

The Response of Catastrophe Insurance Markets to Extreme Events: A Real Option Approach

Alex Boulatov
boulatov@haas.berkeley.edu

Dwight Jaffee
jaffee@haas.berkeley.edu

January 27, 2003

Abstract

This paper connects the traditional financial and insurance literatures in the context of real option theory. We use the analogy between insurance and investment under uncertainty in order to study the general equilibrium in an insurance industry with heterogeneous competing firms. We show that even under the assumption of rationality (defined in the traditional sense) insurance companies should be treated as strategic agents when catastrophic (CAT) events are possible. We derive the equilibrium investment (insurance policies) using a generalization of the standard model of partially reversible investment under uncertainty. Implications of the proposed model for the optimal contract design and future extensions in the framework of dynamic programming approach are discussed.

Haas School of Business, UC Berkeley, Berkeley, CA 94720-1900

We are grateful to Thomas Russell for the fruitful discussions.

1 Introduction

Following each of the last three major US catastrophic events, the insurance markets for these risks responded in distinctive and remarkably similar ways. The three catastrophes were the 1992 Andrew hurricane, the 1994 Northridge earthquake, and the 2001 9/11 terrorist attack. Immediately following each event:

1) Virtually all firms announced they had sufficient capital to pay all expected claims.

2) Many firms then announced they would not renew policies, and left the industry. The result was a disrupted market with rationing, negative impacts on the real estate markets for new construction, real estate finance, and existing property sales.

3) For reinsurance firms and primary insurance firms that remained in the market, premiums were raised dramatically (400% or more).

4) After some time, government entities intervened to stabilize the market; for the natural disasters, state governments acted; for the terrorist event, the federal government acted.

5) As further time passed and subsequent events did not occur, the effects in (2) and (3) appear to be reversed. For the natural disasters, premiums are now even approaching the levels in place prior to the events, and firms are re-entering the industry. For the terrorist attack, 14 months after the event and before the government intervention, we were seeing some modest declines in premiums and the re-entry of some firms into that insurance market.

This stylized description of how catastrophe insurance markets respond to an extreme event indicates distinctive dynamics for how both premiums and market participation respond after the event. On the one hand, prices tend to rise sharply, then to decline slowly as no further event occurs. On the other hand, firms tend to leave the market immediately, then slowly return as no further events occur. In short, we see firms leave the industry when the prices are high, then reenter the industry when the prices are falling.

There are a variety of theories that may explain part, if not all, of the price dynamics. Given that all three events have created unexpectedly large losses, standard Bayesian updating of expected losses could explain at least some of the initial rise in premiums. Bayesian updating, however, provides a less obvious or persuasive explanation of the subsequent decline in premiums. Ambiguity aversion, in which uncertainty about the generating process of the

risk creates an additional risk premium component in prices, can provide a further explanation of the initial jump in premiums, and may even explain the subsequent decline (if the lack of further events reduces the ambiguity aversion).

Insurance premium cycle theories provide an alternative explanation. The idea here is that firms use the large profits derived from the high premiums to replenish their capital following the losses. As the capital base of the industry is restored, premiums fall to lower levels. A key issue for the premium cycle explanation, however, is why the firms do not just directly access the capital markets to replenish their capital. This would seem natural since following the bad event, premia will be at very high levels.

Whatever the efficacy of these explanations for the price dynamics, they do not explain the dynamics of firm exit and later reentry. First, both Bayesian updating and/or ambiguity aversion indicate that premiums can and do respond to the new information concerning risk and uncertainty, leaving no reason for firms to exit the industry. For the premium cycle theories, moreover, firms must remain in the industry in order to earn the exceptional profit margins and thus replenish their capital.

The focus of this paper is to provide quite a different model of price dynamics and of firm exit and reentry, by applying the concepts of real option theory to catastrophe insurance markets.

Traditionally, real option theory is applied to the problem of a firm choosing when to carry out a capital investment expenditure. The well-known result is that firms may delay carrying out investments, even when the project has a positive net present value, in order to gain the benefit of further information that arrives only with time. Recent extensions of the theory, moreover, indicate that the extent of the delay may depend on the degree of competition in the market for the investment project. In particular, the more competitive this market, the less firms will delay the investment, since all firms will fear that some other firm will preempt the market by carrying out the investment sooner rather than later. Thus all firms accelerate their investment.

The decision of an insurance firm to take on catastrophe risks can be considered to be an investment decision. The fact that the insurance firms receive premiums as a cash in flow at date 0, and then face possible payouts on claims at a later date 1—just the opposite of the cash flow pattern of a normal investment project—is not a fundamental difference. Similarly, the arrival of

new information on the distribution of possible insurance losses from a future event corresponds to the arrival of new investment cost information in the traditional real options investment setup. The upshot is that the decisions of a catastrophe insurance firm can readily fit into the traditional real options model. Catastrophe insurance investments, however, are readily reversible in the sense that an insurance firm can allow its current policies to mature, and then leave the industry. Thus, to apply the real option approach to catastrophe insurance, we will have to amend the basic theory to cover the case of reversible investments.

The key benefit of applying real option theory to catastrophe insurance markets is that we obtain a unified theory of price dynamics and exit and entry decisions that fits the stylized facts of catastrophe insurance markets remarkably well.

2 Two-period Model

In order to illustrate the analogy between the optimal insurance and oligopolistic investment strategies, we will first consider a simple two-period model of the insurance market. Following the approach suggested in [7], we assume that there are N risk-neutral insurance companies insuring comparable assets against possible losses due to the random catastrophe (CAT) event. Therefore, each company is facing the same risk, represented as a random loss process. Each company receives the insurance premium from its subscribers in the first period and can borrow/lend money in the financial markets at the interest rate r_f in order to maintain its optimal insurance portfolio. For the clarity of exposition, we will assume in this section that $r_f = 0$. This assumption will be lifted in a general continuous time model considered in next section. In a two-period model, we assume that the CAT event may take place at date 2 after the insurance contract is sold at date 1. At date 2, the subscribers will be compensated by the insurer for all losses, if any. The companies are maximizing their expected wealth, and the loss distribution is assumed to be normal (this assumption will also be lifted in the next section).

