Inattention to Rare Events*

Bartosz Maćkowiak
European Central Bank and CEPR

Mirko Wiederholt
Northwestern University

First version: March 2011. This version: April 2012

Abstract

Why were people so unprepared for the global financial crisis, the European debt crisis, and the Fukushima nuclear accident? To address this question, we study a model in which agents make state-contingent plans – think about actions in different contingencies – subject to the constraint that agents can process only a limited amount of information. The model predicts that agents are unprepared in a state when the state has a low probability, when the optimal action in that state is uncorrelated with the optimal action in normal times, and when actions are strategic complements. We compare the equilibrium allocation of attention to the efficient allocation of attention, and characterize analytically the conditions under which society would be better off if agents thought more carefully about optimal actions in rare events.

Keywords: rare events, disasters, rational inattention, efficiency. (*JEL*: D83, E58, E60).

*Maćkowiak: European Central Bank, Kaiserstrasse 29, 60311 Frankfurt am Main, Germany (e-mail: bartosz.mackowiak@ecb.int); Wiederholt: Department of Economics, Northwestern University, 2001 Sheridan Road, Evanston, IL 60208 (e-mail: m-wiederholt@northwestern.edu). The views expressed in this paper are solely those of the authors and do not necessarily reflect the views of the European Central Bank. We thank for comments Gadi Barlevy, Marco Bassetto, Christian Hellwig, Jean-Paul L'Huillier, Cosmin Ilut, Guido Lorenzoni, Alisdair McKay, Kristoffer Nimark, Laura Veldkamp, and seminar and conference participants at CERGE-EI, Cologne Workshop on Macroeconomics, ECB, Federal Reserve Bank of Chicago, Goethe University Frankfurt, Penn State, NBER Summer Institute 2011, and SED 2011. We thank Giovanni Nicolò for excellent research assistance. Mirko Wiederholt thanks the Federal Reserve Bank of Minneapolis for hospitality.

1 Introduction

In recent years the world experienced several events with disastrous consequences: the global financial crisis, the European sovereign debt crisis, and the Fukushima nuclear accident. These events have in common that people were unprepared for them. How come virtually no one had thought through what to do if an investment bank like Lehman Brothers collapses; what to do if several governments in the euro area are close to default; or what to do if an earthquake and tsunami disable the cooling system of a nuclear reactor? Using a model with informationally constrained agents we ask: Why were people so unprepared for these events? Under which circumstances will people be unprepared again in the future? Would a social planner want people to be better prepared for rare events?

Understanding when people are likely to be unprepared is important. Had people been prepared to take good action in each of these events, these events would have unfolded much less dramatically. However, being well prepared for each contingency is costly and it is unclear, from an ex-ante perspective, if a social planner would want people to be better prepared for rare events. Only if the equilibrium is inefficient a policy encouraging decision-makers to think what to do in rare events can improve welfare.

To address these questions formally, we study a model in which agents make state-contingent plans – think about actions in different contingencies – subject to an information-processing constraint. Agents commit today to actions in the different states tomorrow. This assumption captures the idea that decision-making takes time and once the state realizes agents have to act quickly. Therefore, agents need to plan ahead. Agents have a prior over what the optimal action is in each state and they can process additional information. However, agents can process only a *finite amount* of additional information. Subject to this constraint, agents decide how carefully to think about the optimal action in normal times and in unusual times.

We embed this decision problem into a setup with a continuum of agents where an agent's payoff in each state depends on the agent's own action, the mean action in the population, and a fundamental. The uncertainty about the optimal action in a given state is due to uncertainty about the fundamental and the mean action in the population in that state. We derive the equilibrium allocation of attention and compare it to the efficient allocation of attention.

The model makes three main predictions. First, agents think less carefully about the optimal action in unusual times than in normal times, and thus the mean squared difference between the optimal action and the actual action is larger in unusual times than in normal times. In the model, agents equate the probability-weighted expected loss due to suboptimal actions across states. Thus, if the probability of unusual times is 0.1 percent, the expected loss due to suboptimal action is three orders of magnitude larger in unusual times than in normal times. Agents take on average worse actions in unusual times.

Second, the correlation of optimal actions across states matters for the quality of actions taken in different states. If the optimal action in normal times and the optimal action in unusual times are correlated, thinking about the best action in normal times also improves actions in unusual times. In contrast, if the optimal action in normal times and the optimal action in unusual times are independent, thinking about the best action in normal times fails to improve actions in unusual times.

Third, strategic complementarity in actions makes the allocation of attention more extreme. If actions are strategic complements (that is, the optimal action in a state is increasing in the mean action in the population in that state), the fact that other agents are not thinking carefully about the optimal action in unusual times reduces the incentive for an individual agent to think about the optimal action in unusual times. As a result, the larger the degree of strategic complementarity in actions, the less agents think about the optimal action in unusual times. In fact, for a sufficiently high degree of strategic complementarity in actions, agents are completely inattentive to the rare event.

We look at the recent events from the perspective of the model. Why was the Tokyo Electric Power Company so unprepared for the event that an earthquake and tsunami disable the cooling system of this company's Fukushima Dai-ichi nuclear power plant? The model proposes the following answer: Humans have a limited ability to process information and therefore cannot prepare well for every contingency. A magnitude 9.0 earthquake and tsunami is a low probability event; thinking carefully about how to run a nuclear power plant efficiently in normal times fails to improve actions in times when an earthquake and tsunami disable the plant's cooling system; and strategic complementarity in actions amplifies the effect of a low probability on the allocation of attention. We think the strategic complementarity in actions in this case arose because of relative performance evaluation. Companies tend to be punished less if they fail in times when other companies are failing too.

Why were policy-makers, financial institutions, and academics so unprepared for the collapse of Lehman Brothers? Collapse of one of the most important U.S. financial institutions seemed unlikely a priori; and thinking carefully about how to regulate financial institutions in normal times or how to fine-tune open market operations to achieve a desired level of the federal funds rate helps little when confronted with an imminent collapse of Lehman Brothers. Furthermore, we believe there is strategic complementarity in actions: policy-makers have to push a common agenda in order to get any solution adopted; the management of a financial institution is punished less if it fails in times when other financial institutions are failing too; academics like to work on topics that other academics are working on, because then those other academics are more likely to be interested in the work. This strategic complementarity makes agents even more inattentive to rare events.

Would a planner want people to be more prepared for rare events? To answer this question, we study the following planner problem. The planner maximizes ex-ante utility of agents. The planner chooses the agents' attention allocation, subject to the agents' information processing constraint. We ask: Does the equilibrium allocation of attention equal the efficient allocation of attention? In other words, would society be ex-ante better off if agents allocated their attention differently than they do in equilibrium?

We focus on the case when the economy is efficient under perfect information, that is, inefficiencies, if any, arise due to agents' limited attention. We prove that a simple condition on the payoff function governs the relationship between the equilibrium allocation of attention and the efficient allocation of attention. If this condition is satisfied, society cannot do better by providing incentives for agents to allocate their attention differently; for example, by passing a law requiring nuclear power plants to have a precise plan of action in the case of a tsunami. In contrast, if the condition is not satisfied, the equilibrium allocation of attention differs from the efficient allocation of attention, and society can in principle do better by providing incentives for agents to allocate their attention differently. We also characterize the direction of the inefficiency. We show when the planner wants agents to think more carefully about optimal actions in rare events, and when the opposite is true.

This paper makes contact with three recent strands of literature. It is related to the literature on rational inattention building on Sims (2003).[1] The first main difference to the existing literature on rational inattention is the application. We study how agents make state-contingent plans. Since agents commit to a contingent plan and have a limited ability to process information, the probability of a state affects the quality of the action taken in that state. The second main difference to the existing literature on rational inattention is that we compare the equilibrium allocation of attention to the efficient allocation of attention. To the best of our knowledge, we are the first to do so.

Our work is also related to the literature on rare disasters. See for example Barro (2006), Barro, Nakamura, Steinsson, and Ursua (2010), Gabaix (2010), and Gourio (2010). This literature investigates the implications of rare disasters for asset prices and business cycles when agents act perfectly in a rare event. In contrast, we model agents as acting imperfectly in a rare event and investigate how much incentive agents have to prepare for a rare event. If people had been prepared to take good action in historical rare adverse events, these events would have unfolded less dramatically and perhaps would not be called "disasters" today.

We also make contact with the literature on the efficient use of information. Angeletos and Pavan (2007) study an economy with a continuum of agents in which each agent observes a noisy private and public signal. The precision of the two signals is exogenous. Actions are a linear function of the two signals and Angeletos and Pavan (2007) refer to the coefficients on the two signals as the "use of information." They then compare the equilibrium use of information to the efficient use of information, where the latter is defined as the one that maximizes ex-ante utility. We find that the condition governing the relationship between the equilibrium and efficient use of information in their model with exogenous signal precision equals the condition governing the relationship between the equilibrium and efficient allocation of attention in our model with endogenous signal precision.

Hellwig and Veldkamp (2009) study a beauty contest model with information choice. The payoff of an agent depends on his own action, a fundamental, and the mean action in the population. Agents choose the number of signals that they acquire concerning the fundamental. The main differences to our model are that there is only one regime, agents face a fixed cost per signal

[1] For theoretical papers, see Sims (2003, 2006, 2010), Luo (2008), Maćkowiak and Wiederholt (2009, 2010), Van Nieuwerburgh and Veldkamp (2009, 2010), Woodford (2009), Matejka (2010a,b), Mondria (2010), Paciello (2010), Paciello and Wiederholt (2011), Tutino (2011), and Yang (2011). For empirical papers, see Maćkowiak, Moench, and Wiederholt (2009), Kacperczyk, Van Nieuwerburgh, and Veldkamp (2011), Melosi (2011), and Coibion and Gorodnichenko (2011).

(instead of a limited amount of attention), and their payoff function is less general. An unpublished working paper version of Hellwig and Veldkamp (2009) contains a subsection studying efficiency of information acquisition for a very particular quadratic payoff function.[2] For this payoff function, there exists an equilibrium which is ex ante efficient. This result is consistent with our result concerning efficiency of the equilibrium allocation of attention, because the payoff function that they assume is a special case of the sufficient condition for ex-ante efficiency that we identify.[3]

The rest of the paper is organized as follows. Section 2 presents the model. Section 3 derives the equilibrium allocation of attention. Section 4 compares the equilibrium allocation of attention to the efficient allocation of attention. Section 5 concludes.

2 Model

We study an economy with a continuum of agents indexed by $i \in [0, 1]$ and discrete time indexed by $t = 0, 1, 2, \ldots$. Each period the economy is in one of two regimes. The regime follows a two-state Markov chain. For simplicity, the regime is i.i.d. over time. Let p_n denote the probability of being in regime n. Both regimes have positive probability, that is, $p_1 > 0$ and $p_2 > 0$. It is often helpful to think of p_1 as being smaller than p_2 and therefore of regime one as unusual times and regime two as normal times. Initially we assume that agents know the probabilities of the two regimes. Later we introduce Bayesian learning about these probabilities.

Every period each agent commits to a state-contingent plan for the next period. This assumption captures the idea that decision-making takes time and once the state realizes agents have to act quickly. Therefore, agents need to plan ahead. The contingent plan that agent i commits to in period $t-1$ for period t is denoted $a_{i,t} = (a_{i,t,1}, a_{i,t,2}) \in \mathbb{R}^2$ where $a_{i,t,n}$ denotes the action that agent i will take at time t in regime n. Let $\Psi^{n,t}$ denote the cumulative distribution function for action $a_{i,t,n}$ in the cross-section of the population.

The payoff of agent i at time t in regime n is given by $U^n(a_{i,t,n}, a_{t,n}, z_{t,n})$ where $a_{i,t,n}$ is the own action, $a_{t,n} \equiv \int a_{i,t,n} d\Psi^{n,t}(a_{i,t,n})$ is the mean action in the population, and $z_{t,n}$ is an exogenous

[2] The payoff of an agent is a linear combination of the squared distance between the own action and the fundamental and the squared distance between the own action and the mean action in the population.

[3] Llosa and Venkateswaran (2011) extend the efficiency result in the working paper version of Hellwig and Veldkamp (2009) to a somewhat more general payoff function and study in detail a price setting application.

fundamental. The superscript n indicates that the payoff function may differ across regimes. For tractability, we assume that the payoff function is quadratic

$$\begin{aligned} U^n\left(a_{i,t,n}, a_{t,n}, z_{t,n}\right) &= U^n\left(0,0,0\right) + U^n_{a_i} a_{i,t,n} + U^n_a a_{t,n} + U^n_z z_{t,n} \\ &+ \frac{U^n_{a_i a_i}}{2} a^2_{i,t,n} + \frac{U^n_{aa}}{2} a^2_{t,n} + \frac{U^n_{zz}}{2} z^2_{t,n} \\ &+ U^n_{a_i a} a_{i,t,n} a_{t,n} + U^n_{a_i z} a_{i,t,n} z_{t,n} + U^n_{az} a_{t,n} z_{t,n}. \end{aligned} \tag{1}$$

This assumption can also be viewed as a second-order approximation of any twice differentiable function with these three arguments. Furthermore, we assume that the payoff function is concave in its first argument ($U^n_{a_i a_i} < 0$), the fundamental affects the payoff-maximizing action ($U^n_{a_i z} \neq 0$), and the degree of strategic complementarity or substitutability is below one ($-1 < U^n_{a_i a}/U^n_{a_i a_i} < 1$). In the following, we often exploit the fact that the payoff function can be expressed as[4]

$$U^n\left(a_{i,t,n}, a_{t,n}, z_{t,n}\right) = U^n\left(a^*_{i,t,n}, a_{t,n}, z_{t,n}\right) + \frac{U^n_{a_i a_i}}{2}\left(a_{i,t,n} - a^*_{i,t,n}\right)^2,$$

where $a^*_{i,t,n}$ denotes the optimal action at time t in regime n

$$a^*_{i,t,n} = -\frac{U^n_{a_i}}{U^n_{a_i a_i}} - \frac{U^n_{a_i a}}{U^n_{a_i a_i}} a_{t,n} - \frac{U^n_{a_i z}}{U^n_{a_i a_i}} z_{t,n}.$$

Finally, without loss of generality, we assume that the coefficients on $a_{t,n}$ and $z_{t,n}$ in the equation for the optimal action sum to one.[5] Defining $\varphi_n \equiv -U^n_{a_i}/U^n_{a_i a_i}$, $\gamma_n \equiv -U^n_{a_i a}/U^n_{a_i a_i}$ and $\delta_n \equiv -U^n_{a_i a_i}/2$, the last two equations then become

$$U^n\left(a_{i,t,n}, a_{t,n}, z_{t,n}\right) = U^n\left(a^*_{i,t,n}, a_{t,n}, z_{t,n}\right) - \delta_n\left(a_{i,t,n} - a^*_{i,t,n}\right)^2, \tag{2}$$

with

$$a^*_{i,t,n} = \varphi_n + \gamma_n a_{t,n} + \left(1-\gamma_n\right) z_{t,n}. \tag{3}$$

For simplicity, the vector of fundamentals $z_t = \left(z_{t,1}, z_{t,2}\right)$ is i.i.d. over time. Agents have the common prior belief that the vector of fundamentals is i.i.d. over time and that the fundamental

[4] To obtain this result, compute a Taylor expansion of U^n around $a^*_{i,t,n}$ and notice that the first derivative of U^n with respect to $a_{i,t,n}$ evaluated at $a^*_{i,t,n}$ equals zero and the second derivative of U^n with respect to $a_{i,t,n}$ equals $U^n_{a_i a_i}$.

