Inattention to Rare Events*

Bartosz Maćkowiak
European Central Bank and CEPR

Mirko Wiederholt
Goethe University Frankfurt and CEPR

First version: March 2011. This version: September 2013

Abstract

The world recently experienced several rare events with disastrous consequences: the global financial crisis, the European sovereign debt crisis, and the Fukushima nuclear accident. These events have in common that key decision-makers were unprepared for them, which aggravated these events. Should decision-makers think more about optimal actions in rare events? The paper studies 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 finite amount of information. In the model, different contingencies have different probabilities, mistakes may be more costly in some contingencies than in others, agents may face limited liability, and actions may be strategic complements. We identify the forces that make agents prepare little for rare events. We then study whether a social planner would want agents to be more prepared for rare events. We find that under reasonable assumptions this is the case.

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: Goethe University Frankfurt, Grueneburgplatz 1, 60323 Frankfurt am Main (e-mail: wiederholt@wiwi.uni-frankfurt.de). 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 our discussants Christian Hellwig, Alisdair McKay, and Chris Sims for helpful comments. We also thank Gadi Barlevy, Marco Bassetto, Cosmin Ilut, Jean-Paul L'Huillier, Guido Lorenzoni, Kristoffer Nimark, Laura Veldkamp, and seminar and conference participants at Bonn, CERGE-EI, Cologne Workshop on Macroeconomics, CREI, ECB, Federal Reserve Bank of Chicago, Goethe University Frankfurt, NBER Monetary Economics Program Meeting fall 2011, NBER Summer Institute 2011, NBER Universities Research Conference on Insurance Markets and Catastrophe Risk, Penn State, Rochester, SED 2011, and Toulouse for helpful comments. Giovanni Nicolò provided excellent research assistance. Mirko Wiederholt thanks the Federal Reserve Bank of Minneapolis for hospitality.

1 Introduction

The world recently experienced several rare events with disastrous consequences: the global financial crisis, the European sovereign debt crisis, and the Fukushima nuclear accident. These events have in common that key decision-makers were unprepared for them, which aggravated these events. Should decision-makers think more about optimal actions in unusual times, even if this means that they will think less about optimal actions in normal times?

To address this question formally, this paper studies 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 finite amount of information. In the model, the different contingencies have different probabilities, mistakes may be more costly in some contingencies than in others, agents may face limited liability, and actions may be strategic complements or strategic substitutes. We identify the forces that make agents prepare little for rare events. We then study whether a social planner would want agents to be more prepared for rare events. We find that under reasonable assumptions this is the case.

In the model, agents can process only a finite amount of information. Agents therefore cannot think perfectly about the optimal action in all contingencies. The expected benefit of thinking about a contingency is higher when the contingency is more likely (other things equal). Therefore, the extent to which agents think about a contingency is increasing in the probability of the contingency. The first-order condition for an optimal allocation of attention says: agents allocate attention so as to equate the *probability-weighted* expected loss due to suboptimal actions across contingencies. As a result, the expected loss due to suboptimal action in a contingency is inversely related to the probability of the contingency. If the probability of one contingency is one thousand times smaller than the probability of another contingency (think of the first contingency as a rare event and of the second contingency as normal times), the expected loss due to suboptimal action is one thousand times larger in the first contingency than in the second contingency. The observation that agents take good actions in normal times does not imply that agents will take good actions in unusual times.

The result stated above still holds when mistakes are more costly in some contingencies than in other contingencies. The optimal allocation of attention is still to equate the probability-weighted expected loss due to suboptimal actions across contingencies.

We allow for limited liability in the model because limited liability is a feature of many real world situations. We begin by introducing limited liability symmetrically across contingencies, that is, the extent of limited liability protection is the same in all contingencies. We find that this form of limited liability makes agents prepare even less for rare events. The intuition is the following. Since agents think less about the optimal actions in unusual times than about the optimal actions in normal times, agents take on average worse actions in unusual times than in normal times. Limited liability therefore is more relevant in unusual times than in normal times. Hence, limited liability reduces more strongly the incentive to think about unusual times than the incentive to think about normal times.

We also allow for strategic interactions in the model. Actions may be strategic complements or strategic substitutes. We begin by assuming that the degree of strategic complementarity in actions is the same for all contingencies. We obtain the following result. Start in a situation in which there are no strategic interactions and agents think less about the optimal actions in unusual times than about the optimal actions in normal times. Suppose that actions become strategic complements. Then agents think even less about rare events. Strategic complementarity reduces more strongly the incentive to think about unusual times than the incentive to think about normal times. This is true even though the degree of strategic complementarity is the same in all contingencies.

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. 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.

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

[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).

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 (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. 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.[2]

The rest of the paper is organized as follows. Section 2 presents the model. Section 3 studies the equilibrium allocation of attention. Section 4 looks at recent events from the perspective of the model. Section 5 studies the efficient allocation of attention. Section 6 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 $p_j > 0$ denote the probability of regime $j = r, n$. It is helpful to think of p_r as being close to zero and p_n as being close to one. In words, regime r is "unusual times" (or "the rare event") and regime n is "normal times." In the baseline model presented here, we assume that agents know the true values of p_r and p_n. As an extension, we will consider Bayesian learning about the probabilities of the two regimes.

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 $a_{i,t} = (a_{i,r,t}, a_{i,n,t}) \in \mathbb{R}^2$, where $a_{i,j,t}$ is the action that agent i will take in regime j in period t. Let $\Psi^{j,t}$ denote the cumulative distribution function for action $a_{i,j,t}$ in the cross-section of the population.

The payoff of agent i in regime j in period t is given by $U^j(a_{i,j,t}, a_{j,t}, z_{j,t})$ where $a_{i,j,t}$ is own

[2] 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.

action, $a_{j,t} \equiv \int a_{i,j,t} d\Psi^{j,t}(a_{i,j,t})$ is the mean action in the population, and $z_{j,t}$ is an exogenous fundamental. The superscript j indicates that the payoff function may differ across regimes. For tractability, we assume that the payoff function is quadratic:

$$\begin{aligned} U^j(a_{i,j,t}, a_{j,t}, z_{j,t}) &= U^j(0,0,0) + U^j_{a_i} a_{i,j,t} + U^j_a a_{j,t} + U^j_z z_{j,t} \\ &+ \frac{U^j_{a_i a_i}}{2} a^2_{i,j,t} + \frac{U^j_{aa}}{2} a^2_{j,t} + \frac{U^j_{zz}}{2} z^2_{j,t} \\ &+ U^j_{a_i a} a_{i,j,t} a_{j,t} + U^j_{a_i z} a_{i,j,t} z_{j,t} + U^j_{az} a_{j,t} z_{j,t}. \end{aligned} \quad (1)$$

This assumption can also be viewed as a second-order approximation of any twice differentiable function with the same three arguments. Furthermore, we assume that the payoff function is concave in its first argument ($U^j_{a_i a_i} < 0$), the fundamental affects the optimal action ($U^j_{a_i z} \neq 0$), and the degree of strategic complementarity or substitutability in actions is below one ($-1 < U^j_{a_i a}/U^j_{a_i a_i} < 1$). In the following, we use the fact that the payoff function can be expressed as[3]

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

where $a^*_{i,j,t}$ denotes the optimal action in regime j in period t

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

Finally, without loss of generality, we assume that the coefficients on $a_{j,t}$ and $z_{j,t}$ in the equation for the optimal action sum to one.[4] Defining $\varphi_j \equiv -U^j_{a_i}/U^j_{a_i a_i}$, $\gamma_j \equiv -U^j_{a_i a}/U^j_{a_i a_i}$, and $\delta_j \equiv -U^j_{a_i a_i}/2$, the last two equations then become

$$U^j(a_{i,j,t}, a_{j,t}, z_{j,t}) = U^j\left(a^*_{i,j,t}, a_{j,t}, z_{j,t}\right) - \delta_j\left(a_{i,j,t} - a^*_{i,j,t}\right)^2, \quad (2)$$

with

$$a^*_{i,j,t} = \varphi_j + \gamma_j a_{j,t} + \left(1 - \gamma_j\right) z_{j,t}. \quad (3)$$

For simplicity, the vector of fundamentals $z_t = (z_{r,t}, z_{n,t})$ is i.i.d. over time. Agents have the common prior belief that the vector of fundamentals is i.i.d. over time and normally distributed

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

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

with mean zero and covariance matrix Σ, $z_t = (z_{r,t}, z_{n,t}) \sim i.i.d.N(0, \Sigma)$. There is prior uncertainty about the fundamental and therefore about the optimal action in each regime. In the baseline model, we assume that the fundamentals are independent across regimes, i.e., Σ is diagonal. As an extension, we will relax this assumption.

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

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

where the noise $(\varepsilon_{i,r,t-1}, \varepsilon_{i,n,t-1})$ is independent of the fundamentals, independent across agents 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 that an agent can process more information and therefore reduce uncertainty by more.

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

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

subject to

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

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

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 commits 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 agents only have to solve this problem once.[5]

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 therefore cannot prepare a perfect action plan for each contingency.[6] 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. We begin by abstracting from strategic interactions and limited liability (Section 3.1). Afterwards, we study the effects of strate-

[5] 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).

[6] 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.

gic interactions (Section 3.2) and limited liability (Section 3.3) on the equilibrium allocation of attention. Finally, we consider two extensions of the baseline model (Sections 3.4-3.5).

3.1 The role of probabilities

In this subsection, we characterize the equilibrium allocation of attention in the special case when there are no strategic interactions and agents have unlimited liability. In particular, we derive a mapping from the odds of the rare event to the equilibrium allocation of attention. Furthermore, we show how the odds of the rare event determine the ratio of the expected loss in the rare event to the expected loss in normal times.

When there is no strategic complementarity or substitutability in actions, i.e., $\gamma_r = \gamma_n = 0$, the optimal action of an agent in a regime depends only on the fundamental in that regime, $a^*_{i,j,t} = \varphi_j + z_{j,t}$. The best plan of an agent given his or her posterior equals the conditional expectation of the optimal action, $a_{i,j,t} = \varphi_j + E\left[z_{j,t}|s_{i,t-1}\right]$. Thus, the expected loss in payoff in regime j due to suboptimal action in regime j is given by

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

Equation (8) states that the expected loss in regime j depends on the conditional variance of the fundamental in regime j, Ω_{jj}, and the cost of a mistake in regime j, δ_j. 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.

Since the fundamentals are independent across regimes, it is optimal to think independently about the optimal action in the rare event and the optimal action in normal times. This result is proved in the next subsection. Formally, the optimal covariance matrix of noise in the signal Λ is diagonal, and thus the posterior covariance matrix of the fundamentals Ω is diagonal. As a result, the information-processing constraint (7) reduces to

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

where Σ_{jj} and Ω_{jj} are the prior and posterior variance of the fundamental in regime $j = r, n$. Let

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

denote the attention devoted to regime j. It follows that $\Omega_{jj} = \Sigma_{jj} 2^{-2\kappa_j}$. When no attention is devoted to a regime ($\kappa_j = 0$), the posterior variance equals the prior variance. When positive attention is devoted to a regime ($\kappa_j > 0$), 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

$$\max_{(\kappa_r, \kappa_n) \in \mathbb{R}_+^2} \left(- \sum_{j=r,n} p_j \delta_j \Omega_{jj} \right), \tag{11}$$

subject to

$$\Omega_{jj} = \Sigma_{jj} 2^{-2\kappa_j}, \tag{12}$$

and

$$\kappa_r + \kappa_n \leq \kappa. \tag{13}$$

The unique solution to this problem is

$$\kappa_r = \begin{cases} \kappa & \text{if } \sqrt{\frac{p_r \delta_r \Sigma_{rr}}{p_n \delta_n \Sigma_{nn}}} \geq 2^{\kappa} \\ \frac{1}{2}\kappa + \frac{1}{2} \log_2 \left(\sqrt{\frac{p_r \delta_r \Sigma_{rr}}{p_n \delta_n \Sigma_{nn}}} \right) & \text{if } \sqrt{\frac{p_r \delta_r \Sigma_{rr}}{p_n \delta_n \Sigma_{nn}}} \in [2^{-\kappa}, 2^{\kappa}] \\ 0 & \text{if } \sqrt{\frac{p_r \delta_r \Sigma_{rr}}{p_n \delta_n \Sigma_{nn}}} \leq 2^{-\kappa} \end{cases}. \tag{14}$$

The equilibrium attention allocation is simple. If the ratio of $p_r \delta_r \Sigma_{rr}$ to $p_n \delta_n \Sigma_{nn}$ is equal to one, the attention allocation is fifty-fifty. Starting from this situation, reduce the probability of the rare event, p_r. Agents decide to think *less* about the optimal action in the rare event and *more* about the optimal action in normal times. The reason is that *the expected benefit of thinking about a contingency increases with the probability of that contingency.* Note that a corner solution is possible. If the rare event is sufficiently unlikely, agents decide to *not think at all* about the optimal action in the rare event. Finally, for agents 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 ($\delta_r > \delta_n$) or agents have to be sufficiently more uncertain about the optimal action in unusual times than about the optimal action in normal times ($\Sigma_{rr} > \Sigma_{nn}$).

The extent to which agents think about the optimal action in a regime affects the quality of actions in that regime. It follows from equations (12)-(14) that at an interior solution ($0 < \kappa_r < \kappa$)

$$p_r \delta_r \Omega_{rr} = p_n \delta_n \Omega_{nn}.$$

Agents allocate attention so as to equate the probability-weighted expected loss due to suboptimal action across regimes. Hence

$$\frac{\delta_r \Omega_{rr}}{\delta_n \Omega_{nn}} = \frac{1}{\frac{p_r}{pn}}. \tag{15}$$

The expected loss in the rare event divided by the expected loss in normal times is equal to one over the odds of the rare event. For example, if unusual times have a relative probability of 0.1 percent, the expected loss is *one thousand times* larger in unusual times than in normal times. Hence observing that agents take good actions in normal times does *not* imply that agents will take good actions in unusual times.

3.2 Strategic complementarity in actions

In this subsection, we study the effects of strategic interactions on the equilibrium allocation of the attention. The takeaway is as follows. Start in a situation in which there are no strategic interactions and agents pay *less* attention to the rare event than to normal times. Suppose that actions become strategic complements. Then agents will pay *even less* attention to the rare event. This is true even though the degree of strategic complementarity is *the same* in the two regimes.