Of course, real options play no role in this case, since the investment *must* be made at date 1. Nevertheless, this model provides a useful benchmark for the explicitly dynamic model presented later.

The wealth of the i th insurance company ($i = 1, \dots, N$), W_i , is given by

the budget condition

$$W_i = W_{i0} + q_i R, \quad (1)$$

where W_{i0} is the initial wealth of the insurance company, i , q_i ($i = 1, \dots, N$) are the quantities of contracts for each firm. The “returns” are given by

$$R = p - l, \quad (2)$$

with p being the insurance premia and l is the loss. The expected wealth for the i -th insurer is given by

$$\Phi_i = W_{i0} + q_i E[R], \quad (3)$$

where $i = 1, \dots, N$, q is the vector of allocations and R is the “returns”

$$E[R] = p - E[l]. \quad (4)$$

In order to obtain an equilibrium solution, we have to specify the problem on the subscriber’s side. To avoid technical complications and focus on the impact of the oligopolistic competition we take as given the inverse demand function

$$p = K Q^{-1/\gamma}, \quad (5)$$

with $\{K, \gamma\}$ being parameters of the demand in each of the regions (assumed to be exogenously given) and $Q = \sum_{i=1}^N q_i$, the total number of contracts (aggregate demand) ¹. Making use of (3), we arrive at the optimization problem

$$\max_{q_i} \{L(q_i)\},$$

with

$$L(q_i) = q_i E[R].$$

We will consider several types of equilibrium models: Cournot, monopoly and Walras equilibrium. They correspond to different competition settings: oligopoly, monopoly and perfect competition, respectively.

¹This specification divorces the demand function from dependence on the same parameters of the distribution of losses that are used by the insurance firms. That is, demand may be affected by the loss probability but it not necessarily the same probability as the one used by the firms. The possibility of extending our model to include this form of dependence is discussed in the concluding section.

2.1 Oligopoly

We model oligopoly in the framework of Cournot model. With the above assumptions, we can look for the symmetric equilibrium between the oligopolistic insurance companies with q_i , $\forall i = 1, \dots, N$. This symmetric solution is given by ²

$$q^{(N)} = \left[\frac{K}{E[l]} \left(\frac{\gamma N - 1}{\gamma N} \right) \right]^\gamma, \quad (7)$$

and price,

$$p^{(N)} = E[l] \left(\frac{\gamma N}{\gamma N - 1} \right) = E[l] \left(1 + \frac{1}{\gamma N - 1} \right). \quad (8)$$

From the solution (8), it follows that the equilibrium price breaks into two components, the risk-neutral (expected loss) part and what we will (loosely) call here the risk premium part. Note that the risk premium does not vanish despite the fact that all insurers are risk-neutral. This is because the Cournot competition is imperfect for any finite number of oligopolists. However, when N grows large pricing becomes risk free

$$\lim_{N \rightarrow \infty} p^{(N)} = E[l]. \quad (9)$$

Note, that for all practical purposes it is safe to assume in the CAT market that N is finite and usually $N \ll \infty$. This parallels the comments in [7] that the failure of insurance policies to be priced at the expected losses must be attributed in some dimension to a failure in subdividing the risk among a sufficiently large number of investors. To conclude, even the simple model formulated above shows that for the case of CAT insurance there will be a non-negligible risk-premium component in price.

²The symmetric solution is obtained as a symmetric solution of the optimization problem

$$q = \arg \max_{q_i} \left\{ q_i \left[K \left(q_i + \sum_{j \neq i}^N q_j \right)^{-1/\gamma} - E[l] \right] \right\}. \quad (6)$$

2.2 Monopoly

In this case all subscribers deal with the same insurance company and we therefore have $Q = q$.

$$\max_q P(q)q - E[l]q.$$

The solution is very similar to one derived in the previous subsection for oligopoly. The only correction one needs to make is to set N to 1. We can still use (7) and (8).

2.3 Competitive equilibrium

We derive the equilibrium of this subsection in the framework of Walrasian equilibrium. Let there be a large number of competing insurance companies, N . In this case, the aggregate demand is given by $Q = \sum_{i=1}^N q_i$, $i = 1, \dots, N$. Substituting this into (3), optimizing with respect to $\{q_i\}$ and using the market clearing condition, we obtain a set of FOC equations. Using the same assumptions as for the Cournot equilibrium, we obtain solution for the symmetric competitive equilibrium in the form

$$x^{(N)} = \left(\frac{K}{E[l]} \right)^\gamma, \quad (10)$$

and

$$p^{(N)} = E[l], \quad (11)$$

which is very similar to (7) and (8). From (11), it follows that in this case the risk premium vanishes as opposed to the imperfect competition described by the Cournot model. This is not surprising and is in agreement with (9).

2.4 Optionality in Insurance

The main purpose of the above discussion was to show an analogy between insurance and investment decisions, which will enable us to apply the real options approach to the strategic insurance problem.

These considerations can be further clarified by a simple two-state, two period, example. At date 0, the insurance firm can sell a policy for a premium of 2,200, which commits the firm to pay the annual catastrophe losses of the policy holder forever (the insurance firm cannot disinvest from the policy

once it is sold). The catastrophe loss at date 1 will be either 400 (bad event) or 0 (good event), each occurring with a $\frac{1}{2}$ probability. Whichever event occurs at date 1, this particular event will then be repeated annually forever. The risk-free interest rate is 10%, and we shall assume that the risks are fully diversifiable, so this rate can be used to discount the expected future losses.

The firm can sell the insurance policy as of date 0, in which case it is committed to paying the annual losses forever starting at date 1. The expected net present value (NPV) as of date 0 of this investment is readily computed as:

$NPV = +2,200 - [(.5)(0) + (.5)(400)] / (.10) = 2,200 - 2,000 = 200 > 0$. Thus, if the insurance firm's only choice were to sell the insurance today or never to sell it at all, then it would of course be willing to sell the insurance today.

