 + 2(1-\alpha)\sigma_2^2 e^{-2K} & \text{if } k_1 = 0 \\ -2(1-2\alpha)\sigma_1\sigma_2 e^{-K} & \text{if } k_1 = \frac{1}{2}\left(K + \ln\frac{\alpha\sigma_1}{(1-\alpha)\sigma_2}\right) \\ -2\alpha\sigma_1^2 e^{-2K} + 2(1-\alpha)\sigma_2^2 & \text{if } k_1 = K \end{cases} \tag{55}$$

It is useful to distinguish between three types of equilibria: (i) the equilibrium capacity allocation has the property $k_1 = 0$, (ii) the equilibrium capacity allocation has the property $k_1 = K$, and (iii) the equilibrium capacity allocation has the property $k_1 = \frac{1}{2}\left(K + \ln\frac{\alpha\sigma_1}{(1-\alpha)\sigma_2}\right)$.

Equilibria (i) and (ii) represents situations in which the firm is only able to learn about one factor. In these situations, it is optimal to be relatively more exposed to the factor the firm learns about, and the optimal point of exposure is the one at which the learning adjusted risk exposures are equal. More specifically, if $k_1 = 0$ optimal exposure is implied by $\alpha\sigma_1^2 = (1-\alpha)\sigma_2^2 e^{-2K}$. If $k_1 = K$ optimal exposure is implied by $\alpha\sigma_1^2 e^{-2K} = (1-\alpha)\sigma_2^2$.

$$\alpha^* = \begin{cases} \frac{\sigma_2^2}{\sigma_2^2+\sigma_1^2 e^{2K}} & \text{if} \quad k_1 = 0 \\ \frac{\sigma_2^2}{\sigma_2^2+\sigma_1^2 e^{-2K}} & \text{if} \quad k_1 = K \end{cases} \tag{56}$$

Equilibrium (iii) represents a situation in which the parameter values are such that the firm can learn about both factors. In this case, it is optimal to be as exposed as possible to one factor, where the maximum level of exposure to a factor is implied by the condition for learning about both risks $e^{-K} \leq \frac{\alpha\sigma_1}{(1-\alpha)\sigma_2} \leq e^K$. To determine which risk factor it is optimal to relatively more exposed to, I compare the expected utility of each corner solution for exposure. The maximum exposure or loading to factor 1 is obtained when $\frac{\alpha\sigma_1}{(1-\alpha)\sigma_2} = e^K$, which implies $\overline{\alpha} = \frac{\sigma_2}{\sigma_2+\sigma_1 e^{-K}}$ and the utility of being relatively more exposed to factor 1 is $U_1(\overline{\alpha}) = -2\left(\frac{\sigma_1\sigma_2}{\sigma_1+\sigma_2 e^K}\right)^2$. The maximum exposure or loading to factor 2 is obtained when $\frac{\alpha\sigma_1}{(1-\alpha)\sigma_2} = e^{-K}$, which implies minimum factor 1 exposure $\underline{\alpha} = \frac{\sigma_2}{\sigma_2+\sigma_1 e^K}$ and the utility of being relatively more exposed to factor 2 is $U_1(\underline{\alpha}) = -2\left(\frac{\sigma_1\sigma_2}{\sigma_1 e^K+\sigma_2}\right)^2$. In the case of symmetric equilibria whereby the two factors are ex-ante equally volatile, the firm will be indifferent between the two exposure allocations i.e. $U_1(\overline{\alpha}) = U_1(\underline{\alpha})$ if $\sigma_1 = \sigma_2$. However, in the case of non-symmetric equilibria, it is optimal to be more exposed to the less volatile risk factor i.e. $U_1(\overline{\alpha}) > U_1(\underline{\alpha})$ if $\sigma_1 < \sigma_2$ hence $\alpha^* = \overline{\alpha}$, and $U_1(\overline{\alpha}) < U_1(\underline{\alpha})$ if $\sigma_1 > \sigma_2$ hence $\alpha^* = \underline{\alpha}$.

$$\alpha^* = \begin{cases} \frac{\sigma_2}{\sigma_2+\sigma_1 e^K} & \text{if} \quad k_1 = \frac{1}{2}\left(K + \ln\frac{\alpha\sigma_1}{(1-\alpha)\sigma_2}\right) \text{ and } \sigma_1 > \sigma_2 \\ \frac{\sigma_2}{\sigma_2+\sigma_1 e^{-K}} & \text{if} \quad k_1 = \frac{1}{2}\left(K + \ln\frac{\alpha\sigma_1}{(1-\alpha)\sigma_2}\right) \text{ and } \sigma_1 < \sigma_2 \\ \frac{\sigma_2}{\sigma_2+\sigma_1 e^K} \text{ or } \frac{\sigma_2}{\sigma_2+\sigma_1 e^{-K}} & \text{if} \quad k_1 = \frac{1}{2}\left(K + \ln\frac{\alpha\sigma_1}{(1-\alpha)\sigma_2}\right) \text{ and } \sigma_1 = \sigma_2 \end{cases} \tag{57}$$

However, since the firm is not constrained to learn about both risk factors (i.e. to be in equilibria of the type (iii)), it will optimally choose to learn about one factor only and to be relatively more exposed to the factor it learns about. This follows from the fact that the expected utility associated with the optimal levels of exposure (57) obtained in equilibrium (iii) is lower than the utility associated with the optimal levels of exposure (56) that are obtained in equilibria (i) and (ii). Indifference between these latter exposure allocations (56) arises if the risk factors are ex-ante equally volatile, but it is otherwise optimal to be relatively more exposed to the factor that is ex-ante less volatile.