Formally, we relax the assumption that $\gamma_r = \gamma_n = 0$. For ease of exposition, in the main text we assume that the degree of strategic complementarity or substitutability in actions is the same in unusual times and normal times, $\gamma_r = \gamma_n \equiv \gamma$.[7] When $\gamma > 0$ actions are strategic complements. An individual agent wants to do what other agents do. When $\gamma < 0$ actions are strategic substitutes. An individual agent wants to do the opposite of what other agents do.

The first part of the following proposition states that 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 $\gamma \in (-1, 1)$.

[7] In Appendix A, we cover the case when the degree of strategic complementarity or substitutability in actions differs across regimes.

Proposition 1 *Consider equilibria of the form $a_{j,t} = \psi_j + \phi_j z_{j,t}$ where ψ_j and ϕ_j are coefficients. Since the fundamentals are independent across regimes, each agent decides to receive independent signals about the fundamental in unusual times and the fundamental in normal times (i.e., the equilibrium Λ is diagonal), and the information-processing constraint (7) reduces to*

$$\underbrace{\frac{1}{2}\log_2\left(\frac{\Sigma_{rr}}{\Omega_{rr}}\right)}_{\kappa_r} + \underbrace{\frac{1}{2}\log_2\left(\frac{\Sigma_{nn}}{\Omega_{nn}}\right)}_{\kappa_n} \leq \kappa.$$

Furthermore, if the parameters κ and γ satisfy $2^{\kappa} > \gamma/(1-\gamma)$, the equilibrium is unique and

$$\kappa_r = \begin{cases} \kappa & \text{if } \sqrt{\frac{p_r\delta_r\Sigma_{rr}}{p_n\delta_n\Sigma_{nn}}} \geq (1-\gamma)\,2^{\kappa} + \gamma 2^{-\kappa} \\ \frac{1}{2}\kappa + \frac{1}{2}\log_2(x) & \text{if } \sqrt{\frac{p_r\delta_r\Sigma_{rr}}{p_n\delta_n\Sigma_{nn}}} \in \left[\frac{1}{(1-\gamma)2^{\kappa}+\gamma 2^{-\kappa}}, (1-\gamma)\,2^{\kappa} + \gamma 2^{-\kappa}\right] \\ 0 & \text{if } \sqrt{\frac{p_r\delta_r\Sigma_{rr}}{p_n\delta_n\Sigma_{nn}}} \leq \frac{1}{(1-\gamma)2^{\kappa}+\gamma 2^{-\kappa}} \end{cases}, \tag{16}$$

where

$$x \equiv \frac{\sqrt{\frac{p_r\delta_r\Sigma_{rr}}{p_n\delta_n\Sigma_{nn}}} - \frac{\gamma}{1-\gamma}2^{-\kappa}}{1 - \sqrt{\frac{p_r\delta_r\Sigma_{rr}}{p_n\delta_n\Sigma_{nn}}}\frac{\gamma}{1-\gamma}2^{-\kappa}}. \tag{17}$$

For any values of the parameters κ and γ, the set of equilibria is given in Appendix A.

Proof. See Appendix A. ■

Proposition 1 shows that raising the degree of strategic complementarity in both regimes makes the equilibrium attention allocation *more extreme* (if possible, i.e., if the attention allocation in the absence of strategic interactions is not already a corner solution). Figure 1 illustrates this result by depicting equilibrium attention to the rare event, κ_r, as a function of the square root of $(p_r\delta_r\Sigma_{rr}/p_n\delta_n\Sigma_{nn})$.[8] In the figure, $\gamma = 0$ denotes the case of no strategic interactions, $\gamma >> 0$ 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 point on the horizontal axis *to the left of 1.* In the absence of strategic interactions (i.e., $\gamma = 0$) agents think *less* about the optimal action in unusual times than about the optimal action in normal times (i.e., $\kappa_r < 0.5\kappa$). Raising the degree of strategic complementarity (from $\gamma = 0$ to $\gamma > 0$ and further to $\gamma >> 0$) makes agents think *even less* about the optimal action in unusual times (i.e., κ_r falls).

When actions are strategic complements, the fact that other agents do not think carefully about the optimal action in a regime reduces the incentive for an individual agent to think about

[8] Figure 1 assumes that the parameters κ and γ satisfy $2^{\kappa} > \gamma/(1-\gamma)$.

the optimal action in that regime. This effect is known in the literature.[9] We find that *this effect is stronger for the regime that agents think less about* than for the regime that agents think more about. Therefore, raising the degree of strategic complementarity makes the attention allocation more extreme. This is true although the degree of strategic complementarity is *the same* in the two regimes. In particular, the degree of strategic complementarity need *not* be greater in the regime that agents think less about than in the regime that agents think more about.

As the degree of strategic complementarity rises, *corner solutions occur more easily.* See Figure 1. In fact, for a high degree of strategic complementarity, a *small* change in parameters (e.g., a small change in the probability of the rare event) can have a *large* effect on the equilibrium allocation of attention. In particular, 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.[10]

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

In Appendix A, we characterize in closed form the set of equilibria arising when the degree of strategic complementarity or substitutability differs across regimes. The upshot is that raising the degree of strategic complementarity *in a single regime* makes agents allocate less attention to that regime and more attention to the other regime.

Let us conclude this subsection with the following summary. Whichever regime agents pay *less* attention to in the absence of strategic interactions, agents pay *even less* attention to when actions are strategic complements. Whichever regime agents pay *more* attention to in the absence of strategic interactions, agents pay *even more* attention to when actions are strategic complements. Strategic complementarity in actions makes the allocation of attention more extreme. This is true even though the degree of strategic complementarity is *the same* across regimes.

[9] See Hellwig and Veldkamp (2009) and Maćkowiak and Wiederholt (2009).

[10] 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 Appendix A for the details.

3.3 Limited liability

In this subsection, we study the effects of limited liability on the equilibrium allocation of the attention.

For ease of exposition, we assume that there is no constant in the equation for the optimal action and no strategic complementarity in actions ($\varphi_r = \varphi_n = 0$, $\gamma_r = \gamma_n = 0$, and thus $a^*_{i,j,t} = z_{j,t}$). For tractability, the payoff at the optimal action in a regime, $U^j\left(a^*_{i,j,t}, a_{j,t}, z_{j,t}\right)$, is assumed to be independent of both the fundamental and the average action in the regime. The payoff function with unlimited liability then becomes

$$U^j\left(a_{i,j,t}, z_{j,t}\right) = \bar{u}_j - \delta_j\left(a_{i,j,t} - z_{j,t}\right)^2,$$

where $\bar{u}_j$ is a constant. The payoff function with limited liability is

$$V^j\left(a_{i,j,t}, z_{j,t}\right) = \max\left\{U^j\left(a_{i,j,t}, z_{j,t}\right), \omega_j\right\}.$$

The expected payoff in regime j if the agent commits to action $a_{i,j,t}$ and has received signal $s_{i,t-1}$ equals

$$E\left[V^j\left(a_{i,j,t}, z_{j,t}\right)|s_{i,t-1}\right] = \int_{-\infty}^{\infty} \max\left\{\bar{u}_j - \delta_j\left(a_{i,j,t} - z\right)^2, \omega_j\right\} f\left(z|s_{i,t-1}\right) dz,$$

where $f\left(z|s_{i,t-1}\right)$ denotes the conditional density of the fundamental in regime j given the signal. Since limited liability kicks in if and only if the absolute distance between the action and the optimal action (i.e., the fundamental) exceeds $\Delta_j \equiv \sqrt{\frac{\bar{u}_j - \omega_j}{\delta_j}}$, the expected payoff in regime j can be written as

$$E\left[V^j\left(a_{i,j,t}, z_{j,t}\right)|s_{i,t-1}\right] = \omega_j + \int_{a_{i,j,t}-\Delta_j}^{a_{i,j,t}+\Delta_j} \left[\bar{u}_j - \delta_j\left(a_{i,j,t} - z\right)^2 - \omega_j\right] f\left(z|s_{i,t-1}\right) dz.$$

In the appendix we show that the action that maximizes this expression is the conditional mean of the fundamental, $a_{i,j,t} = E\left[z_{j,t}|s_{i,t-1}\right]$. Furthermore, in the appendix we also show that the maximized expected payoff is independent of the value of $E\left[z_{j,t}|s_{i,t-1}\right]$. We can therefore consider without loss in generality the special case where the density $f\left(z|s_{i,t-1}\right)$ has the property $E\left[z_{j,t}|s_{i,t-1}\right] = 0$. We arrive at the following expression for the maximized expected payoff in regime j

$$\max_{a_{i,j,t}\in\mathbb{R}} E\left[V^j\left(a_{i,j,t},z_{j,t}\right)|s_{i,t-1}\right]=\omega_j+\int_{-\Delta_j}^{\Delta_j}\left[\bar{u}_j-\delta_j z^2-\omega_j\right]f\left(z|s_{i,t-1}\right)dz.$$

One can separate the expected payoff in regime j into the expected payoff with unlimited liability and the expected benefit from limited liability

$$\begin{aligned}\max_{a_{i,j,t}\in\mathbb{R}} E\left[V^j\left(a_{i,j,t},z_{j,t}\right)|s_{i,t-1}\right] &= \bar{u}_j-\delta_j\Omega_{jj}\\ &+\int_{-\infty}^{-\Delta_j}\left[\omega_j-\left(\bar{u}_j-\delta_j z^2\right)\right]f\left(z|s_{i,t-1}\right)dz\\ &+\int_{\Delta_j}^{\infty}\left[\omega_j-\left(\bar{u}_j-\delta_j z^2\right)\right]f\left(z|s_{i,t-1}\right)dz.\end{aligned}$$

The first term on the right-hand side is the expected payoff with unlimited liability. The second term plus the third term is the expected benefit from limited liability. In the following, we denote the expected benefit from limited liability in regime j by $B\left(\Omega_{jj},\bar{u}_j-\omega_j,\delta_j\right)$. Here we recognize that the expected benefit from limited liability depends only on Ω_{jj}, $\bar{u}_j-\omega_j$ and δ_j, which is shown in the appendix. In the following, we use that the expected benefit from limited liability in regime j has the following properties.

Lemma 1 *The expected benefit from limited liability in regime j equals*

$$B\left(\Omega_{jj},\bar{u}_j-\omega_j,\delta_j\right)=\int_{-\infty}^{-\Delta_j}\left[\omega_j-\left(\bar{u}_j-\delta_j z^2\right)\right]f\left(z|s_{i,t-1}\right)dz+\int_{\Delta_j}^{\infty}\left[\omega_j-\left(\bar{u}_j-\delta_j z^2\right)\right]f\left(z|s_{i,t-1}\right)dz,$$

where $\Delta_j=\sqrt{\frac{\bar{u}_j-\omega_j}{\delta_j}}$, and has the following properties:

- $\frac{\partial B(\Omega_{jj},\bar{u}_j-\omega_j,\delta_j)}{\partial\Omega_{jj}}\in\left(0,\delta_j\right)$ *if* $\Delta_j>0$,
- $\frac{\partial^2 B(\Omega_{jj},\bar{u}_j-\omega_j,\delta_j)}{\partial\Omega_{jj}\partial\omega_j}>0$ *if* $\Delta_j>0$,
- $\frac{\partial^2 B(\Omega_{jj},\bar{u}_j-\omega_j,\delta_j)}{\partial^2\Omega_{jj}}>0$ *if* $\Delta_j\geq 1.732\sqrt{\Omega_{jj}}$,
- $\frac{\partial^3 B(\Omega_{jj},\bar{u}_j-\omega_j,\delta_j)}{\partial^2\Omega_{jj}\partial\omega_j}>0$ *if* $\Delta_j>1.732\sqrt{\Omega_{jj}}$ *and* $\frac{\partial^3 B(\Omega_{jj},\bar{u}_j-\omega_j,\delta_j)}{\partial^2\Omega_{jj}\partial\omega_j}<0$ *if* $\Delta_j\in\left(0,1.732\sqrt{\Omega_{jj}}\right)$.

Proof. See Appendix B. ■

The optimal attention allocation in the case of limited liability is the solution to

$$\max_{(\kappa_r,\kappa_n)\in\mathbb{R}_+^2}\left(\sum_{j=r,n} p_j\left[\bar{u}_j-\delta_j\Omega_{jj}+B\left(\Omega_{jj},\bar{u}_j-\omega_j,\delta_j\right)\right]\right), \tag{18}$$

subject to

$$\Omega_{jj}=\Sigma_{jj}2^{-2\kappa_j}, \tag{19}$$

and

$$\kappa_r+\kappa_n\leq\kappa. \tag{20}$$

Using the fact that the attention constraint (20) is binding and substituting constraints (19)-(20) into the objective yields

$$\max_{\kappa_r\in[0,\kappa]} g\left(\kappa_r,\theta\right), \tag{21}$$

where

$$\begin{aligned} g\left(\kappa_r,\theta\right) &= p_r\left[\bar{u}_r-\delta_r\Sigma_{rr}2^{-2\kappa_r}+B\left(\Sigma_{rr}2^{-2\kappa_r},\bar{u}_r-\omega_r,\delta_r\right)\right] \\ &\quad +p_n\left[\bar{u}_n-\delta_n\Sigma_{nn}2^{-2(\kappa-\kappa_r)}+B\left(\Sigma_{nn}2^{-2(\kappa-\kappa_r)},\bar{u}_n-\omega_n,\delta_n\right)\right]. \end{aligned}$$

The following proposition describes how limited liability affects the optimal allocation of attention.

Proposition 2 *(The effect of limited liability) Let κ_r^{LL} and κ_r^{UL} denote the attention allocated to the rare event in the case of limited liability and in the case of unlimited liability, respectively. Furthermore, let $\Omega_{rr}^{UL}=\Sigma_{rr}2^{-2\kappa_r^{UL}}$ denote the optimal posterior variance under unlimited liability. If $\bar{u}_r-\omega_r=\bar{u}_n-\omega_n$, $\delta_r=\delta_n$, $\Sigma_{rr}\geq\Sigma_{nn}$, and $\Delta_r\geq 1.732\sqrt{\Omega_{rr}^{UL}}$, limited liability reduces the attention allocated to the rare event:*

$$\kappa_r^{UL}\in(0,\kappa)\Rightarrow\kappa_r^{LL}<\kappa_r^{UL}.$$

Proof. See Appendix C. ■

3.4 Extension: Correlated optimal actions

In the rest of this section, we consider two extensions of the baseline model studied so far. In this subsection, we relax the assumption that optimal actions are independent across regimes. The upshot is as follows. Start in a situation in which the optimal actions are independent across regimes, the odds of the rare event are small, and the expected loss in the rare event is larger than the expected loss in normal times. Suppose that the optimal actions become correlated. Then the expected loss in the rare event *falls*. More important, the expected loss in the rare event falls *little* so long as the optimal actions are not *strongly* correlated. In other words, so long as the optimal actions are not strongly correlated, the equilibrium differs little from the equilibrium in the case when the optimal actions are independent.