Real option theory, however, offers the insurance firm the opportunity to wait until further information becomes available at date 1, then to decide whether or not to offer the policy. If the insurance firm observes the bad state at date 1 (the loss of 400), it will then bypass the opportunity to sell the policy at all (and it will be grateful it did not sell the policy at date 0). On the other hand, if the firm observes the good state at date 1 (0 loss), then it would compute the expected NPV (as of date 0) of selling the policy at date 1 as:

$$NPV = (.5)[+2,200 - (0) / (.10)] = (.5)(2,000) = 1,000 > 200.$$

That is, deferring the decision till date 1, then selling the policy only if the good state is revealed, creates a net present value that exceeds the (positive) net present value available as of date 0. Thus, in these circumstances the insurance firm would always defer its decision till the information at date 1 was revealed.

Although this example is highly simplified, it illustrates the basic sense in which real option theory may provide a motivation for an insurance firm to defer selling policies pending further information³. For a more general and realistic model, we will suppose that the state variable is presented by a two-state process, where the state may be either "bad" (losses occur) or

³Readers familiar with Dixit and Pindyck [2], pp. 27-28 will also recognize that our example has the same basic structure as their 2-period, 2-state, example of option theory applied to investment decisions. We intentionally set it up this way to illustrate the direct correspondence between insurance and investment applications of real option theory. Dixit and Pindyck [2] also discuss some of the special issues of investment cost uncertainty.

"good" (no losses). Assuming that the transition probabilities between the two states are known, we obtain a complete characterization of the firm's information set. In order to make an analogy with the case of a continuous state variable which is less likely to undergo large jumps, we assume that the transition probability matrix has the form

$$T = \begin{bmatrix} \chi & 1 - \chi \\ 1 - \chi & \chi \end{bmatrix}, \quad (12)$$

with $\chi > 1/2$ implying that the probability of "jump" is less than the probability for the two subsequent states to be of the same type. In other words, after the "good" event it is more likely to observe another "good" event than to switch to the "bad" event on next date, and the same about two subsequent "bad" events. Note that the example considered above can be viewed as a particular case of the two-state jump process with $\chi = 0.5$ at date 0 and $\chi = 1$ on date 1 and later. In a more general case, the insurance company (investor) optimizes by increasing capacity after observing a "good" event and by decreasing it after a "bad" event at date 1, respectively. Therefore, the insurance decisions will depend on a realization of a state variable, which represents the "optionality" feature of the model. Notice that traditional models of insurance have firms first write policies, then draw the results, whereas with the real option structure firms "first draw, then write"⁴. It seems that the real option structure is a better description of what the catastrophe insurance companies actually do.

3 Dynamic Competition and CAT Events

In this section, we address the issue of dynamic oligopolistic competition between different firms as opposed to the static case considered in the previous section.

In the presence of possible future CAT events, and taking into account contract irreversibility (or partial irreversibility), insurance policies can be treated as continuous investments into an industry with risky returns, analogous to the discrete model considered above and suggested in [7]. In this case, the investment policies may have significant "optionality" features when

⁴Thomas Russell suggested this useful distinction to us.

an oligopolistic firm follows its optimal dynamic strategy. In particular, the investment (insurance) may occur only at sufficiently large positive values of NPV (sufficiently low level of risk) implying a significant option value of waiting [2].

However, in the case when all firms are identical and the news come in a continuous stream, this "waiting" effect is diminished in the presence of oligopolistic competition. This result has a simple intuition. All firms are trying to avoid being "pre-empted" by their competitors in the market and therefore tend to wait less than the monopolists subjected to the same level of risk for their investment (insurance) [5]. In other words, the optimal investment occurs at lower (but positive) values of NPV. As we will see below, this is not always true in the presence of sufficiently large CAT events and heterogeneity across the firms.

In what follows, we concentrate on the analysis of general equilibrium in the presence of several competing firms (oligopoly). As the first step, we adopt a standard assumption that all agents are risk-neutral [2], [5]. We will show that the oligopolistic competition among heterogeneous firms in the presence of smooth shocks drives the economy towards the symmetric equilibrium, when all firms follow essentially the same strategy described in [5]. However, the significantly large CAT events may affect the optimal investment (insurance) strategies and reverse the effect in the case when the partial disinvestment (cancellation of insurance policies) is possible.

The intuition for this is the following. It turns out that the firms with smaller capacity (size) have a greater vulnerability to the CAT risk. Therefore, if the industry expects sufficiently large CAT events, it is optimal for small firms to disinvest (cancel policies) and therefore reduce the effects of competition for the larger firms still investing into the industry. As a result, we obtain a non-stationary distribution of firms with respect to their levels of investment, as opposed to the case of small (white noise) shocks which lead to the symmetric strategies for all firms. In order to solve for the fixed-point solution of the multiple-agent dynamical optimization problem arising in this case, we employ the dynamical aggregation approach proposed in [1].

3.1 Dynamic Optimization Problem

Generalizing the approach of the previous section, we extend the analogy between optimal insurance and investment policies to the case of dynamic

oligopolistic competition. We extend the stylized model of the CAT insurance considered above assuming that any insurance company may increase its volume of insurance by the amount dq_i at any time and collect the premium endogenously defined in the dynamic oligopolistic equilibrium analogously to the static case considered above.

To simplify our problem, we assume that all CAT risks are identical a priori, and that each insurance contract provides the policy holder with indemnification for the indefinite future (including the possibility of multiple payouts when multiple events occur). The premium paid is the market premium $p(t)$ at the time each specific policy is initiated. This premium will generally exceed the losses expected on the policy over its infinite life, with the precise amount determined by the oligopolistic structure of the industry. These upfront premiums represent the total resources available to the insurance firm to pay claims. As discussed in [8], there is less need for additional capital, when the firm uses long-term contracts with prepaid premiums. Still, actual claims could exceed the available resources, in which case the firm pays what it can and goes out of business-the same limited liability used by real world firms. We discuss possible extensions of the model to finite-duration contracts in the concluding section.

Once the CAT event occurs, the insurance company compensates the subscribers by the amount of loss represented by a state variable $x(t)$, up to the amount of available resources. At this point, we keep the state variable general, but will specify it below. As new information concerning expected future losses becomes available, firms may wish to raise or lower their exposure q . When firms expand, they sell new policies at the then current market premium $p(t)$. When firms contract, we assume they face penalties consisting of (1) the current market premium $p(t)$ (paid as compensation the policy holder for canceling the coverage) and (2) an additional cost which may be interpreted as the (foregone) opportunity value of staying in business after the shock. In the extreme, a firm may voluntarily go out of business entirely in which case it keeps any net positive balances that remain after paying its penalties. However, one should note that we assume that the firms's strategies $q_i(t)$ are smooth functions of time, implying that there are no "jumps" in $q_i(t)$.