[5] If this assumption is not satisfied, one can always redefine the fundamental $z_{t,n}$ by multiplying it with a constant to ensure that this assumption is satisfied.

in regime one and the fundamental in regime two are normally distributed with mean zero and covariance matrix Σ, that is, $z_t = (z_{t,1}, z_{t,2}) \sim i.i.d.N(0, \Sigma)$. There is prior uncertainty about the fundamental and therefore about the optimal action in each regime. Furthermore, the fundamentals in the two regimes are not perfectly correlated, that is, Σ is non-singular.

Agents can process information before committing to a plan. However, agents can process only a *limited amount* of information. Processing information about the optimal actions in the two regimes in the next period is modeled as receiving a noisy signal concerning the fundamentals in the two regimes in the next period

$$s_{i,t-1} = \begin{pmatrix} z_{t,1} \\ z_{t,2} \end{pmatrix} + \begin{pmatrix} \varepsilon_{i,t-1,1} \\ \varepsilon_{i,t-1,2} \end{pmatrix},$$

where the noise $(\varepsilon_{i,t-1,1}, \varepsilon_{i,t-1,2})$ is independent of the fundamentals, independent across individuals and over time, and normally distributed with mean zero and covariance matrix Λ. Let $\Omega = \Sigma - \Sigma(\Sigma + \Lambda)^{-1}\Sigma$ denote the posterior covariance matrix of z_t after receiving $s_{i,t-1}$. Following Sims (2003), we model the fact that humans have a limited ability to process information as a constraint on uncertainty reduction, where uncertainty is measured by entropy. That is, each agent faces the following constraint on uncertainty reduction:

$$\frac{1}{2}\log_2\left(\frac{|\Sigma|}{|\Omega|}\right) \leq \kappa,$$

where $|\Sigma|$ denotes the determinant of the prior covariance matrix of z_t and $|\Omega|$ denotes the determinant of the posterior covariance matrix of z_t after receiving $s_{i,t-1}$. The parameter $\kappa > 0$ indexes the ability of an agent to process information, where a larger κ means an agent can process more information and can thus reduce uncertainty by more.

Subject to the information-processing constraint, each agent decides how carefully to think about the optimal action in regime one and the optimal action in regime two. Agents aim to maximize the expected payoff in the next period. Formally, agent i solves in period $t-1$

$$\max_{\Lambda} \sum_{n=1}^{2} p_n E\left[U^n\left(a_{i,t,n}, a_{t,n}, z_{t,n}\right)\right], \tag{4}$$

subject to

$$a_{i,t,n} = E\left[\varphi_n + \gamma_n a_{t,n} + (1-\gamma_n) z_{t,n} | s_{i,t-1}\right], \tag{5}$$

$$s_{i,t-1} = \begin{pmatrix} z_{t,1} \\ z_{t,2} \end{pmatrix} + \begin{pmatrix} \varepsilon_{i,t-1,1} \\ \varepsilon_{i,t-1,2} \end{pmatrix}, \tag{6}$$

and

$$\frac{1}{2}\log_2\left(\frac{|\Sigma|}{|\Omega|}\right) \leq \kappa, \tag{7}$$

and the restriction that Λ is a positive semidefinite matrix. Objective (4) is the expected payoff in the next period. Equation (5) states that the agent will commit to the best plan given his or her posterior. Equation (6) is the signal. Inequality (7) is the constraint stating that agents can process only a limited amount of information.

The covariance matrix of noise Λ and the posterior covariance matrix of the fundamentals Ω have no subscripts i and t. The reason is that the solution to problem (4)-(7) is the same for each agent i and every period t. This also means that the equilibrium is symmetric and that agents only have to solve this problem once.[6]

In problem (4)-(7) the informational constraint depends only on the prior covariance matrix of the fundamentals Σ and the posterior covariance matrix of the fundamentals Ω. This setup formalizes the idea that learning is the mental process of absorbing available information. All information required for the agent to take the optimal actions in both regimes is in principle available. The agent, due to limited cognitive ability, cannot attend to all this information and hence cannot prepare a perfect action plan for each contingency.[7] Furthermore, once the agent has formed a conditional expectation of the optimal action there is no physical cost of implementing the action. We think that this setup captures the critical feature of the recent events: people had failed to think through what action to take in certain contingencies, while information about what action to take was available and the physical cost of implementing good action was negligible.

3 The equilibrium allocation of attention

In this section, we derive the equilibrium allocation of attention in the model presented above. In the following section, we study the efficient allocation of attention.

[6] Note that we have assumed that signals are normally distributed. One can show that Gaussian signals are optimal given the quadratic objective, the Gaussian prior, and the constraint on entropy reduction. See Sims (2006).

[7] It is useful to distinguish this setup from a setup in which learning is the discovery of new information via a research-and-development type activity.

3.1 The role of probabilities

To build intuition, we start with the special case of no strategic complementarity in actions and independence of the fundamentals across regimes. Formally, we assume that the payoff function satisfies $\gamma_1 = \gamma_2 = 0$ and the prior covariance matrix of the fundamentals satisfies $\Sigma_{12} = 0$. These assumptions are relaxed below.

When there is no strategic complementarity in actions, the payoff-maximizing action of an agent in a regime depends only on the fundamental in the regime, $a^*_{i,t,n} = \varphi_n + z_{t,n}$. The best plan of an agent given his or her posterior equals the conditional expectation of the payoff-maximizing action, $a_{i,t,n} = \varphi_n + E\left[z_{t,n}|s_{i,t-1}\right]$. Thus, the expected loss in payoff in regime n due to suboptimal action in regime n is given by

$$\begin{aligned} E\left[U^n\left(a_{i,t,n}, a_{t,n}, z_{t,n}\right)\right] - E\left[U^n\left(a^*_{i,t,n}, a_{t,n}, z_{t,n}\right)\right] &= -\delta_n E\left[\left(a_{i,t,n} - a^*_{i,t,n}\right)^2\right] \\ &= -\delta_n E\left[\left(z_{t,n} - E\left[z_{t,n}|s_{i,t-1}\right]\right)^2\right] \\ &= -\delta_n E\left[E\left[\left(z_{t,n} - E\left[z_{t,n}|s_{i,t-1}\right]\right)^2 |s_{i,t-1}\right]\right] \\ &= -\delta_n \Omega_{nn}. \end{aligned} \tag{8}$$

The last line follows from the fact that the fundamental and the signal have a multivariate normal distribution and thus the conditional variance of the fundamental is the same for all signal realizations. Equation (8) states that the expected loss in regime n due to suboptimal action in regime n depends on the conditional variance of the fundamental in regime n and the cost of a mistake in regime n, governed by the parameter δ_n.

When fundamentals are independent across regimes, it is optimal to think independently about the fundamental (optimal action) in regime one and the fundamental (optimal action) in regime two. This result is proved in the next subsection. Formally, the optimal covariance matrix of noise in the signal Λ is diagonal, and thus not only the prior covariance matrix of the fundamentals Σ is diagonal but also the posterior covariance matrix of the fundamentals Ω is diagonal. As a result, the information-processing constraint (7) reduces to

$$\frac{1}{2}\log_2\left(\frac{\Sigma_{11}}{\Omega_{11}}\right) + \frac{1}{2}\log_2\left(\frac{\Sigma_{22}}{\Omega_{22}}\right) \leq \kappa, \tag{9}$$

where Σ_{nn} and Ω_{nn} are the prior and posterior variance of the fundamental in regime n. Let

$$\kappa_n \equiv \frac{1}{2}\log_2\left(\frac{\Sigma_{nn}}{\Omega_{nn}}\right) \tag{10}$$

denote the attention devoted to regime n. When no attention is devoted to a regime, the posterior variance of the fundamental equals the prior variance of the fundamental. When positive attention is devoted to a regime, the posterior is less diffuse than the prior.

Agents decide how carefully to think about the optimal actions in the two regimes. Using equations (8)-(10), the attention problem (4)-(7) can be expressed as

$$\max_{(\kappa_1,\kappa_2)\in\mathbb{R}^2_+}\left[-\sum_{n=1}^{2}p_n\delta_n\Omega_{nn}\right], \tag{11}$$

subject to

$$\Omega_{nn}=\Sigma_{nn}2^{-2\kappa_n}, \tag{12}$$

and

$$\kappa_1+\kappa_2\leq\kappa. \tag{13}$$

Objective (11) is the expected loss in payoff due to suboptimal action in the next period, equation (12) follows directly from the definition of attention devoted to regime n, and inequality (13) is the attention constraint. The fact that the agent chooses the attention devoted to the two regimes rather than the precision of the two signals is just a change of variables. The unique solution to this problem is

$$\kappa_1=\begin{cases}\kappa & \text{if } \sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}}\geq 2^{\kappa}\\ \frac{1}{2}\kappa+\frac{1}{2}\log_2\left(\sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}}\right) & \text{if } \sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}}\in\left[2^{-\kappa},2^{\kappa}\right]\\ 0 & \text{if } \sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}}\leq 2^{-\kappa}\end{cases}. \tag{14}$$

The extent to which a decision-maker thinks about the optimal action in regime one is increasing in the agent's information-processing ability, κ, the probability of the regime, p_1/p_2, the cost of a mistake in the regime, δ_1/δ_2, and the prior variance of the fundamental in the regime, Σ_{11}/Σ_{22}. In discussing the optimal allocation of attention, let us begin with the case of $p_1/p_2=\delta_1/\delta_2=\Sigma_{11}/\Sigma_{22}=1$. In this case, the decision-maker decides to allocate his or her attention equally across the two regimes. Next, starting from that situation, reduce the probability of regime one, $p_1<p_2$. The decision-maker now decides to think less about the optimal action in unusual times than about the optimal action in normal times. The reason is that the benefit of thinking about a contingency that is less likely to occur is smaller. Furthermore, if unusual times are sufficiently unlikely, the decision-maker decides to not think at all about the optimal action in those unusual times. Finally, for the decision-maker to think more about the optimal action in unusual times than about the

optimal action in normal times, the cost of a mistake has to be sufficiently larger in unusual times than in normal times ($\delta_1 > \delta_2$) or the agent has to be sufficiently more uncertain about the optimal action in unusual times than about the optimal action in normal times ($\Sigma_{11} > \Sigma_{22}$).

The extent to which agents think about the optimal action in a regime affects the quality of action in the regime. It follows from equations (12)-(14) that at an interior solution ($0 < \kappa_1 < \kappa$), the expected loss in payoff in unusual times due to suboptimal action in those times equals

$$\delta_1 \Omega_{11} = \delta_1 \Sigma_{11} \left(2^{\kappa} \sqrt{\frac{p_1 \delta_1 \Sigma_{11}}{p_2 \delta_2 \Sigma_{22}}} \right)^{-1},$$

while the expected loss in payoff in normal times due to suboptimal action in those times equals

$$\delta_2 \Omega_{22} = \delta_2 \Sigma_{22} \left(2^{\kappa} \sqrt{\frac{p_2 \delta_2 \Sigma_{22}}{p_1 \delta_1 \Sigma_{11}}} \right)^{-1}.$$

Combining these two equations yields

$$p_1 \delta_1 \Omega_{11} = p_2 \delta_2 \Omega_{22}.$$

In words, agents allocate their attention so as to equate the probability-weighted expected loss due to suboptimal action across regimes. This implies that

$$\frac{\delta_1 \Omega_{11}}{\delta_2 \Omega_{22}} = \frac{1}{\frac{p_1}{p_2}}. \tag{15}$$

The expected loss due to suboptimal action in unusual times divided by the expected loss due to suboptimal action in normal times equals one over the odds of unusual times. If unusual times have a relative probability of 0.1 percent, the expected loss due to suboptimal action is *one thousand times* larger in unusual times than in normal times. Observing that agents take good actions in normal times does not imply that agents will take good actions in unusual times! The quality of action in a regime, measured by the expected loss in payoff due to suboptimal action, is inversely related to the probability of the regime.

Let us use the results of this subsection to think about the global financial crisis. For concreteness, we focus on the defining moment of that crisis which came when policy-makers and bankers met at the Federal Reserve Bank of New York on the weekend of September 12-14, 2008 to consider the future of Lehman Brothers. We think of regime one, unusual times, as the regime in which an investment bank like Lehman fails. First, we argue that the failure of an investment bank like

Lehman was a low-probability event but it was *not* an unthinkable, zero-probability event. Second, we argue that because the probability that an investment bank like Lehman fails was small, policy-makers failed to think through how the economy would function and what action they should take in that event.

Figure 1 plots the probability of default on one-year senior debt of Lehman, based on credit default swap (CDS) premia.[8] Prior to August 9, 2007, the day on which the interbank market first froze up, the probability of default by Lehman was 0.0016 on average. An event with a probability between 0.1 and 0.2 percent is a low-probability event but it is not unthinkable. Moreover, the probability of Lehman's bankruptcy actually rose steadily after August 9, 2007. That probability was about 0.01, on average, between August 9, 2007 and March 2008, the month in which the Federal Reserve helped broker the purchase of Bear Stears by JPMorgan Chase.[9] When the meeting in New York began on September 12, 2008, a reasonable estimate of the probability of Lehman's failure was one-in-ten.

Timothy Geithner, then president of the Federal Reserve Bank of New York, asked one of the working groups formed at the meeting to "put foam on the runway" and "be prepared to do *something*".[10] The decision to form this working group and the words used by Geithner show the meeting's participants were aware that Lehman's collapse was possible and one needed to prepare for it. However, the meeting participants chose to focus on thinking about what to do in the event that Lehman would *not* fail. In particular, inspired by their recent experience with Bear Stears and their experience in 1998 with Long-Term Capital Management, the meeting participants searched for buyers for Lehman and thought about the details of a possible purchase agreement, including any policy support for buyers. In the end, Lehman filed for bankruptcy on September 15, a worldwide panic in financial markets ensued, and the day after policy-makers intervened to

[8] To produce Figure 1, we took from Bloomberg CDS premia on one-year senior debt of Lehman Brothers, at daily frequency, from the beginning of July 2003 to the last trading day, September 12, 2008. We computed the probability of default, plotted in Figure 1, from this data assuming risk neutrality and a recovery rate equal to 8.625 percent, the actual recovery rate reported in Singh and Spackman (2009). The dataset had occasional missing observations which accounts for the missing values in Figure 1.

[9] The date March 11, 2008 in Figure 1 refers to the day on which the Federal Reserve set up Term Securities Lending Facility, a program in which investment banks swap securities with the Fed for Treasury bonds. See Brunnermeier (2009) for a chronology of the global financial crisis in 2007-2008.

[10] We quote Geithner after Wessel (2009, p.17), with emphasis added.

support the biggest victim of the panic, American International Group (AIG). This intervention was seen as a policy reversal and an acknowledgement that policy-makers were unprepared for Lehman's failure. Ben Bernanke, Chairman of the Federal Reserve, admitted as much when he testified to Congress that the policy-makers had failed to foresee the consequences the collapse of Lehman would have on other financial institutions, in particular AIG.[11] Several actions could have been taken by the policy-makers, from carefully going over the balance sheets of financial institutions to giving the Federal Reserve – or some other agency – the authority to wind down a systemically important financial institution in an orderly way, with taxpayers' backing. However, the probability of failure of an investment bank like Lehman was small and, as predicted by the model, the decision-makers failed to think through what action they should take if an investment bank like Lehman fails.