Formally, the decision problem of an individual agent is still given by expressions (4)-(7), except that the prior covariance matrix of the fundamentals Σ is nondiagonal.[11] We solve problem (4)-(7) numerically assuming different values of the covariance between the optimal actions in the two regimes, Σ_{rn}. For simplicity, we consider the case of no strategic interactions, i.e., $\gamma_r = \gamma_n = 0$.

Recall that the agent chooses the covariance matrix of noise in the signal, Λ. When Σ is nondiagonal, we find that the agent chooses a nondiagonal Λ. In words, the agent chooses a signal vector such that each signal contains some information about the optimal action in normal times and some information about the optimal action in unusual times. When the agent thinks about the optimal action in normal times, the agent learns something about the optimal action in unusual times, and vice versa.

Consider a numerical example. We set $\delta_r = \delta_n$, $\Sigma_{rr} = \Sigma_{nn} = 1$, $p_r = 0.01$, and we choose a value of κ such that the posterior variance of the optimal action in normal times, Ω_{nn}, is equal to 0.01 in the case when the optimal actions are independent across regimes ($\Sigma_{rn} = 0$).[12] Figure 2 shows how the solution of the model changes as we raise Σ_{rn} from zero (no prior correlation of the optimal actions) towards one (perfect prior correlation of the optimal actions), holding the other parameters constant. We note the following results: (1) As the prior correlation of the optimal actions becomes *stronger*, the expected loss in the rare event Ω_{rr} *falls*. The stronger the prior correlation of the optimal actions, the more agents learn about what to do in unusual times by thinking about what to do in normal times, and therefore the better agents do on average in the rare event.[13] (2) This effect is nonlinear and sets in *slowly*, i.e., Ω_{rr} is concave in Σ_{rn}. So long as the

[11] We assume that Σ is nonsingular, i.e., the fundamentals are not perfectly correlated across regimes.

[12] In other words, this value of κ means that agents choose to reduce the variance of the optimal action in normal times by a factor of 100.

[13] The expected loss in normal times Ω_{nn} also decreases with Σ_{rn}.

optimal actions are not *strongly* correlated, the expected loss in the rare event falls *little* compared with the case when the optimal actions are independent. What is the source of this nonlinearity? As Figure 2 shows, the posterior correlation of the actual actions, $\Omega_{rn}/\sqrt{\Omega_{rr}\Omega_{nn}}$, is convex in Σ_{rn}. This means that, so long as Σ_{rn} is not large, agents choose signals that contain lots of information about what to do in normal times and little information about what do in the rare event. Hence, so long as Σ_{rn} is not large, the agents' actual actions are correlated but barely so. The expected loss in the rare event falls little compared with the case when the optimal actions are independent.

3.5 Extension: Learning the probability of a rare event

The baseline model assumes that agents know the true probability of the economy being in a particular regime next period. In this subsection, we study a version of the model in which the probability of the economy being in a particular regime next period is a random variable. The following insights emerge. Start in a situation in which agents know the true probabilities and agents pay *less* attention to the rare event than to unusual times. Suppose that agents are Bayesians who must infer the true probabilities over time. When the rare event fails to occur, agents pay *even less* attention to the rare event because the rational estimate of the probability of the rare event falls. Furthermore, once the rare event takes place, agents pay *a lot more* attention to the rare event because the rational estimate of the probability of the rare event occuring again jumps up.

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 attention problem (4)-(7), except that in objective (4) the probability of the economy being in a particular regime next period has been replaced by the agents' posterior expectation of that probability.[14]

Here is 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.[15] The agents' posterior expectation of p evolves over time as shown in Figure 3. Note that between period 1 and period 100, the agents' posterior expectation of p falls slowly. Just before the rare events occurs, the agents' posterior expectation of p is equal to 0.005. Agents underestimate the probability of the rare event by fifty percent. Consequently, agents think *even less* about the optimal action in the rare event. Next, observe that just after the rare event occurs the agents' posterior expectation of p changes by a large amount. The agents' posterior expectation of p doubles to 0.01. Consequently, the occurrence of the rare event causes a large reallocation of attention toward thinking about what to do in the rare event, because agents now find it much more likely that the rare event will occur again.

4 Applications

In this section we use the model to understand the recent events: the global financial crisis, the European sovereign debt crisis, and the Fukushima nuclear accident.

4.1 Global financial crisis

Let us focus on the defining moment of the global financial crisis which came when Lehman Brothers filed for bankruptcy on September 15, 2008. We think of "unusual times" as the regime in which

[14] To evaluate the agents' objective when agents are uncertain about p, in general one must keep track of the agents' posterior distribution of p and perform integration with respect to p. However, if the agents' prior distribution of p and the stochastic process $\{X_t\}$ are independent of the stochastic process $\{z_t, \varepsilon_{i,t}\}$, the agents' objective reduces to expression (4) except that the probability of the economy being in a particular regime next period must be replaced by the agents' posterior expectation of that probability.

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

an investment bank like Lehman Brothers has a sizable negative net present value. We add the word "sizable" to make it clear that in this regime Lehman Brothers cannot be rescued with a small amount of public support; a large amount of public support is necessary. "Normal times" is the regime in which Lehman Brothers has a positive net present value, or at worst a negative net present value close to zero.

We consider the actions of U.S. policy-makers. The policy-makers were uncertain about their optimal action in each regime. The possible actions in normal times were "don't intervene" and "orchestrate a sale of Lehman to another financial institution, possibly with a small amount of public support." To take the optimal action in normal times the policy-makers needed to process information about a few potential buyers of Lehman, the price at which a sale would occur, and the details of any subsidy. On the other hand, the possible actions in unusual times were "don't intervene" and "offer a large amount of public support of some form." Crucially, assessing the consequences of the "don't intervene" action is a very different thought process in unusual times than in normal times. In normal times the shutdown of Lehman will not trigger bankruptcies of other financial institutions. By contrast, in unusual times when Lehman has a sizable negative net present value, its bankruptcy is likely to trigger other bankruptcies. Hence to prepare for unusual times the policy-makers had to think about an entire network of closely connected financial institutions. There was uncertainty about the structure of the network ("who owns whom how much") and about how much more losses financial institutions could still absorb.[16]

The policy-makers reduced their uncertainty as they processed information about what to do in each regime. Our reading of the events is that the policy-makers thought carefully about what to do in normal times. In particular, the policy-makers prepared to orchestrate a sale of Lehman Brothers to another financial institution (like Bank of America or Barclays) with a small amount of public support. Importantly, almost the entire meeting of the policy-makers and bankers at the Federal Reserve Bank of New York on the weekend of September 13-14, 2008, was devoted to planning a sale of Lehman Brothers. By contrast, little time during the meeting was spent thinking about what to do if the hole in Lehman's balance sheet was too deep for Lehman to be sold with a small subsidy. Timothy Geithner, then president of the Federal Reserve Bank of New York, asked

[16] For recent models featuring uncertainty about a financial network, see Caballero and Simsek (2012) and Alvarez and Barlevy (2013).

one of the working groups formed at the meeting to “put foam on the runway,” in case Lehman’s sale could not be orchestrated, and “be prepared to do *something*.”[17]

As the weekend drew to a close, it turned out that unusual times had occurred: Lehman Brothers had a sizable negative net present value and therefore could not be sold with only a small amount of public support. The policy-makers had thought little about what to do in that regime. They had to decide quickly and chose to take the action “don’t intervene.” Lehman filed for bankruptcy. Within days, the policy-makers reversed themselves as they offered a large amount of public support to American International Group, money market funds, and so on. We take this policy reversal as an indication that the optimal action on the weekend of September 13-14, 2008, would have been a different one.

The model proposes the following explanation for why the policy-makers were unprepared for the regime “Lehman Brothers has a sizable negative net present value.” First, this regime was a low-probability event (p_r close to zero). Second, actions of different policy-makers were strategic complements ($\gamma > 0$). Any policy decision had to be made by a committee, and therefore each individual policy-maker had an incentive to propose a policy action acceptable to other policy-makers. Third, the policy-makers faced limited liability because their punishment would be bounded in the event of failure ($\omega > 0$). Fourth, the optimal actions were at most weakly correlated across the two regimes (Σ_{rn} close to zero). The low probability made the policy-makers think little about the contingency “Lehman Brothers has a sizable negative net present value.” Strategic complementarity and limited liability amplified the asymmetry in the allocation of attention. Once Lehman Brothers turned out to have a sizable negative net present value, all the thinking that went into orchestrating a sale of Lehman was of little use.

Of course, assessing the quality of policy-makers’ actions is difficult. Even today we cannot be certain what the optimal action on the weekend of September 13-14, 2008, was. However, the fact that the policy-makers reversed themselves so dramatically indicates that their initial action was far from optimal.

[17] We quote Geithner after Wessel (2009, p.17), with emphasis added. See Wessel (2009) for a description of the meeting at the Federal Reserve Bank of New York on September 13-14, 2008.

4.2 European sovereign debt crisis

We focus on what we see as the defining moment of the European sovereign debt crisis which came in April 2010 when the prime minister of Greece asked other euro-area member states for help in resolving Greece's fiscal crisis.[18] We think of "normal times" as the regime in which the government of a euro-area country is solvent but may be illiquid, i.e., may be unable to roll over its maturing debt. "Unusual times" is the regime in which the government of a euro-area country is insolvent.[19]

We consider the actions of euro-area policy-makers who had to respond to the request of Greece's prime minister. Greece entered the post-Lehman era with a large amount of government debt. In October 2009, a new Greek government announced that the fiscal situation was a lot worse than had previously been understood. The immediate problem was that a sizable amount of public debt was due to mature in May 2010. The new government announced a package of fiscal reforms. As May 2010 approached everyone waited to see if the fiscal reforms would "work" and, consequently, if the regime would turn out to be "normal times" or "unusual times." The possible actions of the euro-area policy-makers in normal times were "don't intervene" and "make Greece a loan." Possible actions in unusual times were likewise "don't intervene" and "make Greece a loan," but possible actions in unusual times also included "give Greece a transfer," "guarantee Greek government debt," and "help Greece organize an orderly default."

Importantly, preparing for unusual times is a very different activity than preparing for normal times. Preparation for normal times involves figuring out the size and conditions of a loan. On the other hand, figuring out how to make government default orderly is a very different task from designing the conditions of a loan. Figuring out the modalities of an inter-country transfer or an inter-country debt guarantee is also a very different thought process because it would change the way European Monetary Union works.

The euro-area policy-makers reduced their uncertainty as they processed information about what to do in each regime. Far-reaching options such as different forms of a public debt guarantee were on the table. However, our interpretation of the events is that the policy-makers spent most of the time between October 2009 and April 2010 thinking about the size and conditions of any

[18] See Bastasin (2012), IMF (2013), and Irwin (2013) for a chronology of the fiscal crisis in Greece.

[19] By the government being insolvent we mean that government debt exceeds the present value of primary budget surpluses in the absence of reform and under any politically feasible reform.

loan to Greece. In particular, much attention was given to figuring out the interest rate on the loan and designing the reform measures Greece would have to promise in order to get the loan. By contrast, during that period (or any time after the creation of the euro) we are not aware of any planning for an orderly default by the government of a euro-area country.

In the end, the news coming from Greece between October 2009 and April 2010 turned out to be bad. Greece found itself in unusual times. In response to the prime minister's request in April 2010, the euro-area policy-makers together with the International Monetary Fund made the Greek government a loan on May 2, 2010. By October 2010, the chancellor of Germany and the president of France decided that government default would have to be an option in the euro area. Preparation for an orderly default by Greece began. In October 2011, a new assistance package for Greece was announced, this time including provisions for default. We take this policy reversal as an indication that the optimal action in the spring of 2010 would have been a different one.

The model proposes the following explanation for why the policy-makers were unprepared for the regime "the government of a euro-area country is insolvent." This regime was a low-probability event. The low probability made the policy-makers think little about the contingency "the government of a euro-area country is insolvent." Strategic complementarity and limited liability amplified the asymmetry in the allocation of attention. Once Greece turned out to be insolvent, all the thinking that went into designing the optimal conditionality for a loan to Greece was of little use.

4.3 Fukushima nuclear accident

We define "unusual times" as the regime in which an earthquake and tsunami disable the cooling system of a nuclear power plant.

The 9.0-magnitude earthquake that struck off the coast of Japan on March 11, 2011 cut all off-site power supply to the Fukushima Dai-ichi nuclear power plant, owned and operated by Tokyo Electric Power Company (Tepco). The ensuing tsunami waves knocked out all of the plant's emergency diesel generators apart from one. After the earthquake and tsunami had cut power supply and thereby disabled the cooling system, workers at the plant tried to avoid a catastrophe. The most severe problem was that the fuel rods inside the reactors were overheating, causing a buildup of steam and hydrogen inside the reactor buildings, which 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. The workers opened the emergency manual and discovered that it 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 ad hoc ways to vent the reactor in the absence of electricity.[20] At about 2:30pm on March 12, the operators confirmed a decrease in pressure inside the reactor, providing some indication that the improvised venting was starting to work. Unfortunately, this good news came too late. Shortly thereafter, a hydrogen explosion destroyed the Unit 1 reactor building.[21]

What happened was *not* that the Tepco staff 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 Tepco staff had not thought about what to do if the cooling system were disabled: The emergency manual contained *no* instructions on how to vent a nuclear reactor in the absence of electricity and therefore the workers on site had to improvise corrective measures. This corner allocation of attention ($\kappa_r = 0$) is consistent with the model if p_r is sufficiently close to zero.