Analogous to the results of the previous section and following [5], the optimization problem for oligopolistic competition among N heterogeneous insurance firms can be formulated in terms of a dynamic programming Cournot

auction problem for the i -th firm's project value V_i in the form

$$e^{-rt}V_i(t) = \max_{q_i(t)} E_t \left[\int_t^{+\infty} dt' e^{-rt'} \left\{ p(t') \frac{d}{dt'} q_i(t') - \varphi(t') \pi_i[q_i(t'), x(t')] \right\} \right], \quad (13)$$

where $q_i(t)$ and $Q(t) = \sum_{i=1}^N q_i(t)$ are the firm i and aggregate investments, respectively, $p(t)$ is a premium, $\pi[q_i(t), x(t)]$ is an instantaneous payoff (loss) of investment (analogous to the dividend process in asset pricing) and $\varphi(t')$ is the probability density of exactly one CAT event occurring at time t' . Note that in (13), we always have a constraint $\frac{d}{dt} q_i(t) \geq 0$, implying that the firm's capacity can either increase or remain on the same level. This constraint will be lifted in the next section where we take the disinvestment into account.

In the optimization problem (13), $x(t)$ is a state variable describing the cost of losses. Note that this variable represents the "bad news", as opposed to the investment case. It reflects the inflow of information on expected losses due to the future and the impact of the CAT events that have already happened. One should expect that $x(t)$ has at least two different components representing the gradual inflow of the information on the risk exposure and the instantaneous change of the payoffs due to the CAT event, respectively. While the gradual inflow of news is presented by a Geometric Brownian Motion, the CAT event is modeled by a Poisson jump process. Since the short-term predictions of CAT events typically do not have a lot of predictive power, we may assume that the Poisson jumps and GBM are only weakly correlated. As we will discuss below, our model can be extended to arbitrary strong correlations and this does not affect our main conclusions. Therefore, we assume

$$dx = x(\mu dt + \sigma dz + d\eta_{\lambda, \Delta}), \quad (14)$$

with μ and σ being the drift and volatility of the GBM, respectively, $z(t)$ is a standard Wiener process and $\eta_{\lambda, \Delta}(t)$ characterizing the Poisson jump process with a constant intensity λ and size Δ ; see Dixit and Pyndyck [2] for a discussion of real options models with combined Poisson and Geometric Brownian Motion stochastic processes. Making use of (13) yields

$$E_t[dV_i(t)] = rV_i(t) dt + \varphi(t) \pi_i[q_i(t), Q(t), x(t)] dt - p(t) dq_i(t), \quad (15)$$

and the probability density of CAT events is given by the probability of Poisson arrival

$$\varphi(t') = \lambda \exp[-\lambda(t' - t)]. \quad (16)$$

Although the extension of our results to arbitrary payoff function is straightforward, we will restrict the analysis to a specific form analogous to the previous section

$$\pi_i[q_i, x] = q_i x, \quad (17)$$

with the premium given by the inverse demand function

$$p(t) = K [Q(t)]^{-1/\gamma}, \quad (18)$$

constant $\gamma > 1$ characterizing the impact of aggregate demand on price

3.2 Solution for Optimal Policies

The optimal investment policy for the continuous investment with random exogenous shocks can be characterized as follows [2]. Each firm does not invest until the expected payoffs achieve a sufficiently high level, which in general is heterogeneous across all firms and has to be defined as a dynamic equilibrium function. This function $x_i(q_i, Q, t)$ is called a trigger function for the firm i [5]. Taking into account (14) making use of the corresponding Dynkin generators and assuming the stationary (infinite horizon) limit, we obtain from (15)

$$\begin{aligned} \left\{ \mu x \frac{\partial}{\partial x} + \frac{\sigma^2}{2} x^2 \frac{\partial^2}{\partial x^2} - r \right\} V_i(t, x) + \lambda [V_i(t, e^\Delta x) - V_i(t, x)] = \\ = \lambda \pi_i[q_i(t), x], \end{aligned} \quad (19)$$

for the firm i 's value function within the waiting region, with the condition at the investment (exercise) boundary given by

$$\frac{\partial}{\partial q_i} V_i(q_i, Q, x_i) = -p(t), \quad (20)$$

where x_i is the trigger function for the i -th firm to start investing. The additional smooth-pasting conditions with respect to x ensure that the value of the firm has been optimized

$$\frac{\partial^2}{\partial q_i \partial x} V_i(q_i, Q, x) \big|_{x=x_i} = 0. \quad (21)$$

Introducing the marginal value of investment $M_i(q_i, Q, x) = \frac{\partial}{\partial q_i} V_i(q_i, Q, x)$ yields

$$\left\{ \mu x \frac{\partial}{\partial x} + \frac{\sigma^2}{2} x^2 \frac{\partial^2}{\partial x^2} - r \right\} M_i(x) + \lambda [M_i(e^\Delta x) - M_i(x)] = \lambda x, \quad (22)$$

for the firm i 's marginal value within the waiting region, with the conditions at the investment (exercise) boundary given by

$$M_i(q_i, Q, x_i) = K [Q(t)]^{-1/\gamma}, \quad \frac{\partial}{\partial x} M_i(q_i, Q, x) \big|_{x=x_i} = 0, \quad (23)$$

where x_i is the trigger function for the i -th firm. Note that in the case of homogeneous firms $q_i \equiv Q/N$ and in absence of the CAT events ($\lambda = 0$), we arrive at the symmetric equilibrium analogous to the case considered in [5]. As we will see below, the heterogeneity across the firms in the presence of CAT events may drive the economy away from the symmetric equilibrium and lead to complex non-stationary optimal policies. Solving (22) with (23), we obtain

$$M_i(q_i, Q, x) = Bx + Dx^{-\delta}, \quad (24)$$

in the investment region $x_i \leq x \leq +\infty$, with

$$B = -\frac{\lambda}{r-\mu} \quad (25)$$

$$D = \frac{1}{\delta} x_i^{1+\delta} \frac{\lambda}{r-\mu},$$

with $\delta = \frac{\mu-\sigma^2/2}{\sigma^2} + \sqrt{\left(\frac{\mu-\sigma^2/2}{\sigma^2}\right)^2 + \frac{2\rho}{\sigma^2}}$ and $\rho = r + \lambda [1 - \exp(-\delta\Delta)]$.