Next, consider the European sovereign debt crisis. We think of regime one, unusual times, as the regime in which a government of a euro area country is on the brink of default. Figure 2 plots the probability of default on one-year senior sovereign debt of Greece, Ireland, Portugal, Spain, and Italy, based on CDS premia.[12] The similarity to Figure 1 is striking. Prior to September 12, 2008, the probability of default by each of these nations was between 0.001 and 0.002, on average. An event with a probability between 0.1 and 0.2 percent is a low-probability event but it is not unthinkable. Furthermore, the probability of default by each of these nations actually rose steadily after September 12, 2008. The probability of default by each country that received official assistance (Greece, Ireland, and Portugal) was about one-in-ten at the time when the assistance was announced.[13]

[11] We quote from the website of the Board of Governors of the Federal Reserve System from Bernanke's testimony to Congress on September 24, 2008, with emphasis added: "While perhaps manageable in itself, Lehman's default was combined with the *unexpectedly rapid* collapse of AIG, which *together* contributed to the development last week of extraordinarily turbulent conditions in global financial markets."

[12] To produce Figure 2, we took from Datastream CDS premia on one-year senior sovereign debt of Greece, Ireland, Portugal, Spain, and Italy, at daily frequency, from the beginning of July 2003 through the end of July 2011. We computed the probability of default, plotted in Figure 2, from this data assuming risk neutrality and a recovery rate equal to 21.5 percent, the actual recovery rate in the case of Greece reported by *Financial Times* in its March 20, 2012 issue.

[13] The dates May 2, 2010, November 28, 2010, and May 3, 2011 in Figure 2 refer to the days when official assistance was agreed for Greece, Ireland, and Portugal, respectively. See the website of the ECB for a chronology of the European sovereign debt crisis.

Although it was not unthinkable that a euro area government would find itself on the brink of default, we are not aware of any planning for that event by policy-makers at least until the spring of 2010 when the fiscal stress in Greece triggered turbulence in euro area financial markets. During the design and after the establishment of European Monetary Union, the rhetoric of "market discipline" was used to argue that if yields on debt of a euro area country were to rise, that country would implement fiscal reform, and yields would fall again. This argument misses the point that yields on debt of a euro area country can rise – relative to yields on debt of other euro area countries – only if creditors attach a non-zero probability to default by that country. Thus every time yield spreads or CDS premia are non-zero, the following question arises: What would a default by a European government, or multiple European governments simultaneously, look like? We are not aware of any policy-maker having thought through this question. Some preparation for default began only after the first official assistance package for Greece was agreed in May 2010. This package included no provisions for default – a policy response later reversed when the second official assistance package for Greece was agreed in February 2012 including provisions for default. The model suggests that the policy-makers failed over the years to prepare for a default by a European nation because the probability of that event was small.

Finally, consider the Fukushima nuclear accident. We think of regime one, unusual times, as the regime in which an earthquake and tsunami disable the cooling system of a nuclear reactor in Japan. We think of staff in Japanese nuclear power companies, Japanese nuclear safety regulators, and Japanese government officials as the relevant decision-makers. The earthquake that struck off the coast of Japan on March 11, 2011 was a 9.0-magnitude earthquake. The earthquake cut all off-site power supply to the Fukushima Dai-ichi nuclear power plant. The ensuing tsunami waves knocked out all of the plant's emergency diesel generators apart from one. Hence, it was the combination of earthquake and tsunami that caused the power failure at the Fukushima Dai-ichi plant disabling the plant's cooling system. This combination of earthquake and tsunami was a low-probability event, but it was not an unthinkable event. The so-called Jogan earthquake of A.D. 869 knocked down a castle and sent a tsunami wave more than two miles inland in the same region. This fact was brought up in a meeting of a commission evaluating the safety of the Fukushima Dai-ichi nuclear power plant in June 2009. Several officials of Tokyo Electric Power Company (Tepco)

– the company owning and operating the plant – attended this meeting.[14]

After the earthquake and tsunami hit the plant, workers at the plant tried to avoid a catastrophe. The most severe problem was that the fuel rods inside the reactors were overheating, which caused a buildup of steam and hydrogen inside the reactor buildings. This meant a possible explosion. After communicating with Tepco officials in Tokyo and the prime minister of Japan, the workers on site decided to vent reactor Unit 1 to reduce pressure. Unfortunately, the emergency manual did not contain *any* instructions on how to vent the reactor in the absence of electricity. Throughout the night, the workers tried to figure out ways to vent the reactor in the absence of electricity.[15] At about 2:30pm on March 12, the operators confirmed a decrease in pressure inside the reactor, providing some indication that venting was starting to work. Shortly thereafter, a hydrogen explosion destroyed the Unit 1 reactor building.[16]

What happened was *not* that the decision-makers had thought carefully what action to take if the cooling system were disabled and then judged that action to be too costly to implement. Instead the facts were: a low-but-non-zero-probability event occurred (the combination of earthquake and tsunami disabled the cooling system), and the emergency manual contained no instructions on how to vent a nuclear reactor in the absence of electricity. These facts are consistent with the model. The model predicts that if $p_1 > 0$ is sufficiently small then $\kappa_1 = 0$.

3.2 Strategic complementarity in actions

The payoff-maximizing action of an agent may depend on the actions of other agents. In this subsection we allow for this possibility, that is, we relax the assumption that $\gamma_1 = \gamma_2 = 0$. For ease of exposition, we assume initially that the degree of strategic complementarity in actions is the same in unusual times and in normal times, $\gamma_1 = \gamma_2 \equiv \gamma$. When $\gamma > 0$ actions are strategic complements. When $\gamma < 0$ actions are strategic substitutes. It turns out that raising the degree of strategic complementarity in actions in both regimes makes the equilibrium allocation of attention more extreme. Whichever regime agents were paying *more* attention to in the ab-

[14] See, for example, the March 23, 2011, issue of *The Washington Post*.

[15] See the program "One year later, inside Japan's nuclear meltdown" that National Public Radio broadcast on February 28, 2012. The program is available at www.npr.org.

[16] See the report by the International Atomic Energy Agency (IAEA) international fact finding expert mission after the Fukushima nuclear accident. The report is available at www.iaea.org.

sence of strategic complementarity, agents are paying *even more* attention to in the presence of strategic complementarity. Whichever regime agents were paying *less* attention to in the absence of strategic complementarity, agents are paying *even less* attention to in the presence of strategic complementarity.

The following proposition characterizes equilibrium. The first part of the proposition states that when fundamentals are independent across regimes, it is optimal to think independently about the optimal action in unusual times and the optimal action in normal times. The second part of the proposition characterizes the equilibrium allocation of attention for any value of $\gamma \in (-1, 1)$.

Proposition 1 *Assume that the fundamental in regime one and the fundamental in regime two are independent (i.e., Σ is diagonal). Consider equilibria of the form $a_{t,n} = \psi_n + \phi_n z_{t,n}$ where ψ_n and ϕ_n are coefficients. In this case, each agent decides to receive independent signals about the fundamental in regime one and the fundamental in regime two (i.e., the equilibrium Λ is diagonal), and the information-processing constraint (7) reduces to*

$$\underbrace{\frac{1}{2}\log_2\left(\frac{\Sigma_{11}}{\Omega_{11}}\right)}_{\kappa_1} + \underbrace{\frac{1}{2}\log_2\left(\frac{\Sigma_{22}}{\Omega_{22}}\right)}_{\kappa_2} \leq \kappa,$$

where Σ_{nn} and Ω_{nn} denote the prior and the posterior variance of the fundamental in state n and κ_n denotes the uncertainty reduction about the fundamental in state n. Next, assume that $\gamma_1 = \gamma_2 \equiv \gamma \in (-1, 1)$. If the parameters κ and γ satisfy $2^\kappa > \frac{\gamma}{1-\gamma}$, the equilibrium is unique and the equilibrium attention allocated to thinking about the optimal action in regime one equals

$$\kappa_1 = \begin{cases} \kappa & \text{if } \sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}} \geq (1-\gamma)\,2^\kappa + \gamma 2^{-\kappa} \\ \frac{1}{2}\kappa + \frac{1}{2}\log_2(x) & \text{if } \sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}} \in \left[\frac{1}{(1-\gamma)2^\kappa + \gamma 2^{-\kappa}}, (1-\gamma)\,2^\kappa + \gamma 2^{-\kappa}\right] \\ 0 & \text{if } \sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}} \leq \frac{1}{(1-\gamma)2^\kappa + \gamma 2^{-\kappa}} \end{cases}, \tag{16}$$

where

$$x \equiv \frac{\sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}} - \frac{\gamma}{1-\gamma}2^{-\kappa}}{1 - \sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}}\frac{\gamma}{1-\gamma}2^{-\kappa}}. \tag{17}$$

Furthermore, for any parameters κ and γ, the set of equilibria is given by the following results: (1) $\kappa_1 = \kappa$ is an equilibrium if $\sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}} \geq (1-\gamma)\,2^\kappa + \gamma 2^{-\kappa}$; (2) $\kappa_1 = 0$ is an equilibrium if $\sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}} \leq \frac{1}{(1-\gamma)2^\kappa + \gamma 2^{-\kappa}}$; (3) $\kappa_1 = \frac{1}{2}\kappa + \frac{1}{2}\log_2(x)$ is an equilibrium if $(1-\gamma)\,2^\kappa + \gamma 2^{-\kappa} \geq 1$, $\sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}} \in \left[\frac{1}{(1-\gamma)2^\kappa + \gamma 2^{-\kappa}}, (1-\gamma)\,2^\kappa + \gamma 2^{-\kappa}\right]$ and $\sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}}\frac{\gamma}{1-\gamma} < 2^\kappa$; or $(1-\gamma)\,2^\kappa + \gamma 2^{-\kappa} \leq 1$,

$\sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}} \in \left[(1-\gamma)\,2^{\kappa}+\gamma 2^{-\kappa}, \frac{1}{(1-\gamma)2^{\kappa}+\gamma 2^{-\kappa}}\right]$ *and* $\sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}}\frac{\gamma}{1-\gamma} > 2^{\kappa}$*; and (4) any* $\kappa_1 \in [0,\kappa]$ *is an equilibrium if* $\sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}} = \frac{\gamma}{1-\gamma}2^{-\kappa} = 1$.

Proof. See Appendix A. ■

Consider first the case when the equilibrium is unique, $2^{\kappa} > \gamma/(1-\gamma)$. Equations (16)-(17) show that raising the degree of strategic complementarity in actions makes the equilibrium allocation of attention more extreme (if possible, that is, if the allocation of attention in the absence of strategic complementarity was not already a corner solution). Figure 3 illustrates this result by depicting equilibrium attention as a function of $\sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}}$ for parameters κ and γ satisfying $2^{\kappa} > \gamma/(1-\gamma)$. In Figure 3, $\gamma = 0$ denotes the case of no strategic complementarity in actions, $\gamma >> 0$ denotes a value of γ close to the value at which $2^{\kappa} = \gamma/(1-\gamma)$, and $\gamma > 0$ denotes a value of γ between these two extremes. Pick any value of $\sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}}$ with the property that $\sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}} \neq 1$. For example, suppose the probability of unusual times is small and therefore $\sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}} < 1$. In the absence of strategic complementarity in actions (i.e., $\gamma = 0$) agents think less about the optimal action in unusual times than about the optimal action in normal times (i.e., $\kappa_1 < \frac{1}{2}\kappa$). Raising the degree of strategic complementarity in actions (from $\gamma = 0$ to $\gamma > 0$ or $\gamma >> 0$) makes agents think even less about the optimal action in unusual times (i.e., κ_1 falls).

The reason is the following. When actions are strategic complements, the fact that other agents are not thinking carefully about the optimal action in a regime reduces the incentive for an individual agent to think about the optimal action in that regime. This effect is stronger for the regime that agents are thinking less about. Therefore, raising the degree of strategic complementarity in actions in both regimes makes the allocation of attention more extreme.

Note also that as the degree of strategic complementarity in actions increases, corner solutions occur more easily. See Figure 3. Hence, for a high degree of strategic complementarity in actions, a small change in parameters (e.g., a small change in the probability of unusual times) can have a large effect on the equilibrium allocation of attention. In fact, as γ approaches the value at which $2^{\kappa} = \gamma/(1-\gamma)$, the parameter region in which the equilibrium allocation of attention is an interior solution collapses to a single point. Moreover, for a sufficiently high degree of strategic complementarity in actions, there exist multiple equilibria. Specifically, whenever $2^{\kappa} \leq \gamma/(1-\gamma)$, there exists more than one equilibrium allocation of attention. See Proposition 1 and note that the condition $2^{\kappa} \leq \gamma/(1-\gamma)$ is equivalent to the condition $(1-\gamma)\,2^{\kappa}+\gamma 2^{-\kappa} \leq 1$.

Finally, strategic substitutability in actions has the opposite effect. As one can see from equations (16)-(17), strategic substitutability in actions (i.e., $\gamma < 0$) makes the equilibrium allocation of attention less extreme.

In Proposition 1 it is assumed that the degree of strategic complementarity in actions is the same across regimes. In Appendix A we also characterize in closed form the set of equilibria when the degree of strategic complementarity differs across regimes. Suppose that actions are strategic complements in both regimes but the degree of strategic complementarity may differ across regimes. Then, there exists a unique equilibrium for all $\sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}} \in \mathbb{R}_+$ if and only if

$$\frac{1}{(1-\gamma_1)\left[2^{\kappa}+\frac{\gamma_2}{1-\gamma_2}2^{-\kappa}\right]} < (1-\gamma_2)\left[2^{\kappa}+\frac{\gamma_1}{1-\gamma_1}2^{-\kappa}\right] < \frac{1}{\frac{\gamma_2}{1-\gamma_2}2^{-\kappa}}. \tag{18}$$

If this condition for a unique equilibrium is satisfied, the attention allocated to regime one equals

$$\kappa_1 = \begin{cases} \kappa & \text{if } \sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}} \geq (1-\gamma_2)\left[2^{\kappa}+\frac{\gamma_1}{1-\gamma_1}2^{-\kappa}\right] \\ \frac{1}{2}\kappa + \frac{1}{2}\log_2(x) & \text{otherwise} \\ 0 & \text{if } \sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}} \leq \frac{1}{(1-\gamma_1)\left[2^{\kappa}+\frac{\gamma_2}{1-\gamma_2}2^{-\kappa}\right]} \end{cases},$$

where

$$x \equiv \frac{\sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}} - \frac{\gamma_1}{1-\gamma_1}2^{-\kappa}}{1-\sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}}\frac{\gamma_2}{1-\gamma_2}2^{-\kappa}}.$$

These equations and condition (18) imply that, if the equilibrium allocation of attention is an interior solution, increasing the probability of a regime raises the attention allocated to that regime. Furthermore, increasing the degree of strategic complementarity in a single regime (an experiment we could not do before because we assumed that the degree of strategic complementarity was the same across regimes) makes agents allocate less attention to that regime.[17]

The model predicts that low probability of a regime and strategic complementarity in actions in that regime together can have dramatic effects on the amount of attention allocated to that regime. This prediction is important because in the real world actions often are strategic complements. Most policy decisions are made by a committee, and therefore policy-makers have to push for a

[17] The final result is similar to the results in Maćkowiak and Wiederholt (2009) and Hellwig and Veldkamp (2009) on how strategic complementarity in actions affects equilibrium attention allocation and equilibrium information acquisition.

common agenda in order to get any solution adopted. In this respect, policy-making in the euro area is extreme because most important policy decisions there must be unanimous. In the United States and throughout the global economy, many individuals and several institutions had to coordinate the policy reaction to the collapse of Lehman. In Japan, the response to the earthquake and tsunami of March 11, 2011 had to be coordinated between staff in private companies, nuclear safety regulators, and government officials.[18] In the private sector, relative performance evaluation is common and therefore a manager is punished less if he or she fails in times when other managers are failing too.[19] Researchers like to work on topics that other researchers are working on, because then those other researchers are more likely to be interested in the work. In many real-world situations, humans tend to compare themselves to others. The model predicts that those kinds of strategic complementarity in actions make agents even more inattentive to rare events.