The model provides the following explanation for why Tepco staff were unprepared for the regime "an earthquake and tsunami disable the cooling system of a nuclear reactor." First, this regime was a low-probability event. Second, optimal day-to-day actions in a nuclear power plant are at most weakly correlated with optimal emergency actions. Thinking carefully about how to run a nuclear power plant optimally in normal times fails to improve actions in times when an earthquake and tsunami have disabled the plant's cooling system.[22] Third, Tepco officials face limited liability. Fourth, actions of decision-makers in different Japanese nuclear power companies were strategic complements. The managers in the different nuclear power companies in Japan were subject to relative performance evaluation, like most managers in the world. Therefore, each manager knew that he or she would be punished less if he or she failed at a time when other managers were failing too. For instance, a Tepco manager could claim that the managers in other nuclear power companies in Japan had also failed to think about what to do if an earthquake and tsunami disable

[20] 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.

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

[22] *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."

the cooling system of a reactor.

4.4 Low-probability events, *not* unthinkable events

It is sometimes argued that the recent events were unthinkable, zero-probability events. Another popular view is that the probability of each of the recent events was impossible to estimate. The evidence suggests otherwise.

The probability of default by Lehman Brothers and the probability of default by Greece were simple to estimate, at least crudely, based on publicly available data. In both cases, publicly available data suggested that the probability of default was small but strictly greater than zero. Figure 4 plots the probability of default on one-year senior debt of Lehman, based on credit default swap (CDS) premia.[23] Prior to August 9, 2007, the day on which the interbank market froze up, the probability of default by Lehman was 0.002 on average. The probability of Lehman's default between August 9, 2007, and the last day on which the Lehman CDS was traded, September 12, 2008, was 0.03 on average. An event with a probability of 0.03 is a low-probability event but it is not unthinkable. Figure 5 plots the probability of default on one-year government debt of Greece, likewise based on CDS premia.[24] Prior to September 15, 2008, the day of Lehman's bankruptcy, the probability of default by the Greek government was 0.002 on average. The probability of Greek default between September 15, 2008, and the day on which the first assistance package for Greece was agreed, May 2, 2010, was 0.03 on average. The similarity between Figure 5 and Figure 4 is striking.

The probability of the combination of a 9.0-magnitude earthquake and tsunami near Fukushima could not be estimated based on financial market data. However, this earthquake-tsunami combi-

[23] To produce Figure 4, 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 4, 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 4.

[24] To produce Figure 5, we took from Datastream CDS premia on one-year government debt of Greece, at daily frequency, from the beginning of 2004 to May 2, 2010, the day on which the first assistance package for Greece was agreed. We computed the probability of default, plotted in Figure 5, 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.

nation had a well-known precedent. The so-called Jogan earthquake of 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 Tepco officials attended this meeting.[25] Thus the earthquake-tsunami combination of March 11, 2011, was a low-probability but not an unthinkable event and this was known inside Tepco.

5 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 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

$$\begin{aligned} \tilde{U}^n\left(a_{t,n}, z_{t,n}\right) &= U^n(0,0,0) + \left(U^n_{a_i} + U^n_a\right) a_{t,n} + U^n_z z_{t,n} \\ &\quad + \frac{U^n_{a_i a_i} + 2U^n_{a_i a} + U^n_{aa}}{2} a^2_{t,n} + \frac{U^n_{zz}}{2} z^2_{t,n} + \left(U^n_{a_i z} + U^n_{az}\right) a_{t,n} z_{t,n}. \end{aligned} \tag{22}$$

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,

$$U^n_{a_i a_i} + 2U^n_{a_i a} + U^n_{aa} < 0. \tag{23}$$

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

Let $a_{t,n}^*$ denote the common action $a_{t,n} \in \mathbb{R}$ that maximizes $\tilde{U}^n(a_{t,n}, z_{t,n})$. It follows from equations (22) and (23) that

$$a_{t,n}^* = -\frac{U_{a_i}^n + U_a^n}{U_{a_i a_i}^n + 2U_{a_i a}^n + U_{aa}^n} - \frac{U_{a_i z}^n + U_{az}^n}{U_{a_i a_i}^n + 2U_{a_i a}^n + U_{aa}^n} z_{t,n}. \tag{24}$$

One can show that expected utility in regime n equals

$$\begin{aligned} E\left[U^n(a_{i,t,n}, a_{t,n}, z_{t,n})\right] &= E\left[\tilde{U}^n\left(a_{t,n}^*, z_{t,n}\right)\right] - \frac{\left|U_{a_i a_i}^n + 2U_{a_i a}^n + U_{aa}^n\right|}{2} E\left[\left(a_{t,n} - a_{t,n}^*\right)^2\right] \\ &\quad - \frac{\left|U_{a_i a_i}^n\right|}{2} E\left[\left(a_{i,t,n} - a_{t,n}\right)^2\right]. \end{aligned} \tag{25}$$

The proof is in Appendix D. 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:

$$\max_{(\kappa_1,\kappa_2)\in\mathbb{R}_+^2} \sum_{n=1}^{2} p_n \left\{ \frac{U_{a_i a_i}^n + 2U_{a_i a}^n + U_{aa}^n}{2} E\left[\left(a_{t,n} - a_{t,n}^*\right)^2\right] + \frac{U_{a_i a_i}^n}{2} E\left[\left(a_{i,t,n} - a_{t,n}\right)^2\right] \right\}, \tag{26}$$

subject to equation (24),

$$a_{t,n} = \psi_n + \phi_n z_{t,n}, \tag{27}$$

$$a_{i,t,n} = (\varphi_n + \gamma_n \psi_n) + (\gamma_n \phi_n + 1 - \gamma_n) \frac{\frac{\Sigma_{nn}}{\Lambda_{nn}}}{\frac{\Sigma_{nn}}{\Lambda_{nn}} + 1} \left(z_{t,n} + \varepsilon_{i,t-1,n}\right), \tag{28}$$

$$\psi_n = \frac{\varphi_n}{1-\gamma_n}, \phi_n = \frac{(1-\gamma_n)\frac{\frac{\Sigma_{nn}}{\Lambda_{nn}}}{\frac{\Sigma_{nn}}{\Lambda_{nn}}+1}}{1-\gamma_n \frac{\frac{\Sigma_{nn}}{\Lambda_{nn}}}{\frac{\Sigma_{nn}}{\Lambda_{nn}}+1}}, \tag{29}$$

$$\frac{\Sigma_{nn}}{\Lambda_{nn}} = 2^{2\kappa_n} - 1, \tag{30}$$

and

$$\kappa_1 + \kappa_2 \leq \kappa. \tag{31}$$

Objective (26) 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 (28) follows from equations (5)-(6) and equation (27).

Equation (29) follows from equations (27)-(28), the definition of $a_{t,n}$, and the assumption that noise washes out in the aggregate. Equation (30) follows from the definition $\kappa_n \equiv \frac{1}{2}\log_2\left(\frac{\Sigma_{nn}}{\Omega_{nn}}\right)$ and $\Omega_{nn} = \Sigma_{nn} - \Sigma_{nn}\left(\Sigma_{nn} + \Lambda_{nn}\right)^{-1}\Sigma_{nn}$. Finally, constraint (31) 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 (24). The condition that the equilibrium actions under perfect information equal the welfare-maximizing actions thus reads

$$-\frac{U^n_{a_i} + U^n_a}{U^n_{a_i a_i} + 2U^n_{a_i a} + U^n_{aa}} = \frac{\varphi_n}{1-\gamma_n}, \tag{32}$$

and

$$-\frac{U^n_{a_i z} + U^n_{az}}{U^n_{a_i a_i} + 2U^n_{a_i a} + U^n_{aa}} = 1. \tag{33}$$

Substituting equation (24), equations (27)-(30), and equations (32)-(33) into the planner's objective (26) gives

$$\max_{(\kappa_1,\kappa_2)\in\mathbb{R}^2_+} -\sum_{n=1}^{2} p_n \delta_n \Sigma_{nn}\left[\left(1 - 2\gamma_n + \frac{U^n_{aa}}{U^n_{a_i a_i}}\right)\frac{1}{\left(\gamma_n + (1-\gamma_n)\,2^{2\kappa_n}\right)^2} + \frac{(1-\gamma_n)^2\left(2^{2\kappa_n}-1\right)}{\left(\gamma_n + (1-\gamma_n)\,2^{2\kappa_n}\right)^2}\right], \tag{34}$$

subject to

$$\kappa_1 + \kappa_2 \leq \kappa. \tag{35}$$

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:

$$\max_{(\kappa_1,\kappa_2)\in\mathbb{R}_+^2} -\sum_{n=1}^{2} p_n\delta_n\Sigma_{nn}\left[\left(1-2\gamma+\frac{U_{aa}}{U_{a_ia_i}}\right)\frac{1}{\left(\gamma+(1-\gamma)\,2^{2\kappa_n}\right)^2}+\frac{(1-\gamma)^2\left(2^{2\kappa_n}-1\right)}{\left(\gamma+(1-\gamma)\,2^{2\kappa_n}\right)^2}\right], \tag{36}$$

subject to

$$\kappa_1+\kappa_2\leq\kappa. \tag{37}$$

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

Proposition 3 *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 (23), conditions (32)-(33) and $\left(U_{aa}^1/U_{a_ia_i}^1\right)=\left(U_{aa}^2/U_{a_ia_i}^2\right)\equiv\left(U_{aa}/U_{a_ia_i}\right)$ hold. The efficient allocation of attention, denoted κ_1^{eff}, is then given by the solution to problem (36)-(37). Finally, suppose that the constraint (37) is binding and the problem (36)-(37) is convex. Then the following result holds. If $\gamma=\left(U_{aa}/U_{a_ia_i}\right)$ 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 E. ■

Proposition 3 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 $-\left(U_{a_ia}/U_{a_ia_i}\right)$ equals the ratio $\left(U_{aa}/U_{a_ia_i}\right)$, 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 3. 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 4 *Assume that the conditions of Proposition 3 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 F. ■

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

$$\frac{U_{a_i a}}{|U_{a_i a_i}|} + \frac{U_{aa}}{|U_{a_i a_i}|} = 0. \tag{38}$$

Moreover, Proposition 4 states that if the left-hand side of equation (38) 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 (38) 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 (38) 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 (38) is strictly negative, the positive externality dominates, and when the left-hand side of equation (38) is strictly positive, the negative externality dominates.

Condition (38) is equivalent to a condition that has already appeared in the literature in a different context. More precisely, condition (38) is equivalent to the following condition which appears in Angeletos and Pavan (2007):

$$-\frac{U_{a_i a}}{U_{a_i a_i}} = 1 - \left(\frac{U_{a_i a_i}}{U_{a_i a_i}} + 2\frac{U_{a_i a}}{U_{a_i a_i}} + \frac{U_{aa}}{U_{a_i a_i}}\right). \tag{39}$$

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 (39) 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 3 assumes that there is a certain degree of symmetry across regimes. The degree of strategic complementarity $\gamma_n \equiv -\left(U^n_{a_i a}/U^n_{a_i a_i}\right)$ is assumed to be the same across regimes and the ratio $\left(U^n_{aa}/U^n_{a_i a_i}\right)$ 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 (38) holds for each regime, that is, $-\left(U^n_{a_i a}/U^n_{a_i a_i}\right) = \left(U^n_{aa}/U^n_{a_i a_i}\right)$ 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 4, 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.

6 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 $n = 2$,

$$a_{t,n} = \psi_n + \phi_n z_{t,n}, \tag{40}$$

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 (40) to substitute for $a_{t,n}$ in the objective yields

$$\max_{\Lambda} \left\{ -\sum_{n=1}^{2} p_n \delta_n \left(\gamma_n \phi_n + 1 - \gamma_n\right)^2 \Omega_{nn} \right\}, \tag{41}$$

subject to

$$\Omega = \Sigma - \Sigma \left(\Sigma + \Lambda\right)^{-1} \Sigma, \tag{42}$$

$$\frac{1}{2} \log_2 \left(\frac{|\Sigma|}{|\Omega|} \right) \leq \kappa, \tag{43}$$

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 (43) can be expressed as

$$\frac{1}{2} \log_2 \left(\frac{\Sigma_{11}\Sigma_{22} - \Sigma_{12}^2}{\Omega_{11}\Omega_{22} - \Omega_{12}^2} \right) \leq \kappa, \tag{44}$$

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 (44) 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 (44) without improving objective (41). 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 (41)-(43) has the property $\Lambda_{12} = 0$. Next, using $\Sigma_{12} = \Lambda_{12} = \Omega_{12} = 0$ the information choice problem (41)-(43) simplifies to

$$\max_{\left(\Lambda_{11}^{-1}, \Lambda_{22}^{-1}\right) \in \mathbb{R}_+^2} \left\{ -\sum_{n=1}^{2} p_n \delta_n \left(\gamma_n \phi_n + 1 - \gamma_n\right)^2 \Omega_{nn} \right\}, \tag{45}$$

subject to

$$\Omega_{nn} = \frac{1}{\frac{\Sigma_{nn}}{\Lambda_{nn}} + 1}\Sigma_{nn}, \tag{46}$$

and

$$\frac{1}{2}\log_2\left(\frac{\Sigma_{11}}{\Omega_{11}}\right) + \frac{1}{2}\log_2\left(\frac{\Sigma_{22}}{\Omega_{22}}\right) \leq \kappa. \tag{47}$$

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 (45)-(47) can be written as

$$\max_{(\kappa_1,\kappa_2)\in\mathbb{R}^2_+}\left\{-\sum_{n=1}^{2} p_n\delta_n\left(\gamma_n\phi_n + 1 - \gamma_n\right)^2\Omega_{nn}\right\}, \tag{48}$$

subject to

$$\Omega_{nn} = 2^{-2\kappa_n}\Sigma_{nn}, \tag{49}$$

and

$$\kappa_1 + \kappa_2 \leq \kappa. \tag{50}$$

The unique solution to this problem is given by

$$\kappa_1 = \begin{cases} \kappa & \text{if } x \geq 2^{\kappa} \\ \frac{1}{2}\kappa + \frac{1}{2}\log_2(x) & \text{if } x \in \left[2^{-\kappa}, 2^{\kappa}\right] \\ 0 & \text{if } x \leq 2^{-\kappa} \end{cases}, \tag{51}$$

where

$$x \equiv \sqrt{\frac{p_1\delta_1\left(\gamma_1\phi_1 + 1 - \gamma_1\right)^2\Sigma_{11}}{p_2\delta_2\left(\gamma_2\phi_2 + 1 - \gamma_2\right)^2\Sigma_{22}}}; \tag{52}$$

and

$$\kappa_2 = \kappa - \kappa_1. \tag{53}$$

The optimal uncertainty reduction about the fundamental in regime one is an increasing function of κ and x. Finally, it follows from equation (46) 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

$$\Lambda_{11}^{-1} = \frac{2^{2\kappa_1} - 1}{\Sigma_{11}}, \tag{54}$$

$$\Lambda_{22}^{-1} = \frac{2^{2\kappa_2} - 1}{\Sigma_{22}}. \tag{55}$$

Step 4: Equations (51)-(53) 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 (40) 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 (54)-(55) 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 (40) is correct if and only if

$$\psi_n = \frac{\varphi_n}{1-\gamma_n}, \tag{56}$$

$$\phi_n = \frac{(1-\gamma_n)\left(1-2^{-2\kappa_n}\right)}{1-\gamma_n\left(1-2^{-2\kappa_n}\right)}. \tag{57}$$