Clearly, the effective exponent in presence of jumps is distorted in comparison to the one for the pure GBM, in which case $\rho = r$. Therefore, the

jump process leads to the shift of the discount rate. This can be interpreted as an effective hazard rate in the presence of the additional risk due to the jump process, and is analogous to the liquidity shock [3]. One should note that in the presence of jumps, the explicit analytical expression for the exponent δ is not available. However, the equation given above gives a complete characterization of the distortion introduced by the CAT events which can be solved numerically. Note that this distortion is small for sufficiently rare CAT events (which is normally the case). Therefore, the above equation can be approximately solved analytically. As we will see below, the main contribution of large but rare CAT events comes from the large distortion of optimal investment dynamics, as opposed to small distortions of the exercise boundary.

Using the boundary condition, we finally obtain the trigger functions in the form

$$x_i(q_i, Q) = kQ^{-1/\gamma}, \quad (26)$$

where $k = K \frac{(r-\mu)}{\lambda} \left(\frac{\delta}{\delta+1} \right) e^{-\Delta}$ is a constant. One should note that the trigger functions are heterogeneous across the firms with different current capacities q_i , even though the function (26) depends only on the aggregate capacity. The reason for this is that the trigger function given by (26) in the case of many competing firms should be understood as a "reaction function" of the firm i to the strategies of all other firms, according to the general concept of the Cournot equilibrium [5]. This is analogous to the situation arising in a simple two-period model considered in the previous section. Therefore, the aggregate demand in (26) should be separated into the capacity of the firm i itself q_i (choice variable) and the contribution of all other firms q' (which is not immediately affected by the firm i). Since $Q = q_i + q'$, (26) yields

$$x_i(q_i, q') = k(q_i + q')^{-1/\gamma}. \quad (27)$$

The dependence (27) is presented in Fig.1, where we take a numerical example for the large and small firms, respectively. The triggers of the large and small firms correspond to the dashed (red) and solid (green) lines, respectively.

Insert Fig.1 about here

In Fig. 1, the trigger functions show, for a given level of bad news x , the optimal amount of the firm's investment q . The trigger function for the large firm (dashed red line) is higher than the one for the small firm (solid green line). This implies that the large firm may continue increasing its capacity at a larger level of the state variable ("bad news") than the small one. This has a simple intuition, since the small firm can not strongly affect the aggregate volume and therefore the price by its decisions. This makes a small firm more vulnerable to the bad news and therefore more "risk averse" than the large one.

Note that both trigger functions belong to the same family parameterized by the current capacity of each firm. This indicates that the symmetric equilibrium is achievable within the same optimal set of strategies. In case of symmetric equilibrium, the two curves in Fig. 1 coincide. As we will see below, the symmetric equilibria are naturally achieved for the slow GBM news inflow. As we will see below, the heterogeneous decision boundaries combined with the large CAT events may drive the system to a different type of equilibrium which is highly asymmetric. In particular, the sufficiently large negative shocks can simply knock some firms out of business (note that we are considering optimal policies based on the assumption of the full rationality).

4 Dynamic Equilibrium with Disinvestment

To illustrate a model with *disinvestment*, we will consider the case where each insurance firm can reduce its level of insurance $q_i(t)$ provided that it pays a "penalty", as introduced earlier. If partial disinvestment is possible, the firm can reduce its capacity in case the bad news (the demand shock) is sufficiently large. Therefore, we obtain two boundaries for the investment and disinvestment, when the firm's capacity increases and decreases, respectively.

We assume that the disinvestment is associated with a marginal cost p_1 , implying that the investment is only partially irreversible. Since this cost corresponds to the value of staying in business for the insurance firm, we assume that p_1 is proportional to the current premium, $p_1(t) = K_1 [Q(t)]^{-1/\gamma}$ with $K_1 \geq K$. The last inequality comes from the fact that the "penalty" should be compared to the capitalized stream of the premia, which is proportional but greater than the one-period premium $p(t)$. In the concluding section, we discuss plans for a more general form of this penalty cost.

Since the firms take into account the disinvestment opportunities in their optimal strategies, the disinvestment decisions affect the investment boundaries for the firms. This means that both investment and disinvestment boundaries have to be defined simultaneously. This situation is typical for the systems that function in multiple regimes [?, ?].

In this case, we have an additional term in (13) describing the disinvestment process when the firm's capacity decreases.

$$e^{-rt}V_i(t) = \max_{q_i(t)} E_t \int_t^{+\infty} \{p(t') \frac{d}{dt'} q_i(t') \theta \left[\frac{d}{dt'} q_i(t') \right] + p_1(t') \frac{d}{dt'} q_i(t') \theta \left[-\frac{d}{dt'} q_i(t') \right] - \varphi(t') \pi_i[q_i(t'), x(t')]\} e^{-rt'} dt', \quad (28)$$

where $\theta(t)$ is an indicator step function, which equals one for positive and zero to negative values of its argument.