3.3 Correlated optimal actions

This subsection relaxes the assumption that optimal actions are independent across regimes. We study how the correlation of optimal actions across regimes affects the quality of actions taken in the different regimes. Consider the special case of no strategic complementarity in actions, that is, $\gamma_1 = \gamma_2 = 0$. In this case the solution to the decision problem of a single agent is also the solution of the model, because there is no interaction between agents. The decision problem of a single agent is given by expressions (4)-(7), where Σ is non-diagonal. We solve the problem (4)-(7) numerically for different values of the covariance between the optimal actions in the two regimes, Σ_{12}.

The solution has the following feature: The larger in absolute value the prior correlation of the optimal actions across the two regimes, the smaller the expected loss in unusual times. Thinking about the best action in normal times now gives some information about the best action in unusual

[18] "Everyone was trying to avoid responsibility", said a person close to the response effort quoted by *Financial Times* in its May 19, 2011 issue.

[19] In 14 out of 15 leading U.S. and European banks, the chief executive officer in 2010 either was already the CEO before September 2008 (12 out of 15) or was a high ranking insider before September 2008 (2 out of 15). The only financial institution with a CEO in 2010 who had been an outsider before September 2008 was Royal Bank of Scotland, effectively nationalized after September 2008. See the June 15, 2011, issue of *Financial Times*. This fact supports the idea that the management of a corporation is punished little when it does poorly at a time when other corporations do poorly too.

times. Consequently, the expected loss in unusual times falls compared to the case when the optimal actions are independent across the two regimes. The stronger the prior correlation, the stronger this effect and thus the smaller the mean squared difference between the actual action and the optimal action in unusual times.

Another feature of the solution is that the posterior correlation of the optimal actions across regimes is a convex function of the prior correlation of the optimal actions across regimes. When the prior correlation rises in absolute value, the posterior correlation stays close to zero at first and then swiftly moves to one in absolute value. So long as the prior correlation is not too far from zero, it is optimal to behave approximately as if the prior correlation were zero.

Figure 4 illustrates these features of the solution through a numerical example. We fix $\Sigma_{11} = \Sigma_{22} = 1$, $p_1 = 0.01$, i.e. the probability of unusual times is 0.01, and we vary Σ_{12}. To begin with, suppose that $\Sigma_{12} = 0$, that is the optimal actions are independent across the two regimes. In this case, Λ and Ω are diagonal. We choose a value of κ such that the posterior variance of the optimal action in normal times, Ω_{22}, equals 0.01.[20] Then it turns out that the posterior variance of the optimal action in unusual times, Ω_{11}, equals 0.99. Now, suppose that Σ_{12} rises, that is the optimal actions become more and more positively correlated a priori, and all other parameters remain unchanged.[21] Figure 4 displays Ω_{11}, Ω_{22}, and the posterior correlation of the optimal actions, $\Omega_{12}/\sqrt{\Omega_{11}\Omega_{22}}$, as functions of Σ_{12}. Ω_{11} falls and is concave. More prior correlation in the optimal actions implies that agents do better on average in unusual times, but concavity means that this effect sets in slowly.[22] Furthermore, $\Omega_{12}/\sqrt{\Omega_{11}\Omega_{22}}$ rises and is convex. For values of Σ_{12} as large as 0.8, Ω_{12} is as small as 0.1 meaning that it is optimal to behave approximately as if Σ_{12} were zero.

The model predicts that people will be least prepared for unusual times when the optimal action in unusual times is independent of the optimal action in normal times. Furthermore, so long as the optimal action in unusual times and the optimal action in normal times are not strongly correlated, people will be much less prepared for unusual times than for normal times. These predictions are important, because in the real world optimal actions in unusual times and normal times are typically independent or weakly related. Thinking carefully about how to regulate financial institutions in

[20] This value of Ω_{22} means that thinking about the optimal action in normal times reduces the variance of that action by a factor of 100.

[21] Since $\Sigma_{11} = \Sigma_{22} = 1$, Σ_{12} is both the prior covariance of the optimal actions and their prior correlation.

[22] Ω_{22} also falls.

normal times or how to fine-tune open market operations to achieve a desired level of the federal funds rate helps little when confronted with the imminent collapse of Lehman Brothers. Thinking carefully about how to run a nuclear power plant efficiently in normal times fails to improve actions in times when an earthquake and tsunami disable the plant's cooling system.[23]

3.4 Extension: Learning the probability of rare events

We have assumed that the probability of the economy being in any given regime at any point in time is known. This subsection studies a version of the model in which the probability of the economy being in any given regime is a random variable.

Consider a random variable X that has a Bernoulli distribution with an unknown parameter p, i.e. X can take only the values 0 and 1, the probabilities are

$$\Pr(X = 1) = p \qquad \text{and} \qquad \Pr(X = 0) = 1 - p,$$

and p itself is a random variable. We think of $X = 1$ as unusual times and we think of $X = 0$ as normal times. Suppose that: (i) agents observe sequentially random variables $X_1, ..., X_s, ...$ that are i.i.d. over time and each has this Bernoulli distribution; (ii) in period 0, the agents' prior distribution of p is a beta distribution with parameters $\alpha > 0$ and $\beta > 0$; and (iii) in every period $t = 1, 2, ...$, agents observe whether $X = 1$ or $X = 0$ and agents update their prior distribution of p. Then the agents' posterior distribution of p given that $X_t = x_t$, $t = 1, ..., s$, is a beta distribution with parameters $\alpha + y$ and $\beta + s - y$, where $y = \sum_{t=1}^{s} x_t$. Furthermore, agents still solve the problem (4)-(7) where the probability of the economy being in any given regime has been replaced by the agents' posterior expectation of that probability.[24]

This version of the model matches what we believe are the following features of reality. When unusual times fail to occur for some time, agents tend to underestimate the probability of unusual times. Consequently, agents think even less about the optimal action in unusual times. Furthermore, when unusual times do occur agents tend to increase significantly their estimate of the

[23] *Financial Times* in its May 7-8, 2011 issue quotes Goshi Hosono, a senior aide to Japan's prime minister, saying "Tepco's job is to deliver a constant supply of electricity – extremely routine work. It is a company for stable times."

[24] This statement is true because the agents' prior distribution of p and the stochastic process $\{X_t\}$ are independent of the stochastic process $\{z_t, \varepsilon_{i,t}\}$.

probability of unusual times occurring again. Consequently, an occurrence of unusual times causes a significant reallocation of attention toward thinking about what to do in unusual times.

Consider a numerical example. Suppose that the true value of p is 0.01. In period 0, the agents' prior distribution of p is a beta distribution with parameters $\alpha = 1$ and $\beta = 99$. Note that the agents' prior expectation of p equals the truth, because the prior expectation of p equals $\alpha/(\alpha+\beta) = 0.01$. Let $X_t = 0$ for $t = 1, ..., s-1$, $X_t = 1$ for $t = s$, and $s = 101$. In words, the regime turns out to be normal times one hundred periods in a row and in period 101 the regime turns out to be unusual times.[25] The agents' posterior expectation of p evolves over time as shown in Figure 5. Note that between period 1 and period 100, the agents' posterior expectation of p falls slowly. Just before unusual times occur, the agents' posterior expectation of p equals 0.005. Agents underestimate the probability of unusual times by fifty percent. Furthermore, note that just after unusual times the agents' posterior expectation of p changes by a large amount. The agents' posterior expectation of p doubles to 0.01.

4 The efficient allocation of attention

Would society be better off from an ex-ante perspective if agents allocated their attention differently than in equilibrium? To answer this question, we study the following planner problem. The planner can tell agents how to allocate their attention (i.e., how carefully to think about the optimal actions in the different regimes). The planner has to respect the agents' information-processing constraint (i.e., the planner has to respect that agents can process only a limited amount of information). Finally, the planner maximizes ex-ante utility of the agents. The propositions in this section characterize analytically the relationship between the equilibrium allocation of attention and the efficient allocation of attention (i.e., the solution to the planner problem). When the two coincide, ex-ante utility cannot be raised by creating incentives for agents to allocate their attention differently, for example, by passing a law that requires companies running nuclear power plants to have a precise plan for actions in the case of an earthquake or tsunami. When the two differ, ex-ante utility can be raised by changing the allocation of attention.

Before stating the planner problem, we derive a simple expression for expected utility in regime

[25] The probability that unusual times fail to occur in one hundred Bernoulli trials with $p = 0.01$ equals about 0.36.

n, that is, $E\left[U^n\left(a_{i,t,n}, a_{t,n}, z_{t,n}\right)\right]$. The derivation follows closely the derivation of a similar expression in Angeletos and Pavan (2007). Let $\tilde{U}^n\left(a_{t,n}, z_{t,n}\right) \equiv U^n\left(a_{t,n}, a_{t,n}, z_{t,n}\right)$ denote the payoff in regime n when all agents take the same action $a_{i,t,n} = a_{t,n}$. It follows from equation (1) that

$$\begin{aligned}\tilde{U}^n\left(a_{t,n}, z_{t,n}\right) &= U^n(0,0,0) + \left(U_{a_i}^n + U_a^n\right) a_{t,n} + U_z^n z_{t,n} \\ &\quad + \frac{U_{a_i a_i}^n + 2U_{a_i a}^n + U_{aa}^n}{2} a_{t,n}^2 + \frac{U_{zz}^n}{2} z_{t,n}^2 + \left(U_{a_i z}^n + U_{az}^n\right) a_{t,n} z_{t,n}. \end{aligned} \tag{19}$$

In the following, we assume that $\tilde{U}^n\left(a_{t,n}, z_{t,n}\right)$ is concave in its first argument, that is,

$$U_{a_i a_i}^n + 2U_{a_i a}^n + U_{aa}^n < 0. \tag{20}$$

Let $a_{t,n}^*$ denote the common action $a_{t,n} \in \mathbb{R}$ that maximizes $\tilde{U}^n\left(a_{t,n}, z_{t,n}\right)$. It follows from equations (19) and (20) that

$$a_{t,n}^* = -\frac{U_{a_i}^n + U_a^n}{U_{a_i a_i}^n + 2U_{a_i a}^n + U_{aa}^n} - \frac{U_{a_i z}^n + U_{az}^n}{U_{a_i a_i}^n + 2U_{a_i a}^n + U_{aa}^n} z_{t,n}. \tag{21}$$

One can show that expected utility in regime n equals

$$\begin{aligned} E\left[U^n\left(a_{i,t,n}, a_{t,n}, z_{t,n}\right)\right] &= E\left[\tilde{U}^n\left(a_{t,n}^*, z_{t,n}\right)\right] - \frac{\left|U_{a_i a_i}^n + 2U_{a_i a}^n + U_{aa}^n\right|}{2} E\left[\left(a_{t,n} - a_{t,n}^*\right)^2\right] \\ &\quad - \frac{\left|U_{a_i a_i}^n\right|}{2} E\left[\left(a_{i,t,n} - a_{t,n}\right)^2\right]. \end{aligned} \tag{22}$$

The proof is in Appendix B. The last equation implies that expected utility is maximized when all agents take the action $a_{t,n}^*$ for all $z_{t,n}$, that is, $a_{i,t,n} = a_{t,n}^*$ for all $z_{t,n}$. There is a loss in expected utility when the mean action in the population does not move one for one with $a_{t,n}^*$ (the second term on the right-hand side of the last equation) and when there is dispersion in actions (the third term on the right-hand side of the last equation).

When Σ is diagonal and the planner considers equilibria of the form $a_{t,n} = \psi_n + \phi_n z_{t,n}$, the problem of the planner who chooses the allocation of attention of the agents so as to maximize expected utility of the agents reads:

$$\max_{(\kappa_1,\kappa_2)\in\mathbb{R}_+^2} \sum_{n=1}^{2} p_n \left\{ \frac{U_{a_i a_i}^n + 2U_{a_i a}^n + U_{aa}^n}{2} E\left[\left(a_{t,n} - a_{t,n}^*\right)^2\right] + \frac{U_{a_i a_i}^n}{2} E\left[\left(a_{i,t,n} - a_{t,n}\right)^2\right] \right\}, \tag{23}$$

subject to equation (21),

$$a_{t,n} = \psi_n + \phi_n z_{t,n}, \tag{24}$$

$$a_{i,t,n} = (\varphi_n + \gamma_n \psi_n) + (\gamma_n \phi_n + 1 - \gamma_n) \frac{\frac{\Sigma_{nn}}{\Lambda_{nn}}}{\frac{\Sigma_{nn}}{\Lambda_{nn}} + 1} (z_{t,n} + \varepsilon_{i,t-1,n}), \tag{25}$$

$$\psi_n = \frac{\varphi_n}{1 - \gamma_n}, \phi_n = \frac{(1-\gamma_n) \frac{\frac{\Sigma_{nn}}{\Lambda_{nn}}}{\frac{\Sigma_{nn}}{\Lambda_{nn}}+1}}{1 - \gamma_n \frac{\frac{\Sigma_{nn}}{\Lambda_{nn}}}{\frac{\Sigma_{nn}}{\Lambda_{nn}}+1}}, \tag{26}$$

$$\frac{\Sigma_{nn}}{\Lambda_{nn}} = 2^{2\kappa_n} - 1, \tag{27}$$

and

$$\kappa_1 + \kappa_2 \leq \kappa. \tag{28}$$

Objective (23) is expected utility of the agents minus $\sum_{n=1}^{2} p_n E\left[\tilde{U}^n\left(a^*_{t,n}, z_{t,n}\right)\right]$, which is a term that the planner cannot affect. Equation (25) follows from equations (5)-(6) and equation (24). Equation (26) follows from equations (24)-(25), the definition of $a_{t,n}$, and the assumption that noise washes out in the aggregate. Equation (27) follows from the definition $\kappa_n \equiv \frac{1}{2}\log_2\left(\frac{\Sigma_{nn}}{\Omega_{nn}}\right)$ and $\Omega_{nn} = \Sigma_{nn} - \Sigma_{nn}(\Sigma_{nn} + \Lambda_{nn})^{-1}\Sigma_{nn}$. Finally, constraint (28) is the information-processing constraint of the agents in the case of diagonal Σ and Λ.

In the following, we focus on the case that the economy is efficient under perfect information, that is, the equilibrium actions under perfect information equal the welfare-maximizing actions. It follows from equation (5) and the definition of $a_{t,n}$ that the equilibrium actions under perfect information are given by

$$a_{i,t,n} = \frac{\varphi_n}{1-\gamma_n} + z_{t,n}.$$

The welfare-maximizing actions are given by $a_{i,t,n} = a^*_{t,n}$ where $a^*_{t,n}$ is given by equation (21). The condition that the equilibrium actions under perfect information equal the welfare-maximizing actions thus reads

$$-\frac{U^n_{a_i} + U^n_a}{U^n_{a_i a_i} + 2U^n_{a_i a} + U^n_{aa}} = \frac{\varphi_n}{1-\gamma_n}, \tag{29}$$

and

$$-\frac{U^n_{a_i z} + U^n_{az}}{U^n_{a_i a_i} + 2U^n_{a_i a} + U^n_{aa}} = 1. \tag{30}$$

Substituting equation (21), equations (24)-(27), and equations (29)-(30) into the planner's objective (23) gives

$$\max_{(\kappa_1,\kappa_2)\in\mathbb{R}^2_+} -\sum_{n=1}^{2} p_n \delta_n \Sigma_{nn} \left[\left(1 - 2\gamma_n + \frac{U^n_{aa}}{U^n_{a_i a_i}}\right)\frac{1}{\left(\gamma_n + (1-\gamma_n)2^{2\kappa_n}\right)^2} + \frac{(1-\gamma_n)^2\left(2^{2\kappa_n}-1\right)}{\left(\gamma_n + (1-\gamma_n)2^{2\kappa_n}\right)^2}\right], \tag{31}$$

subject to

$$\kappa_1 + \kappa_2 \leq \kappa. \tag{32}$$

Increasing the attention allocated to regime n reduces the mean squared difference between the mean action $a_{t,n}$ and the welfare-maximizing action $a^*_{t,n}$ (see the first term in square brackets in the objective), but may increase or decrease the dispersion in actions in regime n (see the second term in square brackets in the objective). The reason for the second effect is that at $\kappa_n = 0$ dispersion in actions in regime n equals zero and as $\kappa_n \to \infty$ dispersion in actions in regime n goes to zero, while for intermediate values of κ_n dispersion in actions is positive.