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 (51)-(53), where ϕ_1 and ϕ_2 are given by equation (57). Using equation (57) to substitute for ϕ_1 and ϕ_2 in equation (52) yields

$$x = \sqrt{\frac{p_1 \delta_1 \Sigma_{11}}{p_2 \delta_2 \Sigma_{22}} \frac{\frac{1-\gamma_1}{1-\gamma_1\left(1-2^{-2\kappa_1}\right)}}{\frac{1-\gamma_2}{1-\gamma_2\left(1-2^{-2\kappa_2}\right)}}}. \tag{58}$$

Thus, an equilibrium allocation of attention is a pair (κ_1, κ_2) satisfying equations (51), (53) and (58). 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 (58) 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 (51) 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

$$\sqrt{\frac{p_1 \delta_1 \Sigma_{11}}{p_2 \delta_2 \Sigma_{22}}} \leq \frac{1}{(1-\gamma_1)\left[2^{\kappa} + \frac{\gamma_2}{1-\gamma_2} 2^{-\kappa}\right]}. \tag{59}$$

Second, consider an equilibrium with the property $\kappa_1 = \kappa$. Substituting $\kappa_1 = \kappa$ and $\kappa_2 = 0$ into equation (58) 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 (51) 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

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

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 (58) 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

$$\left[1 - \sqrt{\frac{p_1 \delta_1 \Sigma_{11}}{p_2 \delta_2 \Sigma_{22}}} \frac{\gamma_2}{1-\gamma_2} 2^{-\kappa}\right] x = \sqrt{\frac{p_1 \delta_1 \Sigma_{11}}{p_2 \delta_2 \Sigma_{22}}} - \frac{\gamma_1}{1-\gamma_1} 2^{-\kappa}. \tag{61}$$

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

$$x = \frac{\sqrt{\frac{p_1 \delta_1 \Sigma_{11}}{p_2 \delta_2 \Sigma_{22}}} - \frac{\gamma_1}{1-\gamma_1} 2^{-\kappa}}{1 - \sqrt{\frac{p_1 \delta_1 \Sigma_{11}}{p_2 \delta_2 \Sigma_{22}}} \frac{\gamma_2}{1-\gamma_2} 2^{-\kappa}}. \tag{62}$$

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 (51) that $\kappa_1=\frac{1}{2}\kappa+\frac{1}{2}\log_2(x)$ is an equilibrium if and only if

$$\frac{\sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}}-\frac{\gamma_1}{1-\gamma_1}2^{-\kappa}}{1-\sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}}\frac{\gamma_2}{1-\gamma_2}2^{-\kappa}}\in\left[2^{-\kappa},2^{\kappa}\right]. \tag{63}$$

Furthermore, when

$$\left[1-\sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}}\frac{\gamma_2}{1-\gamma_2}2^{-\kappa}\right]>0, \tag{64}$$

condition (63) is equivalent to

$$\sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}}\in\left[\frac{\frac{1}{1-\gamma_1}}{2^{\kappa}+\frac{\gamma_2}{1-\gamma_2}2^{-\kappa}},\frac{2^{\kappa}+\frac{\gamma_1}{1-\gamma_1}2^{-\kappa}}{\frac{1}{1-\gamma_2}}\right]. \tag{65}$$

When

$$\left[1-\sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}}\frac{\gamma_2}{1-\gamma_2}2^{-\kappa}\right]<0, \tag{66}$$

condition (63) is equivalent to

$$\sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}}\in\left[\frac{2^{\kappa}+\frac{\gamma_1}{1-\gamma_1}2^{-\kappa}}{\frac{1}{1-\gamma_2}},\frac{\frac{1}{1-\gamma_1}}{2^{\kappa}+\frac{\gamma_2}{1-\gamma_2}2^{-\kappa}}\right]. \tag{67}$$

Finally, if

$$\left[1-\sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}}\frac{\gamma_2}{1-\gamma_2}2^{-\kappa}\right]=0, \tag{68}$$

equation (61) reduces to

$$\sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}}=\frac{\gamma_1}{1-\gamma_1}2^{-\kappa}. \tag{69}$$

In summary, if conditions (64)-(65) or conditions (66)-(67) 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 (62). If conditions (68)-(69) 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 (40). If $\gamma_1=\gamma_2\equiv\gamma$, conditions (59), (60), (64)-(65), (66)-(67) and (68)-(69) and equation (62) reduce to the conditions and equation given in Proposition 1.

B Proof of Lemma 1

Step 1: The expected benefit from limited liability in regime j equals

$$\begin{aligned} B\left(\Omega_{jj}, \bar{u}_j - \omega_j, \delta_j\right) &= \int_{-\infty}^{-\Delta_j} \left[\omega_j - \left(\bar{u}_j - \delta_j z^2\right)\right] f\left(z|s_{i,t-1}\right) dz + \int_{\Delta_j}^{\infty} \left[\omega_j - \left(\bar{u}_j - \delta_j z^2\right)\right] f\left(z|s_{i,t-1}\right) dz \\ &= 2\int_{-\infty}^{-\Delta_j} \left[\omega_j - \left(\bar{u}_j - \delta_j z^2\right)\right] f\left(z|s_{i,t-1}\right) dz. \end{aligned} \tag{70}$$

The second equality is due to the fact that the density has mean zero and is symmetric around its mean.

Step 2: We now compute and sign four derivatives of the expected benefit from limited liability in regime j. The first derivative with respect to Ω_{jj} equals

$$\begin{aligned} \frac{\partial B\left(\Omega_{jj}, \bar{u}_j - \omega_j, \delta_j\right)}{\partial \Omega_{jj}} &= 2\int_{-\infty}^{-\Delta_j} \left[\omega_j - \left(\bar{u}_j - \delta_j z^2\right)\right] \frac{\partial f\left(z|s_{i,t-1}\right)}{\partial \Omega_{jj}} dz \\ &= 2\int_{-\infty}^{-\Delta_j} \left[\omega_j - \left(\bar{u}_j - \delta_j z^2\right)\right] \left[f\left(z|s_{i,t-1}\right)\left(\frac{z^2}{2\Omega_{jj}^2} - \frac{1}{2\Omega_{jj}}\right)\right] dz. \end{aligned} \tag{71}$$

The cross derivative with respect to Ω_{jj} and ω_j equals

$$\frac{\partial^2 B\left(\Omega_{jj}, \bar{u}_j - \omega_j, \delta_j\right)}{\partial \Omega_{jj} \partial \omega_j} = 2\int_{-\infty}^{-\Delta_j} \left[f\left(z|s_{i,t-1}\right)\left(\frac{z^2}{2\Omega_{jj}^2} - \frac{1}{2\Omega_{jj}}\right)\right] dz. \tag{72}$$

The second derivative with respect to Ω_{jj} equals

$$\begin{aligned} \frac{\partial^2 B\left(\Omega_{jj}, \bar{u}_j - \omega_j, \delta_j\right)}{\partial^2 \Omega_{jj}} &= 2\int_{-\infty}^{-\Delta_j} \left[\omega_j - \left(\bar{u}_j - \delta_j z^2\right)\right] \frac{\partial^2 f\left(z|s_{i,t-1}\right)}{\partial^2 \Omega_{jj}} dz \\ &= 2\int_{-\infty}^{-\Delta_j} \left[\omega_j - \left(\bar{u}_j - \delta_j z^2\right)\right] \frac{f\left(z|s_{i,t-1}\right)}{\Omega_{jj}^2} \left[\frac{1}{4}\left(\frac{z^2}{\Omega_{jj}} - 1\right)^2 - \left(\frac{z^2}{\Omega_{jj}} - \frac{1}{2}\right)\right] dz. \end{aligned} \tag{73}$$

The derivative with respect to Ω_{jj}, Ω_{jj}, and ω_j equals

$$\frac{\partial^3 B\left(\Omega_{jj}, \bar{u}_j - \omega_j, \delta_j\right)}{\partial^2 \Omega_{jj} \partial \omega_j} = 2\int_{-\infty}^{-\Delta_j} \frac{f\left(z|s_{i,t-1}\right)}{\Omega_{jj}^2} \left[\frac{1}{4}\left(\frac{z^2}{\Omega_{jj}} - 1\right)^2 - \left(\frac{z^2}{\Omega_{jj}} - \frac{1}{2}\right)\right] dz. \tag{74}$$

Let us start with the cross derivative (72). If $\Delta_j \geq \sqrt{\Omega_{jj}}$, the integral on the right-hand side of equation (72) is strictly positive because the integrand is strictly positive for all $z \in (-\infty, -\Delta_j)$. If $\Delta_j \in \left[0, \sqrt{\Omega_{jj}}\right)$, then

$$\begin{aligned}
\frac{\partial^2 B\left(\Omega_{jj}, \bar{u}_j - \omega_j, \delta_j\right)}{\partial \Omega_{jj} \partial \omega_j} &= 2 \int_{-\infty}^{-\Delta_j} \left[f\left(z | s_{i,t-1}\right) \left(\frac{z^2}{2\Omega_{jj}^2} - \frac{1}{2\Omega_{jj}} \right) \right] dz \\
&\geq 2 \int_{-\infty}^{0} \left[f\left(z | s_{i,t-1}\right) \left(\frac{z^2}{2\Omega_{jj}^2} - \frac{1}{2\Omega_{jj}} \right) \right] dz \\
&= \int_{-\infty}^{\infty} \left[f\left(z | s_{i,t-1}\right) \left(\frac{z^2}{2\Omega_{jj}^2} - \frac{1}{2\Omega_{jj}} \right) \right] dz \\
&= 0.
\end{aligned}$$

The weak inequality is a strict inequality if $\Delta_j > 0$ and the weak inequality is an equality if $\Delta_j = 0$. Collecting results yields

$$\frac{\partial^2 B\left(\Omega_{jj}, \bar{u}_j - \omega_j, \delta_j\right)}{\partial \Omega_{jj} \partial \omega_j} > 0 \text{ for all } \Delta_j > 0, \tag{75}$$

and

$$\frac{\partial^2 B\left(\Omega_{jj}, \bar{u}_j - \omega_j, \delta_j\right)}{\partial \Omega_{jj} \partial \omega_j} = 0 \text{ if } \Delta_j = 0. \tag{76}$$

Let us turn to the first derivative (71). If $\Delta_j \geq \sqrt{\Omega_{jj}}$, the integral on the right-hand side of equation (71) is strictly positive because the integrand is strictly positive for all $z \in (-\infty, -\Delta_j)$. Furthermore, take any value for Δ_j satisfying $\Delta_j \geq \sqrt{\Omega_{jj}}$ and increase ω_j so as to reduce Δ_j to any value $\Delta_j \geq 0$ (recall that $\Delta_j = \sqrt{\frac{\bar{u}_j - \omega_j}{\delta_j}}$). The results about the cross derivative $\frac{\partial^2 B(\Omega_{jj}, \bar{u}_j - \omega_j, \delta_j)}{\partial \Omega_{jj} \partial \omega_j}$ imply that this increase in ω_j raises the first derivative $\frac{\partial B(\Omega_{jj}, \bar{u}_j - \omega_j, \delta_j)}{\partial \Omega_{jj}}$. It follows that

$$\frac{\partial B\left(\Omega_{jj}, \bar{u}_j - \omega_j, \delta_j\right)}{\partial \Omega_{jj}} > 0 \text{ for all } \Delta_j \geq 0.$$

It also follows that, for any Ω_{jj} and δ_j, the derivative $\frac{\partial B(\Omega_{jj}, \bar{u}_j - \omega_j, \delta_j)}{\partial \Omega_{jj}}$ is maximized at $\bar{u}_j - \omega_j = 0$. Furthermore, equation (71) implies

$$\frac{\partial B\left(\Omega_{jj}, 0, \delta_j\right)}{\partial \Omega_{jj}} = \delta_j.$$

Collecting results yields

$$\frac{\partial B\left(\Omega_{jj}, \bar{u}_j - \omega_j, \delta_j\right)}{\partial \Omega_{jj}} \in (0, \delta_j) \text{ for all } \Delta_j > 0, \tag{77}$$

and

$$\frac{\partial B\left(\Omega_{jj},\bar{u}_j-\omega_j,\delta_j\right)}{\partial\Omega_{jj}}=\delta_j \text{ if } \Delta_j=0. \tag{78}$$

Next, consider the second derivative (73). The term

$$\frac{\partial^2 f\left(z|s_{i,t-1}\right)}{\partial^2\Omega_{jj}}=\frac{f\left(z|s_{i,t-1}\right)}{\Omega_{jj}^2}\left[\frac{1}{4}\left(\frac{z^2}{\Omega_{jj}}-1\right)^2-\left(\frac{z^2}{\Omega_{jj}}-\frac{1}{2}\right)\right] \tag{79}$$

has the following properties. The term equals zero for two values of $\left(z^2/\Omega_{jj}\right)$: $3+\sqrt{6}$ and $3-\sqrt{6}$. Furthermore, if $\left(z^2/\Omega_{jj}\right)\notin\left[3-\sqrt{6},3+\sqrt{6}\right]$, the term (79) is strictly positive. If $\left(z^2/\Omega_{jj}\right)\in\left(3-\sqrt{6},3+\sqrt{6}\right)$, the term (79) is strictly negative. We arrive at the following conclusion. If $\Delta_j\geq\sqrt{3+\sqrt{6}}\sqrt{\Omega_{jj}}$, the integral on the right-hand side of equation (73) is strictly positive because the integrand is strictly positive for all $z\in(-\infty,-\Delta_j)$. Thus

$$\frac{\partial^2 B\left(\Omega_{jj},\bar{u}_j-\omega_j,\delta_j\right)}{\partial^2\Omega_{jj}}>0 \text{ for all } \Delta_j\geq\sqrt{3+\sqrt{6}}\sqrt{\Omega_{jj}}. \tag{80}$$