In the continuation region, the equation (22) and conditions on investment boundary (23) for the marginal value of investment still hold, but we have additional conditions on the disinvestment boundary. Analogous to (23), we obtain

$$M_i(q_i, Q, y_i) = p_1, \quad \frac{\partial}{\partial x} M_i(q_i, Q, x) |_{y_i} = 0, \quad (29)$$

where y_i is the disinvestment trigger function for the i -th firm and p_1 is a marginal cost of disinvestment. In this case, both power-law terms in the solution of the HJB equation (22) have to be taken into account. The maximization and pasting conditions are the same, but are to be applied on both boundaries. Analogous to (24), we obtain

$$M_i(q_i, Q, x_i) = Bx + Cx^\beta + Dx^{-\delta}, \quad (30)$$

where

$$B = -\frac{\lambda}{r-\mu},$$

$$Bx_i \left(1 + \frac{1}{\delta}\right) + Cx_i^\beta \left(1 - \frac{\beta}{\delta}\right) = p, \quad (31)$$

$$By_i \left(1 + \frac{1}{\delta}\right) + Cy_i^\beta \left(1 - \frac{\beta}{\delta}\right) = p_1,$$

$$x_i^\delta \left(C\beta x_i^\beta - Bx_i\right) = y_i^\delta \left(C\beta y_i^\beta - By_i\right),$$

with $\beta = -\frac{\mu-\sigma^2/2}{\sigma^2} + \sqrt{\left(\frac{\mu-\sigma^2/2}{\sigma^2}\right)^2 + \frac{2\rho}{\sigma^2}}$. Note that (31) is four equations with four unknowns $\{B, C, x_i, y_i\}$ representing two coefficients and two trigger functions x_i and y_i .

While the exact analytical solutions of the system (31) for arbitrary parameters are not available, it can be solved numerically. Moreover, one can use approximate solutions under assumptions that both investment and disinvestment boundaries are well separated and therefore do not strongly affect each other. In this case, the approximate solution for the investment boundary is still (26) and the disinvestment one is given by

$$y_i(q_i, q') = k_1 (q_i + q')^{-1/\gamma}, \quad (32)$$

where $k_1 = K_1 \frac{(r-\mu)}{\lambda} \left(\frac{\delta}{\delta+1}\right) e^{-\Delta} \geq \varkappa$ is a constant. As before, the triggers for both regimes are different across the firms with different capacities.

4.1 Comparative Dynamics of Investment

From (27) and (32), it follows that the investment decisions depend on the size of the firm (current level of investment). Therefore, some firms may decide to invest while the optimal strategies for the others will be to stay in the waiting region. This is in contrast to [5], where only symmetric investment strategies have been considered. The heterogeneous triggers have been considered in a different context in a paper [10], where the heterogeneity arises due to different investment costs across the firms. Note that in our case there is no such exogenous heterogeneity. In our model it arises due to the different capacities across the firms and ultimately due to the CAT events that occur in the economy.

In the oligopolistic equilibrium, the investment decisions of each firm is affected by the decisions of all competitors. This happens since the investment decisions of all firms change the aggregate capacity Q and therefore the price (inverse demand) which in its turn affects the investment boundaries (triggers) for different firms depending on their current capacities. Making use of (27) and (32), we obtain the elasticities of the firms with respect to the aggregate capacity of its competitors in the form

$$\frac{q'}{x_i} \frac{\partial x_i}{\partial q'} = -\frac{1}{\gamma} \left(1 - \frac{q_i}{Q}\right). \quad (33)$$

From (33), it follows that the elasticity is negative which is a consequence of the fact that the state variable x represents "bad news". The absolute value of the elasticity decreases in the firm's capacity. Analogously, the slope of the trigger function defined as the trigger elasticity with respect to the capacity of the firm is given by

$$\frac{q_i}{x_i} \frac{\partial x_i}{\partial q_i} = -\frac{1}{\gamma} \frac{q_i}{Q}, \quad (34)$$

and its absolute value is increasing in the firm's capacity.

From (33), it follows that the small firms are stronger affected by negative shocks than the large ones, which may be even made better-off in case if their competitors go out of business. Therefore, if some firm changes its capacities, it mostly affects the strategies of small but not of very large firms.

As it follows from (27) and (32), the small firms typically have higher values of the triggers since they are effectively more "risk averse" than the large ones. On the other hand, (34) indicates that the small firms' triggers have lower absolute elasticities with respect to their capacities. This is because the small firms act more like price-takers (they do not affect the price), whereas the large ones act more like strategic monopolists (since they know that they will strongly move the price). This creates additional distortion of the oligopolistic equilibrium.

As we will see below, these distortions remain small and do not prevent the economy from achieving the symmetric equilibrium for the GBM-type news, but become large and may strongly affect the equilibrium in case when the CAT events are present. The optimal investment policies can be characterized by a closed set of i non-linear SDEs (one for each firm)

$$\begin{aligned}
dq_i(t, q') &= \left\{ \frac{1}{\sqrt{k}} \theta(x - x_i(q_i, q')) \left[\frac{dq_i}{dx_i} \right] - \frac{1}{\sqrt{k_1}} \theta(y_i(q_i, q') - x) \left[\frac{dq_i}{dy_i} \right] \right\} dx, \\
\left[\frac{dq_i}{dx_i} \right] &= -\frac{\gamma}{k} (q_i + q')^{1+1/\gamma}, \\
\left[\frac{dq_i}{dy_i} \right] &= -\frac{\gamma}{k_1} (q_i + q')^{1+1/\gamma},
\end{aligned} \tag{35}$$

where $\theta(s)$ is an indicator step function, $\left[\frac{dq_i}{dx_i} \right]$ and $\left[\frac{dq_i}{dy_i} \right]$ are the marginal rates of investment and disinvestment at the boundaries, respectively. The random shock process is described by (14) and Q is the aggregate investment level, $Q(t) = \sum q_i(t)$. Dynamical equations (35) are completed by appropriate initial conditions $q_i^{(0)}$ representing initial level of investment for each firm i .

As we will see below, firms may have different (sub-optimal) levels of investments due to the CAT events that they previously experienced. In case of such events, it may be optimal for some firms to go out of business, i.e. to close their insurance policies. In principle, the set (35) with appropriate initial conditions gives a complete characterization of the optimal investment strategies for the oligopolistic investment economy. For example, it can be analyzed numerically for different paths of the random shocks $x(t)$. However, it is not in general analytically tractable. In order to get a more intuitive description of the economy, we will consider some approximate analytical solutions and discuss the simplifying assumptions required for these solutions to be valid. It turns out that these conditions have a simple economic intuition and may therefore be quite reasonable.

From (35), it follows that the optimal strategy of each firm depends on the actions of all others through the aggregate level of investment Q . As it follows from (33), the firms' reaction to the dynamics of the aggregate investment depends on size and is therefore heterogeneous across different firms. This makes the firms strategies potentially asymmetric in presence of large shocks.

In order to obtain an intuition on the nature of dynamical equilibrium, let us first consider the case of two competing firms (duopoly).