Finally, in the following, we focus on the case where the degree of strategic complementarity is the same across regimes and the ratio $U^n_{aa}/U^n_{a_i a_i}$ is the same across regimes. The planner problem then reduces to:

$$\max_{(\kappa_1,\kappa_2)\in\mathbb{R}^2_+} -\sum_{n=1}^{2} p_n \delta_n \Sigma_{nn} \left[\left(1 - 2\gamma + \frac{U_{aa}}{U_{a_i a_i}}\right) \frac{1}{\left(\gamma + (1-\gamma)\, 2^{2\kappa_n}\right)^2} + \frac{(1-\gamma)^2 \left(2^{2\kappa_n} - 1\right)}{\left(\gamma + (1-\gamma)\, 2^{2\kappa_n}\right)^2} \right], \tag{33}$$

subject to

$$\kappa_1 + \kappa_2 \leq \kappa. \tag{34}$$

The next two propositions state results concerning the relationship between the equilibrium allocation of attention and the efficient allocation of attention.

Proposition 2 *Assume that Σ is diagonal, $\gamma_1 = \gamma_2 \equiv \gamma$, and $2^\kappa > \frac{\gamma}{1-\gamma}$. The equilibrium allocation of attention, denoted κ_1^{equ}, is then given by equation (16). Furthermore, assume that condition (20), conditions (29)-(30) and $\left(U^1_{aa}/U^1_{a_i a_i}\right) = \left(U^2_{aa}/U^2_{a_i a_i}\right) \equiv \left(U_{aa}/U_{a_i a_i}\right)$ hold. The efficient allocation of attention, denoted κ_1^{eff}, is then given by the solution to problem (33)-(34). Finally, suppose that the constraint (34) is binding and the problem (33)-(34) is convex. Then the following result holds. If $\gamma = \left(U_{aa}/U_{a_i a_i}\right)$ or $\kappa_1^{equ} = \frac{1}{2}\kappa$, the equilibrium allocation of attention equals the efficient allocation of attention: $\kappa_1^{equ} = \kappa_1^{eff}$.*

Proof. See Appendix C. ■

Proposition 2 can be interpreted as a welfare theorem for the allocation of attention. The proposition states conditions under which the equilibrium allocation of attention equals the efficient allocation of attention. The setup is the following: The conditions of Proposition 1 hold; agents take the welfare-maximizing actions under perfect information; there is a certain degree of symmetry across regimes; and the planner problem is convex. In this case, the equilibrium allocation of attention equals the efficient allocation of attention if either the payoff function has the property that the ratio $-\left(U_{a_i a}/U_{a_i a_i}\right)$ equals the ratio $\left(U_{aa}/U_{a_i a_i}\right)$, or in equilibrium agents allocate their attention equally across regimes (i.e., $\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}=1$), or both.

A few comments on the setup are in order. The conditions of Proposition 1 imply that there exists a unique equilibrium and a closed form solution for the equilibrium allocation of attention. This simplifies the proof of Proposition 2. The condition that the economy is efficient under perfect information is a natural benchmark. It means that inefficiencies, if any, arise due to limited attention by agents. The requirement that there is a certain degree of symmetry across regimes will be relaxed later.

The following proposition characterizes the direction of the inefficiency when the payoff function does not have the property $-\left(U_{a_i a}/U_{a_i a_i}\right)=\left(U_{aa}/U_{a_i a_i}\right)$ and in equilibrium agents do no allocate their attention equally across regimes.

Proposition 3 *Assume that the conditions of Proposition 2 are satisfied. If* $\gamma \neq \left(U_{aa}/U_{a_i a_i}\right)$, $\kappa_1^{equ} \neq \frac{1}{2}\kappa$, *and* $\kappa_1^{equ} \in (0,\kappa)$, *the equilibrium allocation of attention differs from the efficient allocation of attention. More precisely, when* $\gamma < \left(U_{aa}/U_{a_i a_i}\right)$ *the planner would prefer agents to pay more attention to the regime that they are allocating less attention to (i.e., when* $\gamma < \left(U_{aa}/U_{a_i a_i}\right)$ *then* $0<\kappa_n^{equ}<\frac{1}{2}\kappa$ *implies* $\kappa_n^{eff}>\kappa_n^{equ}$*). In contrast, when* $\gamma > \left(U_{aa}/U_{a_i a_i}\right)$ *the planner would prefer agents to pay even less attention to the regime that they are allocating less attention to (i.e., when* $\gamma > \left(U_{aa}/U_{a_i a_i}\right)$ *then* $0<\kappa_n^{equ}<\frac{1}{2}\kappa$ *implies* $\kappa_n^{eff}<\kappa_n^{equ}$*).*

Proof. See Appendix D. ■

When agents allocate their attention to some extent to both regimes, agents do not allocate their attention equally across regimes, and the payoff function does not have the property $-\left(U_{a_i a}/U_{a_i a_i}\right)=\left(U_{aa}/U_{a_i a_i}\right)$, the equilibrium allocation of attention differs from the efficient allocation of attention. In addition, the direction of the inefficiency can be seen directly from the payoff function. If $-\left(U_{a_i a}/U_{a_i a_i}\right)<\left(U_{aa}/U_{a_i a_i}\right)$ the planner would prefer agents to pay *more* attention to the regime that they are devoting less attention to. If $-\left(U_{a_i a}/U_{a_i a_i}\right)>\left(U_{aa}/U_{a_i a_i}\right)$ the planner would prefer agents to pay *even less* attention to the regime that they are devoting less attention to.

For example, suppose that $\frac{1}{(1-\gamma)2^{\kappa}+\gamma 2^{-\kappa}} < \sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}} < 1$, implying that in equilibrium agents think to some extent about the optimal action in unusual times, but less than about the optimal action in normal times. Then, if $-\left(U_{a_i a}/U_{a_i a_i}\right) < \left(U_{aa}/U_{a_i a_i}\right)$ the planner would prefer agents to think more about the optimal action in unusual times and less about the optimal action in normal times than is the case in equilibrium.

Proposition 2 gives two conditions under which the equilibrium allocation of attention equals the efficient allocation of attention. One of these conditions reads

$$\frac{U_{a_i a}}{|U_{a_i a_i}|} + \frac{U_{aa}}{|U_{a_i a_i}|} = 0. \tag{35}$$

Moreover, Proposition 3 states that if the left-hand side of equation (35) is strictly *negative*, agents in equilibrium allocate *too little* attention to unusual times from an ex-ante welfare perspective. By contrast, if the left-hand side of equation (35) is strictly *positive*, agents in equilibrium allocate *too much* attention to unusual times from an ex-ante welfare perspective. To understand these results, note that there are two externalities in the model – a positive and a negative externality. When agents think more carefully about the optimal action in a regime, the mean action in the regime moves more with the fundamental in that regime, which directly increases ex-ante utility. This positive externality is present for both regimes and is stronger for the regime that agents are paying less attention to. Hence, if this positive externality were the only externality, the planner would want agents to think *more* about unusual times. On the other hand, when agents think more carefully about the optimal action in a regime and therefore the mean action in the regime moves more with the fundamental in that regime, the problem of other agents becomes more complicated. This negative externality is present for both regimes and is stronger for the regime that agents are paying less attention to. Hence, if this negative externality were the only externality, the planner would want agents to think *less* about unusual times. When condition (35) holds, the positive externality and the negative externality exactly cancel and the equilibrium allocation of attention equals the efficient allocation of attention. By contrast, when the left-hand side of equation (35) is strictly negative, the positive externality dominates, and when the left-hand side of equation (35) is strictly positive, the negative externality dominates.

Condition (35) is equivalent to a condition that has already appeared in the literature in a different context. More precisely, condition (35) is equivalent to the following condition which appears in Angeletos and Pavan (2007):

$$-\frac{U_{a_i a}}{U_{a_i a_i}} = 1 - \left(\frac{U_{a_i a_i}}{U_{a_i a_i}} + 2\frac{U_{a_i a}}{U_{a_i a_i}} + \frac{U_{aa}}{U_{a_i a_i}} \right). \tag{36}$$

Angeletos and Pavan (2007) study an economy with a continuum of agents in which each agent observes a noisy private and public signal. The precision of the two signals is exogenous. Due to the quadratic Gaussian structure of the economy, actions are a linear function of the two signals and Angeletos and Pavan (2007) refer to the coefficients on the two signals as the "use of information." They then compare the equilibrium use of information to the efficient use of information, where the latter is defined as the one that maximizes ex-ante utility. For economies that are efficient under perfect information, it turns out that the equilibrium use of information equals the efficient use of information if and only if condition (36) is satisfied. We thus arrive at the following conclusion. The same condition that governs the relationship between the equilibrium use of information and the efficient use of information in the model in Angeletos and Pavan (2007) also governs the relationship between the equilibrium allocation of attention and the efficient allocation of attention in our model with an endogenous signal precision. Our intuition for this finding is the following. If the use of information is efficient, then the acquisition of information is also efficient, so long as there is no direct externality in the acquisition of information (which is the case here).

Proposition 2 assumes that there is a certain degree of symmetry across regimes. The degree of strategic complementarity $\gamma_n \equiv -\left(U_{a_i a}^n / U_{a_i a_i}^n\right)$ is assumed to be the same across regimes and the ratio $\left(U_{aa}^n / U_{a_i a_i}^n\right)$ is assumed to be the same across regimes. When this symmetry requirement is not satisfied, a sufficient condition for the equilibrium allocation of attention to equal the efficient allocation of attention is that condition (35) holds for each regime, that is, $-\left(U_{a_i a}^n / U_{a_i a_i}^n\right) = \left(U_{aa}^n / U_{a_i a_i}^n\right)$ for $n = 1, 2$. The proof is the same as before. The agents' first-order condition then equals the planners' first-order condition, and the conditions for corner solutions are the same for the agents and the planner.

Finally, Proposition 3, which characterizes the direction of the inefficiency when $\gamma \neq (U_{aa}/U_{a_i a_i})$, $\kappa_1^{equ} \neq \frac{1}{2}\kappa$, and $\kappa_1^{equ} \in (0, \kappa)$, does not cover the case of corner solutions. We now cover this case. When $\gamma > (U_{aa}/U_{a_i a_i})$ and $\kappa_1^{equ} = 0$ or $\kappa_1^{equ} = \kappa$, the equilibrium allocation of attention equals the efficient allocation of attention. The planner would prefer agents to pay even less attention to the regime that they are devoting less attention to. However, this is impossible because the equilibrium allocation of attention is already a corner solution. Hence, the equilibrium allocation of attention equals the efficient allocation of attention. They are both corner solutions. When $\gamma < (U_{aa}/U_{a_i a_i})$ and $\kappa_1^{equ} = 0$ or $\kappa_1^{equ} = \kappa$, the equilibrium allocation of attention may equal or differ from the efficient allocation of attention. If the efficient allocation of attention is a corner solution, the two coincide. If the efficient allocation of attention is not a corner solution, the two differ.

5 Conclusions and future research

This paper proposes an explanation for why people were unprepared for the global financial crisis, the European debt crisis, and the Fukushima nuclear accident. The explanation has four features: (1) Humans have a limited ability to process information and therefore cannot prepare well for every contingency. (2) These events seemed unlikely a priori. (3) Thinking carefully about the optimal action in normal times does not improve much actions in unusual times. (4) Actions are strategic complements. The model identifies the circumstances under which people will be unprepared for contingent events also in the future.

We study a rational inattention model in which agents decide how carefully to think about optimal actions in different contingencies, subject to an information-processing constraint. We find that agents are unprepared in a state when the state has a low probability, the optimal action in that state is uncorrelated with the optimal action in normal times, and actions are strategic complements. We then use the model to ask the following question: Would society be better off if agents allocated their attention differently than in equilibrium? To answer this question, we compare the equilibrium allocation of attention to the efficient allocation of attention. We find that the same condition that governs the relationship between the equilibrium use of information and the efficient use of information in Angeletos and Pavan (2007) governs the relationship between the equilibrium allocation of attention and the efficient allocation of attention in our model with an endogenous information structure.

In the real world, there exists regulation that affects the allocation of attention. For example, aviation regulations force passengers on every flight to think about the optimal action in the rare event of a landing on water. Does this increase ex-ante utility? At the same time, there does not seem to be regulation in Japan that requires companies running nuclear power plants to have a precise plan of what to do when an earthquake and tsunami disable a plant's cooling system.

Should this be changed? The efficiency results in this paper help understand when regulation that affects the allocation of attention can improve welfare and when it cannot improve welfare.

The efficiency question asked in this paper – whether the equilibrium allocation of attention equals the efficient allocation of attention – is new to the best of our knowledge; has a clear answer; and could be asked in a wide range of other contexts. For example, one could ask whether the extent to which investors think about payoffs of their assets in different states of the world is efficient.

The model is simple in some dimensions. For instance, in future research one could relax the assumption that the probability of unusual times is independent of actions taken by agents in the model. We think of this assumption as a reasonable approximation because, no matter what humans do, failure of a systematically important financial institution, severe fiscal stress, and a nuclear emergency will probably remain low-but-non-zero probability events.

A Proof of Proposition 1

Step 1: We consider equilibria where the average action in a regime is an affine function of the fundamental in that regime. Formally, for $n = 1$ and $n = 2$,

$$a_{t,n} = \psi_n + \phi_n z_{t,n}, \tag{37}$$

where ψ_1, ϕ_1, ψ_2, and ϕ_2 are undetermined coefficients that we need to solve for.

Step 2: The information choice problem (4)-(7) can now be stated as follows. Substituting equations (2), (3) and (5) into objective (4), deducting a constant that the agent cannot affect from the objective, and using equation (37) to substitute for $a_{t,n}$ in the objective yields

$$\max_{\Lambda} \left\{ -\sum_{n=1}^{2} p_n \delta_n \left(\gamma_n \phi_n + 1 - \gamma_n\right)^2 \Omega_{nn} \right\}, \tag{38}$$

subject to

$$\Omega = \Sigma - \Sigma \left(\Sigma + \Lambda\right)^{-1} \Sigma, \tag{39}$$

$$\frac{1}{2} \log_2 \left(\frac{|\Sigma|}{|\Omega|} \right) \leq \kappa, \tag{40}$$

and the restriction that Λ is a positive semidefinite matrix. Here Ω_{nn} denotes the posterior variance of the fundamental in regime n. Furthermore, using the formula for the determinant of a two-by-two matrix, the information flow constraint (40) can be expressed as

$$\frac{1}{2} \log_2 \left(\frac{\Sigma_{11}\Sigma_{22} - \Sigma_{12}^2}{\Omega_{11}\Omega_{22} - \Omega_{12}^2} \right) \leq \kappa, \tag{41}$$

where Ω_{12} denotes the posterior covariance of the fundamental in the two regimes.