Finally, let us turn to the cross derivative (74). We already showed that the integrand on the right-hand side of equation (74) has the following properties: it equals zero if $\left(z^2/\Omega_{jj}\right)\in\left\{3-\sqrt{6},3+\sqrt{6}\right\}$, it is strictly positive if $\left(z^2/\Omega_{jj}\right)\notin\left[3-\sqrt{6},3+\sqrt{6}\right]$, and it is strictly negative if $\left(z^2/\Omega_{jj}\right)\in\left(3-\sqrt{6},3+\sqrt{6}\right)$. Furthermore, the integral (74) equals zero if $\Delta_j=0$, because the fourth central moment of a normal distribution equals three times the squared variance. We arrive at the following conclusion. There exists a unique threshold value $\bar{\Delta}>0$ with the property

$$\left.\frac{\partial^3 B\left(\Omega_{jj},\bar{u}_j-\omega_j,\delta_j\right)}{\partial^2\Omega_{jj}\partial\omega_j}\right|_{\Delta_j=\bar{\Delta}}=2\int_{-\infty}^{-\bar{\Delta}}\frac{f\left(z|s_{i,t-1}\right)}{\Omega_{jj}^2}\left[\frac{1}{4}\left(\frac{z^2}{\Omega_{jj}}-1\right)^2-\left(\frac{z^2}{\Omega_{jj}}-\frac{1}{2}\right)\right]dz=0. \tag{81}$$

Furthermore, $\bar{\Delta}\in\left(\sqrt{3-\sqrt{6}}\sqrt{\Omega_{jj}},\sqrt{3+\sqrt{6}}\sqrt{\Omega_{jj}}\right)$ and the integral (74) is strictly positive if $\Delta_j>\bar{\Delta}$ while the integral (74) is strictly negative if $\Delta_j\in\left(0,\bar{\Delta}\right)$. This threshold value $\bar{\Delta}$ is linear in $\sqrt{\Omega_{jj}}$ and numerical integration yields that $\bar{\Delta}=1.732\sqrt{\Omega_{jj}}$.

C Proof of Proposition 2

Step 1: The attention choice problem with limited liability reads

$$\max_{\kappa_r\in[0,\kappa]} g\left(\kappa_r,\theta\right), \tag{82}$$

where κ_r is the choice variable, θ is a vector of parameters, and $g(\kappa_r, \theta)$ is the objective function:

$$\begin{aligned} g(\kappa_r, \theta) &= p_r \left[\bar{u}_r - \delta_r \Sigma_{rr} 2^{-2\kappa_r} + B\left(\Sigma_{rr} 2^{-2\kappa_r}, \bar{u}_r - \omega_r, \delta_r\right)\right] \\ &\quad + p_n \left[\bar{u}_n - \delta_n \Sigma_{nn} 2^{-2(\kappa - \kappa_r)} + B\left(\Sigma_{nn} 2^{-2(\kappa - \kappa_r)}, \bar{u}_n - \omega_n, \delta_n\right)\right]. \end{aligned}$$

For comparison, the attention choice problem with unlimited liability reads

$$\max_{\kappa_r \in [0,\kappa]} h(\kappa_r, \theta), \tag{83}$$

where

$$h(\kappa_r, \theta) = p_r \left[\bar{u}_r - \delta_r \Sigma_{rr} 2^{-2\kappa_r}\right] + p_n \left[\bar{u}_n - \delta_n \Sigma_{nn} 2^{-2(\kappa - \kappa_r)}\right].$$

The only difference is that the expected benefit from limited liability in the two regimes (i.e., the term $B\left(\Sigma_{jj} 2^{-2\kappa_j}, \bar{u}_j - \omega_j, \delta_j\right)$ with $j = r, n$) only appears in the first objective function.

Step 2: Let us first study the attention choice problem with unlimited liability. The objective function $h : \mathbb{R} \times \mathbb{R}^{11} \to \mathbb{R}$ is twice continuously differentiable and strictly concave in its first argument. Furthermore, the set $[0, \kappa]$ is compact. Hence, the maximization problem has a unique solution and the solution is given by

$$\begin{aligned} &\kappa_r^{UL} = 0 && \text{if } \left.\frac{\partial h(\kappa_r, \theta)}{\partial \kappa_r}\right|_{\kappa_r = 0} \leq 0, \\ &\kappa_r^{UL} = \kappa && \text{if } \left.\frac{\partial h(\kappa_r, \theta)}{\partial \kappa_r}\right|_{\kappa_r = \kappa} \geq 0, \\ &\left.\frac{\partial h(\kappa_r, \theta)}{\partial \kappa_r}\right|_{\kappa_r = \kappa_r^{UL}} = 0 && \text{otherwise.} \end{aligned}$$

The partial derivative of the objective function with respect to κ_r equals

$$\frac{\partial h(\kappa_r, \theta)}{\partial \kappa_r} = p_r \delta_r \Sigma_{rr} 2^{-2\kappa_r} 2\ln(2) - p_n \delta_n \Sigma_{nn} 2^{-2(\kappa - \kappa_r)} 2\ln(2).$$

Combining results yields

$$\kappa_r^{UL} = \begin{cases} 0 & \text{if } \sqrt{\frac{p_r \delta_r \Sigma_{rr}}{p_n \delta_n \Sigma_{nn}}} \leq 2^{-\kappa} \\ \frac{1}{2}\kappa + \frac{1}{2}\log_2\left(\sqrt{\frac{p_r \delta_r \Sigma_{rr}}{p_n \delta_n \Sigma_{nn}}}\right) & \text{if } \sqrt{\frac{p_r \delta_r \Sigma_{rr}}{p_n \delta_n \Sigma_{nn}}} \in \left(2^{-\kappa}, 2^{\kappa}\right) \\ \kappa & \text{if } \sqrt{\frac{p_r \delta_r \Sigma_{rr}}{p_n \delta_n \Sigma_{nn}}} \geq 2^{\kappa} \end{cases}.$$

Step 3: Let us turn to the attention choice problem with limited liability. The objective function $g : \mathbb{R} \times \mathbb{R}^{11} \to \mathbb{R}$ is twice continuously differentiable in its first argument and the set

$[0, \kappa]$ is compact. Hence, the maximization problem has a solution. The partial derivative of the objective function with respect to κ_r equals

$$\begin{aligned}\frac{\partial g\left(\kappa_r,\theta\right)}{\partial \kappa_r} &= p_r\delta_r\Sigma_{rr}2^{-2\kappa_r}2\ln\left(2\right)\left[1-\frac{\frac{\partial B\left(\Sigma_{rr}2^{-2\kappa_r},\bar{u}_r-\omega_r,\delta_r\right)}{\partial\Sigma_{rr}2^{-2\kappa_r}}}{\delta_r}\right] \\ &\quad -p_n\delta_n\Sigma_{nn}2^{-2(\kappa-\kappa_r)}2\ln\left(2\right)\left[1-\frac{\frac{\partial B\left(\Sigma_{nn}2^{-2(\kappa-\kappa_r)},\bar{u}_n-\omega_n,\delta_n\right)}{\partial\Sigma_{nn}2^{-2(\kappa-\kappa_r)}}}{\delta_n}\right]. \qquad (84)\end{aligned}$$

First, consider the case $\kappa_r^{UL} \in (0, \kappa)$. In this case, we have

$$\begin{array}{ll}\frac{\partial h(\kappa_r,\theta)}{\partial\kappa_r}=0 & \text{if } \kappa_r=\kappa_r^{UL} \\ \frac{\partial h(\kappa_r,\theta)}{\partial\kappa_r}<0 & \text{if } \kappa_r>\kappa_r^{UL}\end{array},$$

which implies

$$\begin{array}{ll}p_r\delta_r\Sigma_{rr}2^{-2\kappa_r}2\ln\left(2\right)=p_n\delta_n\Sigma_{nn}2^{-2(\kappa-\kappa_r)}2\ln\left(2\right) & \text{if } \kappa_r=\kappa_r^{UL} \\ p_r\delta_r\Sigma_{rr}2^{-2\kappa_r}2\ln\left(2\right)<p_n\delta_n\Sigma_{nn}2^{-2(\kappa-\kappa_r)}2\ln\left(2\right) & \text{if } \kappa_r>\kappa_r^{UL}\end{array}. \qquad (85)$$

Furthermore, let $\kappa_r^{equality} \in \mathbb{R}_+$ denote the attention allocation at which the posterior uncertainty about the optimal action in the rare event equals the posterior uncertainty about the optimal action in normal times, that is,

$$\Sigma_{rr}2^{-2\kappa_r^{equality}}=\Sigma_{nn}2^{-2\left(\kappa-\kappa_r^{equality}\right)},$$

or equivalently

$$\kappa_r^{equality}=\frac{1}{2}\left[\kappa+\log_2\left(\sqrt{\frac{\Sigma_{rr}}{\Sigma_{nn}}}\right)\right].$$

The assumption $\delta_r = \delta_n$ implies $\kappa_r^{UL} < \kappa_r^{equality}$. The assumptions $\bar{u}_r - \omega_r = \bar{u}_n - \omega_n$, $\delta_r = \delta_n$, and $\Delta_r \geq 1.732\sqrt{\Omega_{rr}^{UL}}$ imply that

$$\begin{array}{lll}\frac{\frac{\partial B\left(\Sigma_{rr}2^{-2\kappa_r},\bar{u}_r-\omega_r,\delta_r\right)}{\partial\Sigma_{rr}2^{-2\kappa_r}}}{\delta_r} > \frac{\frac{\partial B\left(\Sigma_{nn}2^{-2(\kappa-\kappa_r)},\bar{u}_n-\omega_n,\delta_n\right)}{\partial\Sigma_{nn}2^{-2(\kappa-\kappa_r)}}}{\delta_n} & \text{if } \kappa_r\in\left[\kappa_r^{UL},\kappa_r^{equality}\right) \\ \frac{\frac{\partial B\left(\Sigma_{rr}2^{-2\kappa_r},\bar{u}_r-\omega_r,\delta_r\right)}{\partial\Sigma_{rr}2^{-2\kappa_r}}}{\delta_r} = \frac{\frac{\partial B\left(\Sigma_{nn}2^{-2(\kappa-\kappa_r)},\bar{u}_n-\omega_n,\delta_n\right)}{\partial\Sigma_{nn}2^{-2(\kappa-\kappa_r)}}}{\delta_n} & \text{if } \kappa_r=\kappa_r^{equality}\end{array}. \qquad (86)$$

Namely, the assumption $\Delta_r \geq 1.732\sqrt{\Omega_{rr}^{UL}}$ implies that the function B is strictly convex in its first argument on $\left(0, \Omega_{rr}^{UL}\right]$, and we have $\Sigma_{nn}2^{-2(\kappa-\kappa_r)} \leq \Sigma_{rr}2^{-2\kappa_r} \leq \Omega_{rr}^{UL}$ for all $\kappa_r \in \left[\kappa_r^{UL}, \kappa_r^{equality}\right]$. Combining results (84)-(86) yields

$$\forall\kappa_r\in\left[\kappa_r^{UL},\kappa_r^{equality}\right]:\frac{\partial g\left(\kappa_r,\theta\right)}{\partial\kappa_r}<0. \qquad (87)$$

Hence, any $\kappa_r \in \left[\kappa_r^{UL}, \kappa_r^{equality}\right]$ cannot be a solution to the attention choice problem (82). Next, we show that any $\kappa_r > \kappa_r^{equality}$ cannot be a solution to the attention choice problem with limited liability. If $\kappa_r^{equality} \geq \kappa$, this result follows from the fact that κ_r cannot exceed κ. If $\kappa_r^{equality} < \kappa$, this result follows from the following argument. Let Ω denote the posterior uncertainty about the optimal action in the two regimes at $\kappa_r = \kappa_r^{equality}$

$$\Omega = \Sigma_{rr} 2^{-2\kappa_r^{equality}} = \Sigma_{nn} 2^{-2\left(\kappa - \kappa_r^{equality}\right)}.$$

One can express the posterior uncertainty in the two regimes as

$$\Omega_{rr} = \Omega 2^{-2\left(\kappa_r - \kappa_r^{equality}\right)},$$

and

$$\Omega_{nn} = \Omega 2^{2\left(\kappa_r - \kappa_r^{equality}\right)}.$$

For all $\kappa_r > \kappa_r^{equality}$, we have $\Omega_{rr} < \Omega_{nn}$ and one can swap the value of Ω_{rr} and the value of Ω_{nn} by changing the sign of $\kappa_r - \kappa_r^{equality}$. Changing the sign of $\kappa_r - \kappa_r^{equality}$ without violating $\kappa_r \in [0, \kappa]$ is always feasible because the assumption $\Sigma_{rr} \geq \Sigma_{nn}$ implies $\kappa_r^{equality} \geq \frac{1}{2}\kappa$. Furthermore, the objective function under limited liability can be written as

$$\begin{aligned} g(\kappa_r, \theta) \quad = \quad & p_r \bar{u}_r - p_r \left[\delta_r \Omega_{rr} - B\left(\Omega_{rr}, \bar{u}_r - \omega_r, \delta_r\right)\right] \\ & + p_n \bar{u}_n - p_n \left[\delta_n \Omega_{nn} - B\left(\Omega_{nn}, \bar{u}_n - \omega_n, \delta_n\right)\right]. \end{aligned}$$

The first square bracket on the right-hand side is the expected loss due to suboptimal action in the rare event. This expected loss is strictly positive and strictly increasing in Ω_{rr}. See Lemma 1. The second square bracket on the right-hand side is the expected loss due to suboptimal action in normal times. This expected loss is strictly positive and strictly increasing in Ω_{nn}. Recall that for all $\kappa_r > \kappa_r^{equality}$ we have $\Omega_{rr} < \Omega_{nn}$ and swaping the values of Ω_{rr} and Ω_{nn} is feasible. Note that swaping yields a higher value of the objective because $p_r < p_n$, $\bar{u}_r - \omega_r = \bar{u}_n - \omega_n$, and $\delta_r = \delta_n$. Hence, any $\kappa_r > \kappa_r^{equality}$ cannot be a solution to the attention choice problem (82). Combining results we arrive at the conclusion stated in Proposition 2: If $\kappa_r^{UL} \in (0, \kappa)$, every solution to the attention choice problem with limited liability satisfies $\kappa_r < \kappa_r^{UL}$.