4.2 Optimal Policies for Duopoly

In this case, general equations (35) take the form

$$\begin{aligned}
 dq_1(t, Q) &= \left\{ \frac{1}{\sqrt{k}} \theta(x - x_1(q_1, Q)) \left[\frac{dq_1}{dx_1} \right] - \frac{1}{\sqrt{k_1}} \theta(y_1(q_1, Q) - x) \left[\frac{dq_1}{dy_1} \right] \right\} dx, \\
 dq_2(t, Q) &= \left\{ \frac{1}{\sqrt{k}} \theta(x - x_2(q_2, Q)) \left[\frac{dq_2}{dx_2} \right] - \frac{1}{\sqrt{k_1}} \theta(y_2(q_2, Q) - x) \left[\frac{dq_2}{dy_2} \right] \right\} dx, \\
 \left[\frac{dq_1}{dx_1} \right] &= \left[\frac{dq_2}{dx_2} \right] = -\frac{\gamma}{k} Q^{1+1/\gamma}, \quad \left[\frac{dq_1}{dy_1} \right] = \left[\frac{dq_2}{dy_2} \right] = -\frac{\gamma}{k_1} Q^{1+1/\gamma},
 \end{aligned} \tag{36}$$

with

$$\begin{aligned}
 x_1(q_1, q_2) &= k(q_1 + q_2)^{-1/\gamma}, \\
 x_2(q_2, q_1) &= k(q_1 + q_2)^{-1/\gamma}.
 \end{aligned} \tag{37}$$

Suppose that the initial levels of investment for the first and second firms are $q_1^{(0)}$ and $q_2^{(0)}$, respectively.

Graphically, the equilibrium investment strategies are presented in Fig.2 for the case with sufficiently smooth inflow of news (no CAT events are present) and without the opportunity for disinvestment. The investment boundaries depend on size (q_i) and are therefore different for the two firms. The lines $X_i(q)$ with $i = \{1, 2\}$ represent equilibrium investment boundaries for the firms 1 and 2, respectively.

Inset Fig.2 about here

The initial level of the signal is x_0 , then it decreases to a minimum value x_1 . The initial states of the small and large firms are presented in Fig.2 by the points $S1$ and $S2$, respectively. The final states of the small and large firms are $S1'$ and $S2'$, respectively. This demonstrates that if the lowest level of the signal is sufficient to move at least one of the firms onto the investment boundary (the level of bad news is low), then the investment capacities of both firms get closer at the end of the process. Since the negative news does not affect the level of investment, this process should ultimately

drive the investment levels of both firms together. As discussed above, in the limit when the two firms have the same capacities, their trigger functions coincide. Note that when the firms' capacities change, the trigger functions subsequently change according to the current capacity of each firm. The graph presented above gives a qualitative picture of this convergence to the symmetric strategy.

Since the disinvestment is not allowed, the negative news (large x) does not affect the firms' capacities. For example, suppose that in the situation considered in Fig. 2, the state variable returns back to the original value x_0 at the end of a process. This does not affect the firms' final capacities after x_1 and therefore the conversion still takes place. Therefore, the positive inflow of news always moves different firms closer in their investment levels due to the fact that the exercise boundary $X(q)$ is decreasing in q , and this is not affected by the negative news if the disinvestment is not allowed. Clearly, an analogous conclusion should hold for arbitrary number of firms under these conditions.

In Fig.3, we consider a similar case in presence of sufficiently rare but strong CAT events taking into account the possibility of partial (costly) disinvestment. In this case, we obtain a second (disinvestment) boundary along which the disinvestment process. Since we assume that the disinvestment is costly, the disinvestment curves $X_-(q)$ (dotted lines in Fig.3), go above the investment curves $X_+(q)$ (solid and dashed lines). The investment and disinvestment curves for the small and large firms are given by the solid green, solid red, dashed green and dashed red lines, respectively. The main difference between Figs.2 and 3 is that in case of a sufficiently strong negative CAT event, the firms do not get closer with respect to their capacities (as opposed to the small GBM shock described by Fig.2). In particular, firm 1 gets "kicked out of business" by a sufficiently large negative shock considered in Fig.3. On the other hand, firm 2 with a larger initial capacity does not even achieve its disinvestment boundary and stays in the waiting region.

Insert Fig.3 about here

The important feature of the optimal dynamics illustrated in Fig.3 is that the investment strategies "remember" their history. According to [2], this can be interpreted as the "hysteresis" effects in the economy with strategically

competing firms. For example, in a situation illustrated in Fig.3, the firm 2 (the one with a larger initial capacity) does not disinvest since it was caught by the CAT event far from the disinvestment boundary and ends up at the initial level of investment. The firm 1 (with a small initial capacity) on the other hand is sufficiently close to the boundary at the moment of the CAT shock and therefore it completely disinvests with a final level $q_1(t_2) = 0$. As we see from Fig.3, the initial capacity difference between the two firms increases at the end of the shock. Clearly, the same effect should take place for arbitrary number of the firms. This indicates that in presence of sufficiently strong CAT events, the symmetric equilibrium is not dynamically stable (actually, it is not dynamically achievable). One should expect that the equilibrium distribution of firms with respect to their investment levels should be non-stationary and sufficiently complex.

5 Size Distribution of the Firms

In this section, we will consider some analytical results regarding the size distribution of the insurance companies subjected to both expected and unexpected losses arriving in a form of shocks. The optimal investment policies can be characterized by a closed set of i non-linear SDEs (one for each firm)

$$\begin{aligned} dq_i(t, q') &= \left\{ \theta(x_i(q_i, q') - x) \left[\frac{dq_i}{dx_i} \right] - \theta(x - y_i(q_i, q')) \left[\frac{dq_i}{dy_i} \right] \right\} dx, \\ \left[\frac{dq_i}{dx_i} \right] &= -\frac{\gamma}{k} (q_i + q')^{1+1/\gamma}; \quad x_i(q_i, q') = k (q_i + q')^{-1/\gamma}, \\ \left[\frac{dq_i}{dy_i} \right] &= -\frac{\gamma}{k_1} (q_i + q')^{1+1/\gamma}; \quad y_i(q_i, q') = k_1 (q_i + q')^{-1/\gamma}, \end{aligned} \quad (38)$$