Step 3: When the optimal action in regime one and the optimal action in regime two are independent (i.e., $\Sigma_{12} = 0$), it is optimal to receive independent signals concerning the optimal action in regime one and the optimal action in regime two (i.e., $\Lambda_{12} = 0$). The proof is as follows. First, the information flow constraint (41) is always binding. Second, increasing Ω_{12}^2 for a given Ω_{11} and Ω_{22} raises the information flow on the left-hand side of constraint (41) without improving objective (38). Third, when $\Sigma_{12} = 0$, then $\Omega_{12} = 0$ if and only if $\Lambda_{12} = 0$. Hence, when $\Sigma_{12} = 0$, the solution to the information choice problem (38)-(40) has the property $\Lambda_{12} = 0$. Next, using $\Sigma_{12} = \Lambda_{12} = \Omega_{12} = 0$ the information choice problem (38)-(40) simplifies to

$$\max_{\left(\Lambda_{11}^{-1}, \Lambda_{22}^{-1}\right) \in \mathbb{R}_+^2} \left\{ -\sum_{n=1}^{2} p_n \delta_n \left(\gamma_n \phi_n + 1 - \gamma_n\right)^2 \Omega_{nn} \right\}, \tag{42}$$

subject to

$$\Omega_{nn} = \frac{1}{\frac{\Sigma_{nn}}{\Lambda_{nn}} + 1} \Sigma_{nn}, \tag{43}$$

and

$$\frac{1}{2} \log_2 \left(\frac{\Sigma_{11}}{\Omega_{11}} \right) + \frac{1}{2} \log_2 \left(\frac{\Sigma_{22}}{\Omega_{22}} \right) \leq \kappa. \tag{44}$$

Let $\kappa_n \equiv \frac{1}{2} \log_2 \left(\frac{\Sigma_{nn}}{\Omega_{nn}} \right)$ denote the uncertainty reduction about the fundamental in regime n. The information choice problem (42)-(44) can be written as

$$\max_{(\kappa_1, \kappa_2) \in \mathbb{R}_+^2} \left\{ - \sum_{n=1}^{2} p_n \delta_n \left(\gamma_n \phi_n + 1 - \gamma_n \right)^2 \Omega_{nn} \right\}, \tag{45}$$

subject to

$$\Omega_{nn} = 2^{-2\kappa_n} \Sigma_{nn}, \tag{46}$$

and

$$\kappa_1 + \kappa_2 \leq \kappa. \tag{47}$$

The unique solution to this problem is given by

$$\kappa_1 = \begin{cases} \kappa & \text{if } x \geq 2^{\kappa} \\ \frac{1}{2}\kappa + \frac{1}{2} \log_2 (x) & \text{if } x \in [2^{-\kappa}, 2^{\kappa}] \\ 0 & \text{if } x \leq 2^{-\kappa} \end{cases}, \tag{48}$$

where

$$x \equiv \sqrt{\frac{p_1 \delta_1 \left(\gamma_1 \phi_1 + 1 - \gamma_1 \right)^2 \Sigma_{11}}{p_2 \delta_2 \left(\gamma_2 \phi_2 + 1 - \gamma_2 \right)^2 \Sigma_{22}}}; \tag{49}$$

and

$$\kappa_2 = \kappa - \kappa_1. \tag{50}$$

The optimal uncertainty reduction about the fundamental in regime one is an increasing function of κ and x. Finally, it follows from equation (43) and $\kappa_n \equiv \frac{1}{2} \log_2 \left(\frac{\Sigma_{nn}}{\Omega_{nn}} \right)$ that the optimal signal precisions are then given by

$$\Lambda_{11}^{-1} = \frac{2^{2\kappa_1} - 1}{\Sigma_{11}}, \tag{51}$$

$$\Lambda_{22}^{-1} = \frac{2^{2\kappa_2} - 1}{\Sigma_{22}}. \tag{52}$$

Step 4: Equations (48)-(50) give the optimal allocation of attention as a function of the parameters of the model and the undetermined coefficients ϕ_1 and ϕ_2. The next step is to solve for the undetermined coefficients ϕ_1 and ϕ_2 as a function of the optimal allocation of attention. Combining results one then obtains the equilibrium of the model. The actions by agent i are given by equation (5). Substituting the guess (37) into equation (5) yields

$$a_{i,t,n} = (\varphi_n + \gamma_n \psi_n) + (\gamma_n \phi_n + 1 - \gamma_n) E[z_{t,n} | s_{i,t-1}].$$

Calculating the conditional expectation in the last equation using equation (6), $\Sigma_{12} = \Lambda_{12} = 0$, and equations (51)-(52) yields

$$a_{i,t,n} = (\varphi_n + \gamma_n \psi_n) + (\gamma_n \phi_n + 1 - \gamma_n) \left(1 - 2^{-2\kappa_n}\right) (z_{t,n} + \varepsilon_{i,t-1,n}).$$

Calculating the mean action in the population gives

$$a_{t,n} = (\varphi_n + \gamma_n \psi_n) + (\gamma_n \phi_n + 1 - \gamma_n) \left(1 - 2^{-2\kappa_n}\right) z_{t,n}.$$

It follows that, for a given allocation of attention (i.e., for a pair κ_1 and κ_2), the guess (37) is correct if and only if

$$\psi_n = \frac{\varphi_n}{1 - \gamma_n}, \tag{53}$$

$$\phi_n = \frac{(1 - \gamma_n)\left(1 - 2^{-2\kappa_n}\right)}{1 - \gamma_n \left(1 - 2^{-2\kappa_n}\right)}. \tag{54}$$

The last two equations give the undetermined coefficients ψ_1, ψ_2, ϕ_1, and ϕ_2 as a function of the allocation of attention κ_1 and κ_2 and the parameters φ_1, φ_2, γ_1, and γ_2.

Step 5: An equilibrium allocation of attention is a pair (κ_1, κ_2) satisfying equations (48)-(50), where ϕ_1 and ϕ_2 are given by equation (54). Using equation (54) to substitute for ϕ_1 and ϕ_2 in equation (49) yields

$$x = \sqrt{\frac{p_1 \delta_1 \Sigma_{11}}{p_2 \delta_2 \Sigma_{22}} \frac{\frac{1-\gamma_1}{1-\gamma_1\left(1-2^{-2\kappa_1}\right)}}{\frac{1-\gamma_2}{1-\gamma_2\left(1-2^{-2\kappa_2}\right)}}}. \tag{55}$$

Thus, an equilibrium allocation of attention is a pair (κ_1, κ_2) satisfying equations (48), (50) and (55). It is useful to distinguish three types of equilibria: (i) the equilibrium allocation of attention has the property $\kappa_1 = 0$, (ii) the equilibrium allocation of attention has the property $\kappa_1 = \kappa$, and (iii) the equilibrium allocation of attention has the property $\kappa_1 = \frac{1}{2}\kappa + \frac{1}{2}\log_2(x)$.

First, turn to an equilibrium with the property $\kappa_1 = 0$. Substituting $\kappa_1 = 0$ and $\kappa_2 = \kappa$ into equation (55) yields

$$x = \sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}} \frac{1-\gamma_1}{1-\gamma_2} \left[1 - \gamma_2 \left(1 - 2^{-2\kappa}\right)\right].$$

It follows from the last equation and equation (48) that $\kappa_1 = 0$ is an equilibrium if and only if

$$\sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}} \frac{1-\gamma_1}{1-\gamma_2} \left[1 - \gamma_2 \left(1 - 2^{-2\kappa}\right)\right] \leq 2^{-\kappa}.$$

This condition can be stated as

$$\sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}} \leq \frac{1}{(1-\gamma_1)\left[2^{\kappa} + \frac{\gamma_2}{1-\gamma_2} 2^{-\kappa}\right]}. \tag{56}$$

Second, consider an equilibrium with the property $\kappa_1 = \kappa$. Substituting $\kappa_1 = \kappa$ and $\kappa_2 = 0$ into equation (55) yields

$$x = \sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}} \frac{1-\gamma_1}{1-\gamma_2} \frac{1}{1-\gamma_1\left(1-2^{-2\kappa}\right)}.$$

It follows from the last equation and equation (48) that $\kappa_1 = \kappa$ is an equilibrium if and only if

$$\sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}} \frac{1-\gamma_1}{1-\gamma_2} \frac{1}{1-\gamma_1\left(1-2^{-2\kappa}\right)} \geq 2^{\kappa}.$$

This condition can be stated as

$$\sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}} \geq (1-\gamma_2)\left[2^{\kappa} + \frac{\gamma_1}{1-\gamma_1} 2^{-\kappa}\right]. \tag{57}$$

Third, turn to an equilibrium with the property $\kappa_1 = \frac{1}{2}\kappa + \frac{1}{2}\log_2(x)$. Substituting $\kappa_1 = \frac{1}{2}\kappa + \frac{1}{2}\log_2(x)$ and $\kappa_2 = \kappa - \kappa_1$ into equation (55) yields

$$x = \sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}} \frac{\frac{1-\gamma_1}{1-\gamma_1\left(1-2^{-\kappa}\frac{1}{x}\right)}}{\frac{1-\gamma_2}{1-\gamma_2\left(1-2^{-\kappa}x\right)}}.$$

Rearranging the last equation yields

$$\left[1 - \sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}} \frac{\gamma_2}{1-\gamma_2} 2^{-\kappa}\right] x = \sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}} - \frac{\gamma_1}{1-\gamma_1} 2^{-\kappa}. \tag{58}$$

If $\left[1 - \sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}} \frac{\gamma_2}{1-\gamma_2} 2^{-\kappa}\right] \neq 0$, the unique solution to the last equation is

$$x = \frac{\sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}} - \frac{\gamma_1}{1-\gamma_1} 2^{-\kappa}}{1 - \sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}} \frac{\gamma_2}{1-\gamma_2} 2^{-\kappa}}. \tag{59}$$

Thus, when $\left[1-\sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}}\frac{\gamma_2}{1-\gamma_2}2^{-\kappa}\right]\neq 0$, it follows from the last equation and equation (48) that $\kappa_1=\frac{1}{2}\kappa+\frac{1}{2}\log_2(x)$ is an equilibrium if and only if

$$\frac{\sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}}-\frac{\gamma_1}{1-\gamma_1}2^{-\kappa}}{1-\sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}}\frac{\gamma_2}{1-\gamma_2}2^{-\kappa}}\in\left[2^{-\kappa},2^{\kappa}\right]. \tag{60}$$

Furthermore, when

$$\left[1-\sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}}\frac{\gamma_2}{1-\gamma_2}2^{-\kappa}\right]>0, \tag{61}$$

condition (60) is equivalent to

$$\sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}}\in\left[\frac{\frac{1}{1-\gamma_1}}{2^{\kappa}+\frac{\gamma_2}{1-\gamma_2}2^{-\kappa}},\frac{2^{\kappa}+\frac{\gamma_1}{1-\gamma_1}2^{-\kappa}}{\frac{1}{1-\gamma_2}}\right]. \tag{62}$$

When

$$\left[1-\sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}}\frac{\gamma_2}{1-\gamma_2}2^{-\kappa}\right]<0, \tag{63}$$

condition (60) is equivalent to

$$\sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}}\in\left[\frac{2^{\kappa}+\frac{\gamma_1}{1-\gamma_1}2^{-\kappa}}{\frac{1}{1-\gamma_2}},\frac{\frac{1}{1-\gamma_1}}{2^{\kappa}+\frac{\gamma_2}{1-\gamma_2}2^{-\kappa}}\right]. \tag{64}$$

Finally, if

$$\left[1-\sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}}\frac{\gamma_2}{1-\gamma_2}2^{-\kappa}\right]=0, \tag{65}$$

equation (58) reduces to

$$\sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}}=\frac{\gamma_1}{1-\gamma_1}2^{-\kappa}. \tag{66}$$

In summary, if conditions (61)-(62) or conditions (63)-(64) hold, a unique equilibrium with the property $\kappa_1=\frac{1}{2}\kappa+\frac{1}{2}\log_2(x)$ exists and in this equilibrium x is given by equation (59). If conditions (65)-(66) hold, a continuum of equilibria with the property $\kappa_1=\frac{1}{2}\kappa+\frac{1}{2}\log_2(x)$ exist; namely any $\kappa_1\in[0,\kappa]$ is such an equilibrium.

This completes the characterization of equilibria of the form (37). If $\gamma_1=\gamma_2\equiv\gamma$, conditions (56), (57), (61)-(62), (63)-(64) and (65)-(66) and equation (59) reduce to the conditions and equation given in Proposition 1.

B Proof of Equation (22)

Step 1: A Taylor expansion of U^n around $a_{i,t,n} = a_{t,n}$ gives

$$U^n(a_{i,t,n}, a_{t,n}, z_{t,n}) = U^n(a_{t,n}, a_{t,n}, z_{t,n}) + \left[U^n_{a_i} + \left(U^n_{a_i a_i} + U^n_{a_i a}\right) a_{t,n} + U^n_{a_i z} z_{t,n}\right](a_{i,t,n} - a_{t,n}) + \frac{U^n_{a_i a_i}}{2}(a_{i,t,n} - a_{t,n})^2 . \quad (67)$$

Let $W^n\left(a_{t,n}, \sigma_{a_{i,t,n}}, z_{t,n}\right)$ denote welfare in state n under a utilitarian aggregator

$$W^n\left(a_{t,n}, \sigma_{a_{i,t,n}}, z_{t,n}\right) \equiv \int U^n(a_{i,t,n}, a_{t,n}, z_{t,n})\, d\Psi^{n,t}(a_{i,t,n}) . \quad (68)$$

Combining the last two equations gives

$$W^n\left(a_{t,n}, \sigma_{a_{i,t,n}}, z_{t,n}\right) = U^n(a_{t,n}, a_{t,n}, z_{t,n}) + \frac{U^n_{a_i a_i}}{2}\sigma^2_{a_{i,t,n}}, \quad (69)$$

where $\sigma^2_{a_{i,t,n}} \equiv \int (a_{i,t,n} - a_{t,n})^2 d\Psi^{n,t}(a_{i,t,n})$ denotes the dispersion of individual actions in the population. Next, a Taylor expansion of $W^n\left(a_{t,n}, \sigma_{a_{i,t,n}}, z_{t,n}\right)$ around $a_{t,n} = a^*_{t,n}$ and $\sigma_{a_{i,t,n}} = 0$, where $a^*_{t,n}$ is given by equation (21), yields

$$W^n\left(a_{t,n}, \sigma_{a_{i,t,n}}, z_{t,n}\right) = W^n\left(a^*_{t,n}, 0, z_{t,n}\right) + \frac{U^n_{a_i a_i} + 2U^n_{a_i a} + U^n_{aa}}{2}\left(a_{t,n} - a^*_{t,n}\right)^2 + \frac{U^n_{a_i a_i}}{2}\sigma^2_{a_{i,t,n}}. \quad (70)$$

Here we used equation (69) to compute the first and second derivatives of W^n and exploited the fact that the first derivative of W^n with respect to $a_{t,n}$ evaluated at $a^*_{t,n}$ equals zero.