Second, consider the case $\kappa_r^{UL} = 0$. In this case, the unique solution to the attention choice problem with limited liability is $\kappa_r = 0$. The arguments are almost identical to the arguments in the case of $\kappa_r^{UL} \in (0, \kappa)$. There are two differences. The first difference is that result (85) is replaced by

$$\begin{array}{ll} p_r \delta_r \Sigma_{rr} 2^{-2\kappa_r} 2\ln(2) \leq p_n \delta_n \Sigma_{nn} 2^{-2(\kappa-\kappa_r)} 2\ln(2) & \text{if } \kappa_r = \kappa_r^{UL} \\ p_r \delta_r \Sigma_{rr} 2^{-2\kappa_r} 2\ln(2) < p_n \delta_n \Sigma_{nn} 2^{-2(\kappa-\kappa_r)} 2\ln(2) & \text{if } \kappa_r > \kappa_r^{UL} \end{array}.$$

The second difference is that result (87) and the result that any $\kappa_r > \kappa_r^{equality}$ cannot be a solution to the attention choice problem with limited liability now imply that the unique solution to the attention choice problem with limited liability is $\kappa_r = 0$.

Third, consider the case $\kappa_r^{UL} = \kappa$. In this case, the fact that $\kappa_r \in [0, \kappa]$ implies that every solution to the attention choice problem with limited liability satisfies $\kappa_r \leq \kappa_r^{UL}$.

D Proof of Equation (25)

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

$$\begin{aligned} U^n(a_{i,t,n}, a_{t,n}, z_{t,n}) &= U^n(a_{t,n}, a_{t,n}, z_{t,n}) + \left[U^n_{a_i} + \left(U^n_{a_i a_i} + U^n_{a_i a}\right) a_{t,n} + U^n_{a_i z} z_{t,n}\right](a_{i,t,n} - a_{t,n}) \\ &\quad + \frac{U^n_{a_i a_i}}{2}(a_{i,t,n} - a_{t,n})^2. \end{aligned} \tag{88}$$

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

$$W^n\left(a_{t,n}, \sigma_{a_{i,t,n}}, z_{t,n}\right) \equiv \int U^n(a_{i,t,n}, a_{t,n}, z_{t,n})\, d\Psi^{n,t}(a_{i,t,n}). \tag{89}$$

Combining the last two equations gives

$$W^n\left(a_{t,n}, \sigma_{a_{i,t,n}}, z_{t,n}\right) = U^n(a_{t,n}, a_{t,n}, z_{t,n}) + \frac{U^n_{a_i a_i}}{2}\sigma^2_{a_{i,t,n}}, \tag{90}$$

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 (24), yields

$$W^n\left(a_{t,n}, \sigma_{a_{i,t,n}}, z_{t,n}\right) = W^n\left(a^*_{t,n}, 0, z_{t,n}\right) + \frac{U^n_{a_i a_i} + 2U^n_{a_i a} + U^n_{aa}}{2}\left(a_{t,n} - a^*_{t,n}\right)^2 + \frac{U^n_{a_i a_i}}{2}\sigma^2_{a_{i,t,n}}. \tag{91}$$

Here we used equation (90) 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

$$E\left[U^n(a_{i,t,n}, a_{t,n}, z_{t,n})\right] = \int_{z_t}\int_{s_{i,t-1}} U^n\left(a_{i,t,n}(s_{i,t-1}), a_{t,n}(z_t), z_{t,n}\right) dP(s_{i,t-1}|z_t)\, dP(z_t), \tag{92}$$

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 (88) into equation (92) and using equation (90) gives

$$E\left[U^n\left(a_{i,t,n}, a_{t,n}, z_{t,n}\right)\right] = \int_{z_t} W^n\left(a_{t,n}(z_t), \sigma_{a_{i,t,n}}, z_{t,n}\right) dP(z_t). \tag{93}$$

Substituting equation (91) into the last equation yields

$$\begin{aligned} E\left[U^n\left(a_{i,t,n}, a_{t,n}, z_{t,n}\right)\right] &= E\left[W^n\left(a^*_{t,n}, 0, z_{t,n}\right)\right] + \frac{U^n_{a_i a_i} + 2U^n_{a_i a} + U^n_{aa}}{2} E\left[\left(a_{t,n} - a^*_{t,n}\right)^2\right] \\ &\quad + \frac{U^n_{a_i a_i}}{2} E\left[\left(a_{i,t,n} - a_{t,n}\right)^2\right]. \end{aligned} \tag{94}$$

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.

E Proof of Proposition 3

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 (36) and setting the first derivative of the objective with respect to κ_1 equal to zero yields the first-order condition

$$\begin{aligned} &p_1\delta_1\Sigma_{11}\left[\tfrac{(1-\gamma)^2 2^{2\kappa_1}}{\left[\gamma+(1-\gamma)2^{2\kappa_1}\right]^2} + 2\tfrac{(1-\gamma)2^{2\kappa_1}}{\left[\gamma+(1-\gamma)2^{2\kappa_1}\right]^3}\left(\tfrac{U_{aa}}{U_{a_i a_i}} - \gamma\right)\right] 2\ln(2) \\ &-p_2\delta_2\Sigma_{22}\left[\tfrac{(1-\gamma)^2 2^{2(\kappa-\kappa_1)}}{\left[\gamma+(1-\gamma)2^{2(\kappa-\kappa_1)}\right]^2} + 2\tfrac{(1-\gamma)2^{2(\kappa-\kappa_1)}}{\left[\gamma+(1-\gamma)2^{2(\kappa-\kappa_1)}\right]^3}\left(\tfrac{U_{aa}}{U_{a_i a_i}} - \gamma\right)\right] 2\ln(2) = 0. \end{aligned} \tag{95}$$

Let $F_{\kappa_1=0}$ and $F_{\kappa_1=\kappa}$ denote the value of the left-hand side of equation (95) at $\kappa_1 = 0$ and $\kappa_1 = \kappa$, respectively. When the constraint (37) is binding and the planner problem (36)-(37) is convex, the solution to the planner problem is given by

$$\kappa_1^{eff} = \begin{cases} \kappa & \text{if } F_{\kappa_1=\kappa} \geq 0 \\ \kappa_1^{FOC} & \text{if } F_{\kappa_1=0} > 0 > F_{\kappa_1=\kappa} \\ 0 & \text{if } F_{\kappa_1=0} \leq 0 \end{cases}, \tag{96}$$

where κ_1^{FOC} denotes the unique solution to equation (95) in the case of $F_{\kappa_1=0} > 0 > F_{\kappa_1=\kappa}$.

Step 3: If $\gamma = (U_{aa}/U_{a_i a_i})$, the first-order condition (95) reduces to

$$p_1\delta_1\Sigma_{11}\frac{(1-\gamma)^2 2^{2\kappa_1}}{\left[\gamma + (1-\gamma)2^{2\kappa_1}\right]^2} 2\ln(2) - p_2\delta_2\Sigma_{22}\frac{(1-\gamma)^2 2^{2(\kappa-\kappa_1)}}{\left[\gamma + (1-\gamma)2^{2(\kappa-\kappa_1)}\right]^2} 2\ln(2) = 0. \tag{97}$$

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 (97) 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

$$\kappa_1^{eff} = \begin{cases} \kappa & \text{if } \sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}} \geq \gamma 2^{-\kappa} + (1-\gamma)\, 2^{\kappa} \\ \frac{1}{2}\kappa + \frac{1}{2}\log_2\left(\frac{\sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}} - \frac{\gamma}{1-\gamma}2^{-\kappa}}{1-\sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}}\frac{\gamma}{1-\gamma}2^{-\kappa}}\right) & \text{otherwise} \\ 0 & \text{if } \sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}} \leq \frac{1}{\gamma 2^{-\kappa}+(1-\gamma)2^{\kappa}} \end{cases}. \tag{98}$$

Comparing equation (98) 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 (95) reduces to

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

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 $\kappa_1^{equ} = \kappa_1^{eff}$.

F Proof of Proposition 4

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

$$\kappa_1^{equ} = \frac{1}{2}\kappa + \frac{1}{2}\log_2(x), \tag{99}$$

where

$$x \equiv \frac{\sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}} - \frac{\gamma}{1-\gamma}2^{-\kappa}}{1-\sqrt{\frac{p_1\delta_1\Sigma_{11}}{p_2\delta_2\Sigma_{22}}}\frac{\gamma}{1-\gamma}2^{-\kappa}}, \tag{100}$$

and

$$x \in \left(2^{-\kappa}, 2^{\kappa}\right). \tag{101}$$

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 (95) at $\kappa_1 = \kappa_1^{equ} \in (0,\kappa)$. Substituting equation (99) into the left-hand side of equation (95) 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 (100) 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

$$\begin{aligned} F_{\kappa_1=\kappa_1^{equ}\in(0,\kappa)} &= p_1\delta_1\Sigma_{11}\frac{(1-\gamma)2^{\kappa}x}{\left[\gamma+(1-\gamma)2^{\kappa}x\right]^2} \\ &\quad \left[\frac{2}{\gamma+(1-\gamma)2^{\kappa}x}-\frac{2}{\gamma+(1-\gamma)\frac{2^{\kappa}}{x}}\right]\left(\frac{U_{aa}}{U_{a_ia_i}}-\gamma\right)2\ln(2). \end{aligned} \tag{102}$$

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 (99). 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] Alvarez, Fernando and Gadi Barlevy (2013): "Mandatory Disclosure and Contagion for Interconnected Banks." Discussion paper, University of Chicago and Federal Reserve Bank of Chicago.

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

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

[4] 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.

[5] Bastasin, Carlo (2012): "Saving Europe: How National Politics Nearly Destroyed the Euro." Brookings Institution Press, Washington, D.C.

[6] Caballero, Ricardo J. and Alp Simsek (2012): "Fire Sales in a Model of Complexity." Discussion paper, Massachusetts Institute of Technology and Harvard University.

[7] 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.

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

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

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

[11] International Monetary Fund (2013): "Greece: Ex Post Evaluation of Exceptional Access under the 2010 Stand-By Agreement." International Monetary Fund Country Report 13/156.

[12] Irwin, Neil (2013): "The Alchemists: Three Central Bankers and the World on Fire." Penguin Press, New York.

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

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

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

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

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

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

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

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

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

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

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

[24] 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.