Step 2: Given any strategy $a_{i,t,n} : \mathbb{R}^2 \to \mathbb{R}$, expected utility in state n is given by

$$E\left[U^n(a_{i,t,n}, a_{t,n}, z_{t,n})\right] = \int_{z_t}\int_{s_{i,t-1}} U^n\left(a_{i,t,n}(s_{i,t-1}), a_{t,n}(z_t), z_{t,n}\right) dP(s_{i,t-1}|z_t)\, dP(z_t), \quad (71)$$

where $a_{t,n}(z_t) = \int_{s_{i,t-1}} a_{i,t,n}(s_{i,t-1})\, dP(s_{i,t-1}|z_t)$. Substituting equation (67) into equation (71) and using equation (69) gives

$$E\left[U^n(a_{i,t,n}, a_{t,n}, z_{t,n})\right] = \int_{z_t} W^n\left(a_{t,n}(z_t), \sigma_{a_{i,t,n}}, z_{t,n}\right) dP(z_t) . \quad (72)$$

Substituting equation (70) into the last equation yields

$$E\left[U^n(a_{i,t,n}, a_{t,n}, z_{t,n})\right] = E\left[W^n\left(a^*_{t,n}, 0, z_{t,n}\right)\right] + \frac{U^n_{a_i a_i} + 2U^n_{a_i a} + U^n_{aa}}{2} E\left[\left(a_{t,n} - a^*_{t,n}\right)^2\right] + \frac{U^n_{a_i a_i}}{2} E\left[(a_{i,t,n} - a_{t,n})^2\right] . \quad (73)$$

Noting that $W^n\left(a^*_{t,n}, 0, z_{t,n}\right) = \tilde{U}^n\left(a^*_{t,n}, z_{t,n}\right)$ gives the desired result.

C Proof of Proposition 2

Step 1: The first two sentences of Proposition 2 follow from Proposition 1. The next two sentences of Proposition 2 follow from the text above Proposition 2.

Step 2: Substituting $\kappa_2 = \kappa - \kappa_1$ into objective (33) and setting the first derivative of the objective with respect to κ_1 equal to zero yields the first-order condition

$$\begin{aligned} & p_1\delta_1\Sigma_{11}\left[\frac{(1-\gamma)^2 2^{2\kappa_1}}{\left[\gamma+(1-\gamma)2^{2\kappa_1}\right]^2} + 2\frac{(1-\gamma)2^{2\kappa_1}}{\left[\gamma+(1-\gamma)2^{2\kappa_1}\right]^3}\left(\frac{U_{aa}}{U_{a_ia_i}} - \gamma\right)\right] 2\ln(2) \\ & -p_2\delta_2\Sigma_{22}\left[\frac{(1-\gamma)^2 2^{2(\kappa-\kappa_1)}}{\left[\gamma+(1-\gamma)2^{2(\kappa-\kappa_1)}\right]^2} + 2\frac{(1-\gamma)2^{2(\kappa-\kappa_1)}}{\left[\gamma+(1-\gamma)2^{2(\kappa-\kappa_1)}\right]^3}\left(\frac{U_{aa}}{U_{a_ia_i}} - \gamma\right)\right] 2\ln(2) = 0. \end{aligned} \tag{74}$$

Let $F_{\kappa_1=0}$ and $F_{\kappa_1=\kappa}$ denote the value of the left-hand side of equation (74) at $\kappa_1 = 0$ and $\kappa_1 = \kappa$, respectively. When the constraint (34) is binding and the planner problem (33)-(34) is convex, the solution to the planner problem is given by

$$\kappa_1^{eff} = \begin{cases} \kappa & \text{if } F_{\kappa_1=\kappa} \geq 0 \\ \kappa_1^{FOC} & \text{if } F_{\kappa_1=0} > 0 > F_{\kappa_1=\kappa} \\ 0 & \text{if } F_{\kappa_1=0} \leq 0 \end{cases}, \tag{75}$$

where κ_1^{FOC} denotes the unique solution to equation (74) in the case of $F_{\kappa_1=0} > 0 > F_{\kappa_1=\kappa}$.

Step 3: If $\gamma = (U_{aa}/U_{a_ia_i})$, the first-order condition (74) reduces to

$$p_1\delta_1\Sigma_{11}\frac{(1-\gamma)^2\, 2^{2\kappa_1}}{\left[\gamma+(1-\gamma)\, 2^{2\kappa_1}\right]^2} 2\ln(2) - p_2\delta_2\Sigma_{22}\frac{(1-\gamma)^2\, 2^{2(\kappa-\kappa_1)}}{\left[\gamma+(1-\gamma)\, 2^{2(\kappa-\kappa_1)}\right]^2} 2\ln(2) = 0. \tag{76}$$

Now the condition $F_{\kappa_1=0} \leq 0$ reads

$$\sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}} \leq \frac{1}{\gamma 2^{-\kappa} + (1-\gamma)\, 2^{\kappa}},$$

and the condition $F_{\kappa_1=\kappa} \geq 0$ reads

$$\sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}} \geq \gamma 2^{-\kappa} + (1-\gamma)\, 2^{\kappa}.$$

Furthermore, solving equation (76) for κ_1 in the case of $F_{\kappa_1=0} > 0 > F_{\kappa_1=\kappa}$ yields

$$\kappa_1 = \frac{1}{2}\kappa + \frac{1}{2}\log_2\left(\frac{\sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}} - \frac{\gamma}{1-\gamma}2^{-\kappa}}{1 - \sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}}\frac{\gamma}{1-\gamma}2^{-\kappa}}\right).$$

Hence, if $\gamma = (U_{aa}/U_{a_ia_i})$, the efficient allocation of attention is given by

$$\kappa_1^{eff} = \begin{cases} \kappa & \text{if } \sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}} \geq \gamma 2^{-\kappa} + (1-\gamma)\, 2^{\kappa} \\ \frac{1}{2}\kappa + \frac{1}{2}\log_2\left(\frac{\sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}} - \frac{\gamma}{1-\gamma}2^{-\kappa}}{1-\sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}}\frac{\gamma}{1-\gamma}2^{-\kappa}}\right) & \text{otherwise} \\ 0 & \text{if } \sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}} \leq \frac{1}{\gamma 2^{-\kappa} + (1-\gamma)2^{\kappa}} \end{cases}. \tag{77}$$

Comparing equation (77) to equation (16) shows that if $\gamma = (U_{aa}/U_{a_ia_i})$ then $\kappa_1^{equ} = \kappa_1^{eff}$.

Step 4: If $\kappa_1^{equ} = \frac{1}{2}\kappa$, then $\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}} = 1$. See equation (16). Furthermore, when $\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}} = 1$, the first-order condition (74) reduces to

$$\begin{aligned} &\left[\frac{(1-\gamma)^2 2^{2\kappa_1}}{\left[\gamma+(1-\gamma)2^{2\kappa_1}\right]^2} + 2\frac{(1-\gamma)2^{2\kappa_1}}{\left[\gamma+(1-\gamma)2^{2\kappa_1}\right]^3}\left(\frac{U_{aa}}{U_{a_ia_i}} - \gamma\right)\right] \\ -&\left[\frac{(1-\gamma)^2 2^{2(\kappa-\kappa_1)}}{\left[\gamma+(1-\gamma)2^{2(\kappa-\kappa_1)}\right]^2} + 2\frac{(1-\gamma)2^{2(\kappa-\kappa_1)}}{\left[\gamma+(1-\gamma)2^{2(\kappa-\kappa_1)}\right]^3}\left(\frac{U_{aa}}{U_{a_ia_i}} - \gamma\right)\right] = 0. \end{aligned}$$

A solution to the last equation is $\kappa_1^{FOC} = \frac{1}{2}\kappa$. When the planner problem is convex, this implies that $\kappa_1^{eff} = \frac{1}{2}\kappa$. It follows that if $\kappa_1^{equ} = \frac{1}{2}\kappa$ then $\kappa_1^{equ} = \kappa_1^{eff}$.

D Proof of Proposition 3

If $\kappa_1^{equ} \in (0,\kappa)$, then

$$\kappa_1^{equ} = \frac{1}{2}\kappa + \frac{1}{2}\log_2(x), \tag{78}$$

where

$$x \equiv \frac{\sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}} - \frac{\gamma}{1-\gamma}2^{-\kappa}}{1-\sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}}\frac{\gamma}{1-\gamma}2^{-\kappa}}, \tag{79}$$

and

$$x \in \left(2^{-\kappa}, 2^{\kappa}\right). \tag{80}$$

See equations (16)-(17). Let $F_{\kappa_1=\kappa_1^{equ}\in(0,\kappa)}$ denote the value of the left-hand side of the planner's first-order condition (74) at $\kappa_1 = \kappa_1^{equ} \in (0,\kappa)$. Substituting equation (78) into the left-hand side of equation (74) gives

$$\begin{aligned} F_{\kappa_1=\kappa_1^{equ}\in(0,\kappa)} &= p_1\delta_1\Sigma_{11}\left[\frac{(1-\gamma)^2 2^{\kappa}x}{\left[\gamma+(1-\gamma)2^{\kappa}x\right]^2} + 2\frac{(1-\gamma)2^{\kappa}x}{\left[\gamma+(1-\gamma)2^{\kappa}x\right]^3}\left(\frac{U_{aa}}{U_{a_ia_i}} - \gamma\right)\right]2\ln(2) \\ &\quad -p_2\delta_2\Sigma_{22}\left[\frac{(1-\gamma)^2\frac{2^{\kappa}}{x}}{\left[\gamma+(1-\gamma)\frac{2^{\kappa}}{x}\right]^2} + 2\frac{(1-\gamma)\frac{2^{\kappa}}{x}}{\left[\gamma+(1-\gamma)\frac{2^{\kappa}}{x}\right]^3}\left(\frac{U_{aa}}{U_{a_ia_i}} - \gamma\right)\right]2\ln(2). \end{aligned}$$

Furthermore, equation (79) implies

$$p_1\delta_1\Sigma_{11}\frac{(1-\gamma)^2\,2^\kappa x}{\left[\gamma+(1-\gamma)\,2^\kappa x\right]^2}=p_2\delta_2\Sigma_{22}\frac{(1-\gamma)^2\,\frac{2^\kappa}{x}}{\left[\gamma+(1-\gamma)\,\frac{2^\kappa}{x}\right]^2}.$$

Substituting the last equation into the previous equation gives

$$\begin{aligned}F_{\kappa_1=\kappa_1^{equ}\in(0,\kappa)} &= p_1\delta_1\Sigma_{11}\frac{(1-\gamma)\,2^\kappa x}{\left[\gamma+(1-\gamma)\,2^\kappa x\right]^2}\\ &\quad\left[\frac{2}{\gamma+(1-\gamma)\,2^\kappa x}-\frac{2}{\gamma+(1-\gamma)\,\frac{2^\kappa}{x}}\right]\left(\frac{U_{aa}}{U_{a_ia_i}}-\gamma\right)2\ln(2)\,. \qquad (81)\end{aligned}$$

Since $p_1\delta_1\Sigma_{11}>0$, $\gamma\in(-1,1)$, and $x\in(2^{-\kappa},2^\kappa)$, the last expression equals zero if and only if $\frac{U_{aa}}{U_{a_ia_i}}=\gamma$ or $x=1$. Furthermore, when $\frac{U_{aa}}{U_{a_ia_i}}>\gamma$, then $x<1$ implies $F_{\kappa_1=\kappa_1^{equ}\in(0,\kappa)}>0$ while $x>1$ implies $F_{\kappa_1=\kappa_1^{equ}\in(0,\kappa)}<0$. By contrast, when $\frac{U_{aa}}{U_{a_ia_i}}<\gamma$, then $x<1$ implies $F_{\kappa_1=\kappa_1^{equ}\in(0,\kappa)}<0$ while $x>1$ implies $F_{\kappa_1=\kappa_1^{equ}\in(0,\kappa)}>0$. In addition, $x<1$ means $\kappa_1<\frac{1}{2}\kappa$, and $x>1$ means $\kappa_1>\frac{1}{2}\kappa$. See equation (78). Finally, by assumption $\kappa_1^{equ}\in(0,\kappa)$ and the planner problem is convex. Hence, when $\frac{U_{aa}}{U_{a_ia_i}}>\gamma$, then $\kappa_n<\frac{1}{2}\kappa$ implies $\kappa_n^{eff}>\kappa_n^{equ}$. By contrast, when $\frac{U_{aa}}{U_{a_ia_i}}<\gamma$, then $\kappa_n<\frac{1}{2}\kappa$ implies $\kappa_n^{eff}<\kappa_n^{equ}$.

References

[1] Angeletos, George-Marios and Alessandro Pavan (2007): "Efficient Use of Information and Social Value of Information." Econometrica, 75(4), 1103-1142.

[2] Barro, Robert J. (2006): "Rare Disasters and Asset Markets in the Twentieth Century." Quarterly Journal of Economics, 121(3), 823-866.

[3] Barro, Robert J., Emi Nakamura, Jón Steinsson, and José F. Ursua (2010): "Crises and Recoveries in an Empirical Model of Consumption Disasters." Discussion paper, Harvard University and Columbia University.

[4] Brunnermeier, Markus K. (2009): "Deciphering the Liquidity and Credit Crunch 2007-2008." Journal of Economic Perspectives, 23(1), 77-100.

[5] Coibion, Olivier and Yuriy Gorodnichenko (2011): "Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts." Discussion paper, College of William & Mary and UC Berkeley.

[6] Gabaix, Xavier (2010): "Variable Rare Disasters: An Exactly Solved Framework for Ten Puzzles in Macro-Finance." Discussion paper, New York University.

[7] Gourio, François (2010): "Disaster Risk and Business Cycles." Discussion paper, Boston University, forthcoming in American Economic Review.

[8] Hellwig, Christian and Laura Veldkamp (2009): "Knowing What Others Know: Coordination Motives in Information Acquisition." Review of Economic Studies, 76(1), 223-251.

[9] Kacperczyk, Marcin, Stijn Van Nieuwerburgh, and Laura Veldkamp (2011): "Rational Attention Allocation over the Business Cycle." Discussion paper, New York University.

[10] Llosa, Luis Gonzalo, and Venky Venkateswaran (2011): "Efficiency of Information Acquisition in a Price-Setting Model." Discussion paper, UCLA.

[11] Luo, Yulei (2008): "Consumption Dynamics under Information Processing Constraints." Review of Economic Dynamics, 11(2), 366-385.

[12] Maćkowiak, Bartosz, Emanuel Moench, and Mirko Wiederholt (2009): "Sectoral Price Data and Models of Price Setting." Journal of Monetary Economics, 56(S), 78-99.

[13] Maćkowiak, Bartosz and Mirko Wiederholt (2009): "Optimal Sticky Prices under Rational Inattention." American Economic Review, 99(3), 769-803.

[14] Maćkowiak, Bartosz and Mirko Wiederholt (2010): "Business Cycle Dynamics under Rational Inattention." Discussion paper, ECB and Northwestern University.

[15] Matejka, Filip (2010a): "Rationally Inattentive Seller: Sales and Discrete Pricing." Discussion paper, Princeton University and CERGE-EI.

[16] Matejka, Filip (2010b): "Rigid Pricing and Rationally Inattentive Consumers." Discussion paper, Princeton University and CERGE-EI.

[17] Melosi, Leonardo (2011): "Estimating Models with Information Frictions." Discussion paper, London Business School.

[18] Mondria, Jordi (2010): "Portfolio Choice, Attention Allocation, and Price Comovement." Journal of Economic Theory, 145(5), 1837-1864.

[19] Paciello, Luigi (2010): "Monetary Policy Activism and Price Responsiveness to Aggregate Shocks under Rational Inattention." Discussion paper, Einaudi Institute for Economics and Finance.

[20] Paciello, Luigi and Mirko Wiederholt (2011): "Exogenous Information, Endogenous Information and Optimal Monetary Policy." Discussion paper, Einaudi Institute for Economics and Finance and Northwestern University.

[21] Sims, Christopher A. (2003): "Implications of Rational Inattention." Journal of Monetary Economics, 50(3), 665-690.

[22] Sims, Christopher A. (2006): "Rational Inattention: Beyond the Linear-Quadratic Case." American Economic Review Papers and Proceedings, 96(2), 158-163.

[23] Sims, Christopher A. (2010): "Rational Inattention and Monetary Economics." Prepared for Handbook of Monetary Economics, Elsevier.

[24] Singh, Manmohan and Carolyne Spackman (2009): "The Use (and Abuse) of CDS Spreads During Distress." International Monetary Fund Working Paper 09/62.

[25] Tutino, Antonella (2011): "Rationally Inattentive Consumption Choices." Discussion paper, Federal Reserve Bank of Dallas.

[26] Van Nieuwerburgh, Stijn and Laura Veldkamp (2009): "Information Immobility and the Home Bias Puzzle." Journal of Finance, 64(3), 1187-1215.

[27] Van Nieuwerburgh, Stijn and Laura Veldkamp (2010): "Information Acquisition and Under-Diversification." Review of Economic Studies, 77(2), 779-805.

[28] Wessel, David (2009): "In Fed We Trust: Ben Bernanke's War on the Great Panic." Crown Business, New York.

[29] Woodford, Michael (2009): "Information-Constrained State-Dependent Pricing." Journal of Monetary Economics, 56(S), 100-124.

[30] Yang, Ming (2011): "Coordination with Rational Inattention." Discussion paper, Princeton University.

Figure 1: Probability of default, Lehman Brothers

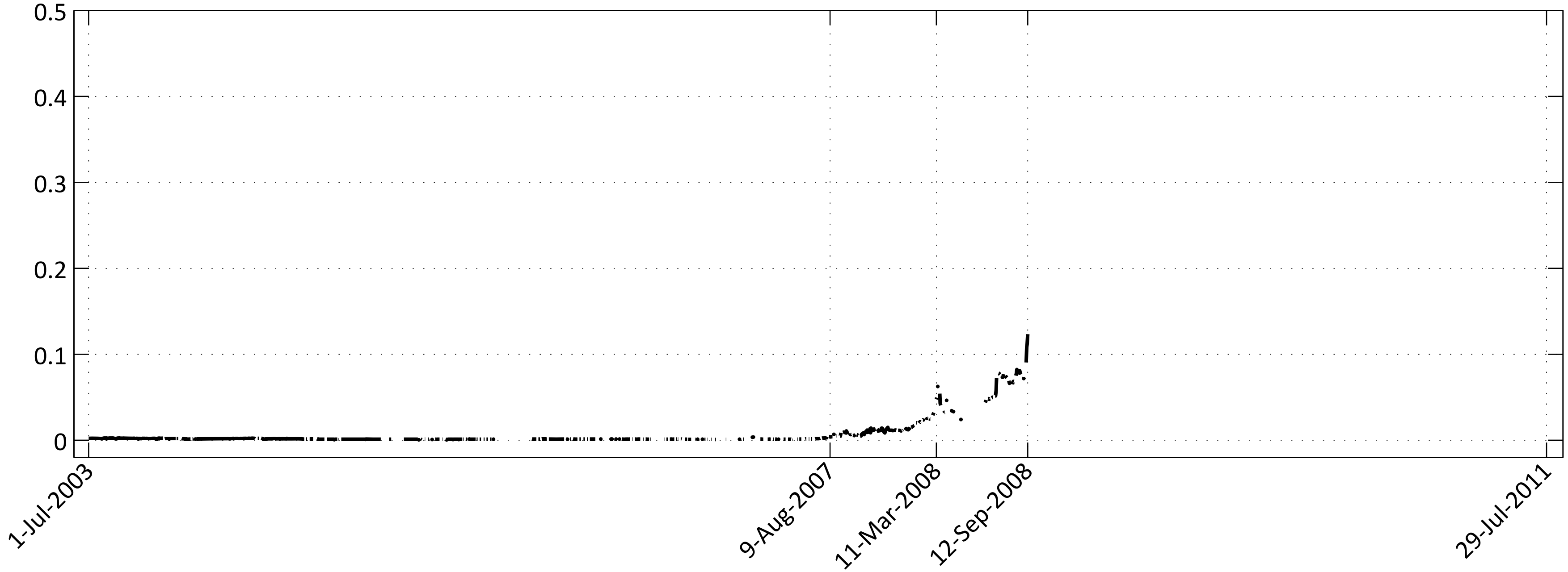
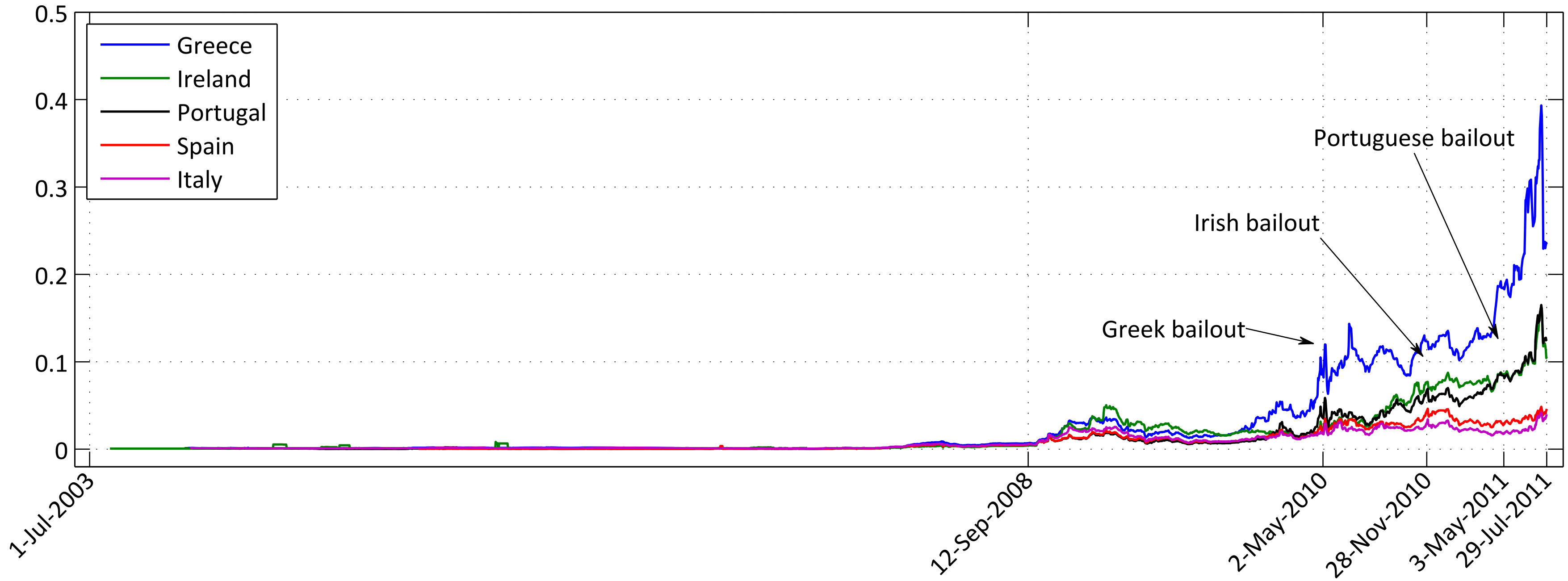


Figure 2: Probability of default, euro area sovereigns



Note: The probabilities of default in Figures 1-2 are derived from CDS premia. See Section 3.1 for the details.

Figure 3: Attention to state one as function of relative likelihood

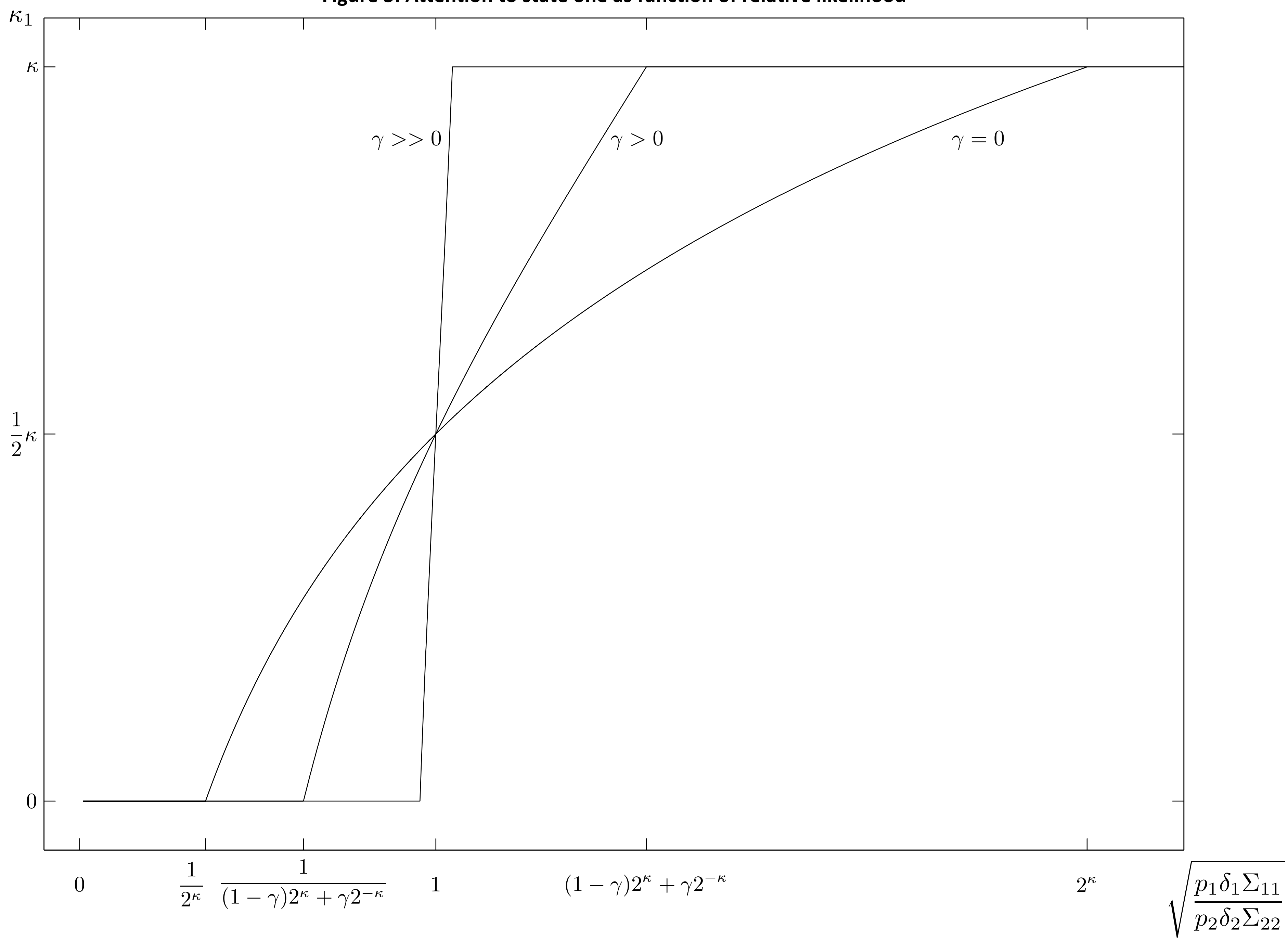
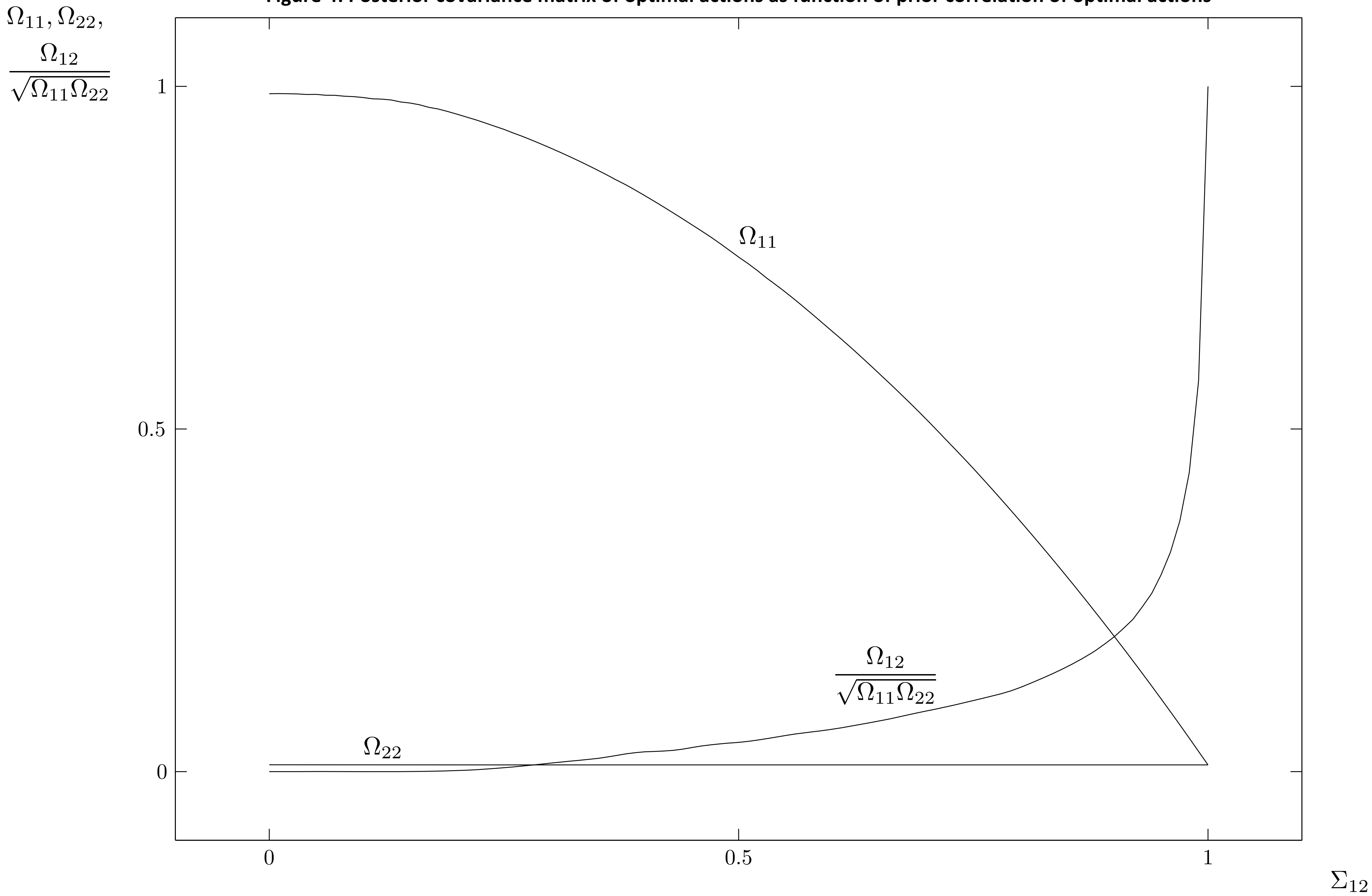


Figure 4: Posterior covariance matrix of optimal actions as function of prior correlation of optimal actions



This figure assumes: $\gamma_1 = \gamma_2 = 0, \delta_1 = \delta_2, \Sigma_{11} = \Sigma_{22} = 1, p_1 = 0.01$

Figure 5: Posterior expectation of the probability of unusual times