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

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

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

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

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

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

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

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

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

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

Figure 1: Equilibrium attention to the rare event

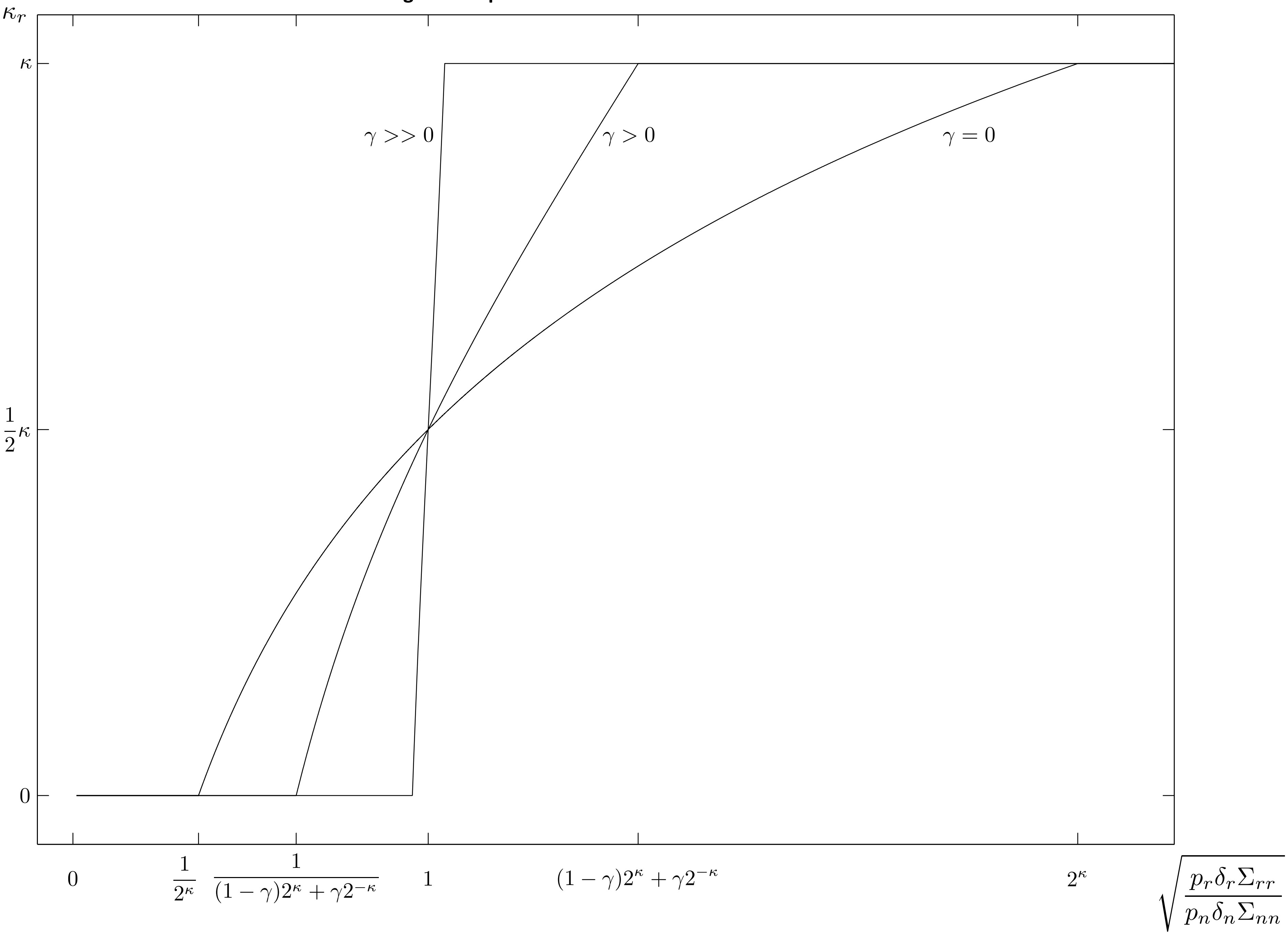
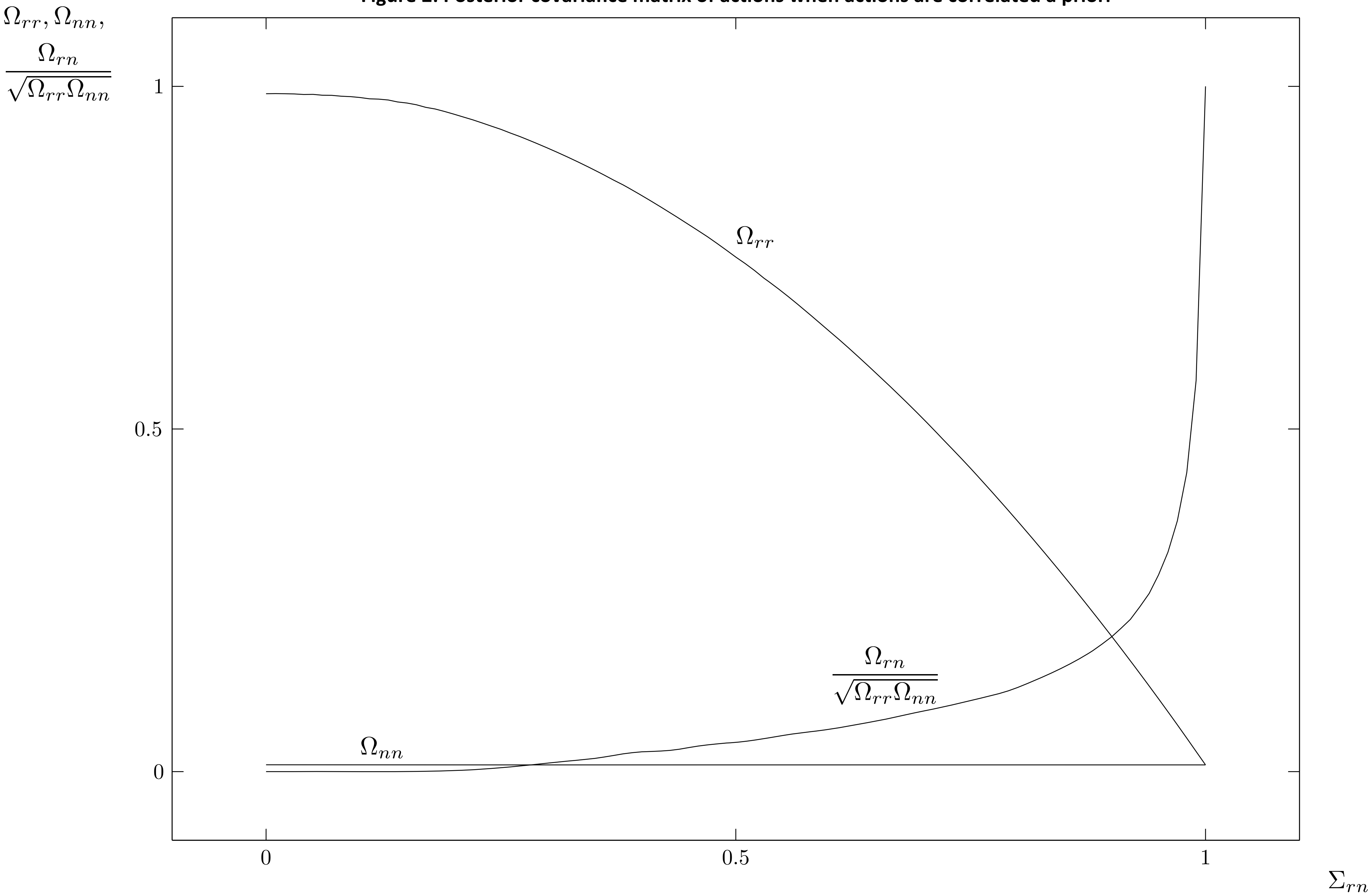


Figure 2: Posterior covariance matrix of actions when actions are correlated a priori



Note: This figure assumes $\Sigma_{rr} = \Sigma_{nn} = 1$ so that Σ_{rn} is the prior correlation of actions.

Figure 3: Posterior expectation of the probability of unusual times

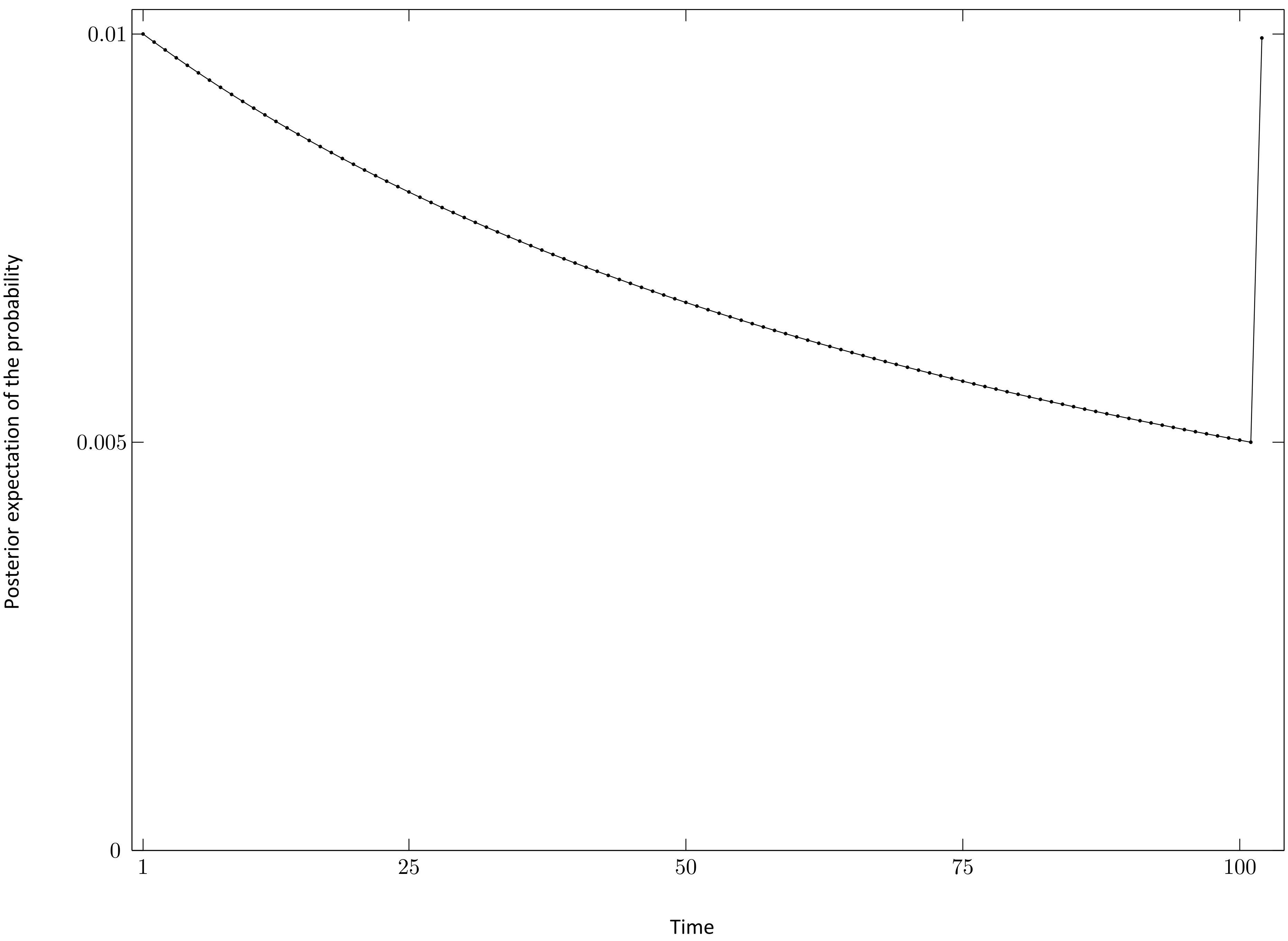


Figure 4: Probability of default, Lehman Brothers

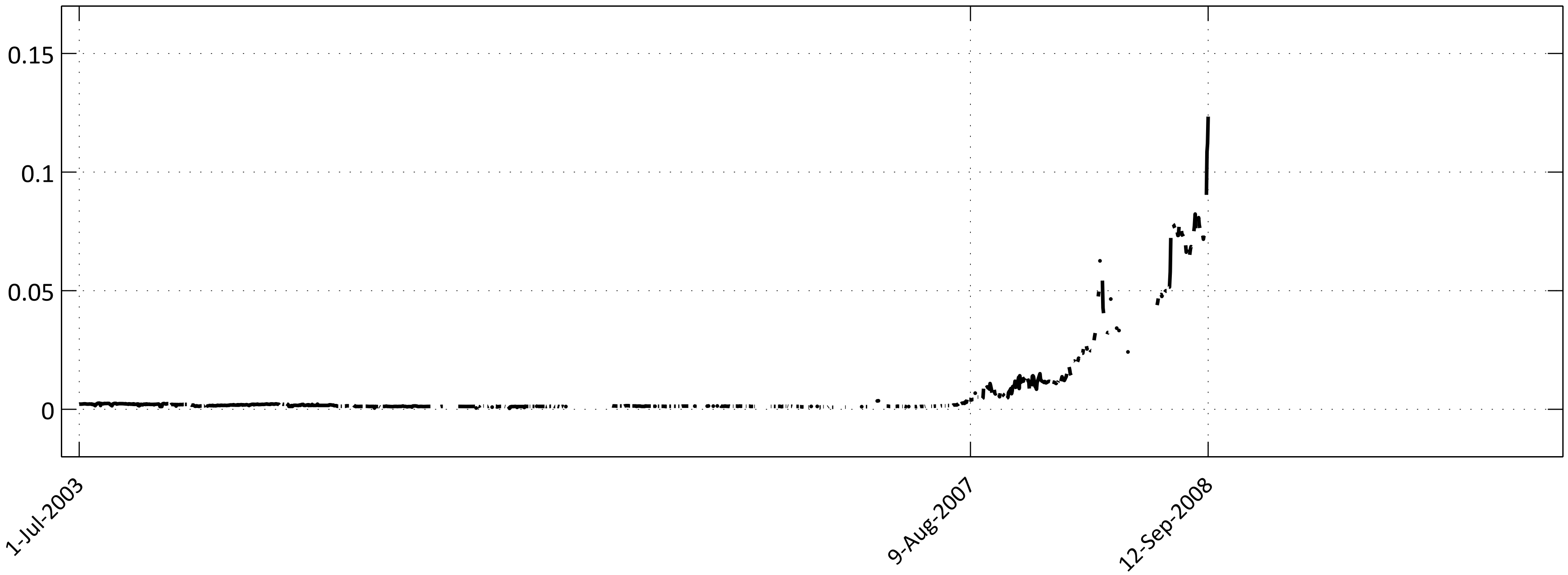
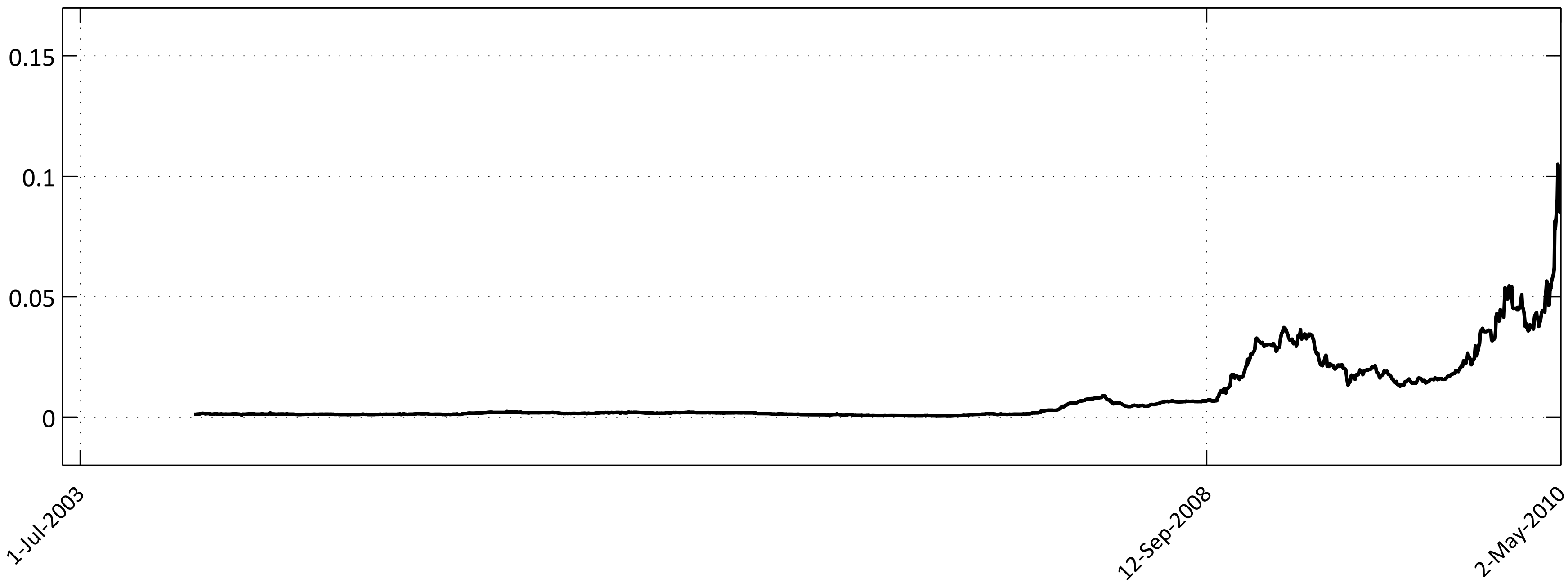


Figure 5: Probability of default, Greece



Note: The probabilities of default in Figures 4-5 are derived from CDS premia. See Section 4.4 for the details.