where $\theta(s)$ is an indicator step function, $\left[\frac{dq_i}{dx_i} \right]$ and $\left[\frac{dq_i}{dy_i} \right]$ are the marginal rates of investment and disinvestment at the boundaries, respectively. The random shock process is described by (14) and Q is the aggregate investment level, $Q(t) = \sum q_i(t)$. Dynamical equations (38) are completed by appropriate initial conditions $q_i^{(0)}$ representing initial level of investment for each firm i . The equation (38) can be considered as a set of equations modelling the dynamics of the whole insurance economy, in a spirit of the approach

described in [2]. From this prospective, the solution of (38) describes the general equilibrium in the economy. One should note that this equilibrium may not be of a static nature depending on the characteristics of the shocks. In other words, the set of dynamic equations (38) may or may not achieve a steady state regime, when the characteristics of the size distribution of the firms are time-independent. As we will see below, the steady-state solutions formally exist, but may be achieved only after a sufficiently long time period (the time "lag").

With the substitution

$$\psi_i(q_i, q') = (q_i + q')^{-1/\gamma}, \quad (39)$$

the system (38) transforms into

$$d\psi_i = \left\{ \frac{1}{k} \theta \left(\psi_i - \frac{1}{k} x \right) - \frac{1}{k_1} \theta \left(\frac{1}{k_1} x - \psi_i \right) \right\} dx. \quad (40)$$

Given (40), we obtain the following forward Kolmogorov equation for the distribution function $F(\psi, x, t)$

$$\begin{aligned} \frac{\partial F}{\partial t} = & \frac{\partial}{\partial x} \left\{ \mu x F + \lambda x [F(xe^\Delta) - F(x)] + \frac{\sigma^2}{2} x^2 \frac{\partial F}{\partial x} \right\} + \\ & + \frac{\partial}{\partial \psi} \left\{ \varphi(\psi, x) [\mu x F + \lambda x [F(xe^\Delta) - F(x)]] \right\} + \\ & + \frac{\partial}{\partial \psi} \left\{ [\varphi(\psi, x)]^2 \frac{\sigma^2}{2} x^2 \frac{\partial F}{\partial \psi} \right\}, \end{aligned} \quad (41)$$

with

$$\varphi(\psi, x) = \frac{1}{k} \theta \left(\psi_i - \frac{1}{k} x \right) - \frac{1}{k_1} \theta \left(\frac{1}{k_1} x - \psi_i \right). \quad (42)$$

The partial differential equation (41) reduces in the limit of large CAT shocks $\Delta \gg 1$ to

$$\begin{aligned} \frac{\partial F}{\partial t} = & \frac{\partial}{\partial x} \left\{ (\mu - \lambda) x F + \frac{\sigma^2}{2} x^2 \frac{\partial F}{\partial x} \right\} + \\ & + \frac{\partial}{\partial \psi} \left\{ \varphi(\psi, x) (\mu - \lambda) x F + [\varphi(\psi, x)]^2 \frac{\sigma^2}{2} x^2 \frac{\partial F}{\partial \psi} \right\}. \end{aligned} \quad (43)$$

The partial differential equation (41) describes the time evolution of the joint probability distribution of shocks x and the variable $\psi(q)$ related to the capacity of the firm. One should note that the solution of (41) should also satisfy the self-consistency condition at the upper boundary $\psi_m = (q')^{-1/\gamma}$

$$\begin{aligned} E[\psi^{-\gamma}] &= E[q + q'] = Q, \\ (\psi_m)^{-\gamma} &= q' = \frac{N-1}{N}Q, \end{aligned} \quad (44)$$

and therefore

$$\frac{N-1}{N} = \frac{(\psi_m)^{-\gamma}}{E[\psi^{-\gamma}]}, \quad (45)$$

implying that the expected aggregate capacity of all firms except for a given one is related to the expected aggregate capacity. The probability distribution in case of the large shocks described by the PDE (43) is characterized by

Theorem 1 *The solution of the PDE (43) is given by*

$$\begin{aligned} \Psi(\psi, x, t) &= J(t) \Gamma(t) \exp \left[-\frac{(\psi - \bar{\psi}(t))^2}{2\Delta(t)} \right] \exp \left[\frac{(\psi_m - \bar{\psi}(t))^2}{2\Delta(t)} \right], \\ \bar{\psi}(t) &= \int_0^t dt' \alpha(t') + \psi_a, \quad \alpha(t) = \eta A(t), \\ \Delta(t) &= \int_0^t dt' \beta(t') + \sigma_0^2, \quad \beta(t) = \sigma^2 B(t), \\ \Gamma(t) &= \frac{2\Delta(t)}{\omega(t) - \rho(t) \left(\frac{\psi_m - \bar{\psi}(t)}{\sqrt{\sigma^2 t}} \right)}, \\ \omega(t) &= 2\alpha(t) \sigma_0^2 + 2\beta \int_0^t dt' \alpha(t'), \\ \rho(t) &= \beta(t) \sqrt{\sigma^2 t}, \end{aligned} \quad (46)$$

where the upper cut-off parameter ψ_m is defined through the self-consistency ("fixed point") condition (45).

Proof: See Appendix.

Given the solution (46), we can define the total inflow of the active firms into the industry $J(t)$ through the self-consistency condition at the upper boundary ψ_m . This inflow of the firms into the economy is presented as $J = \frac{d}{dt}N$ and therefore the consistency conditions can be viewed as Ordinary Differential Equation (ODE) describing the dynamics of the total equilibrium number of the firms in the economy as a function of time and the state variable x . These arguments are similar to the fixed-point type considerations which are widely used in the General Equilibrium models and can be viewed as an analogue of the "market clearing" conditions. The ODE may lead to a rich variety of possible solutions depending on the parameters of the economy and is characterized by

Corollary 1 *The aggregate number of the active firms $N(t)$ in the industry is given by a solution of the following ODE*

$$\begin{aligned} \left(1 - \frac{1}{N}\right)^{-1/\gamma} &= \left(1 + \frac{1}{\rho\psi}\theta\right) \left(1 + \frac{\rho^2\sigma^2 t}{4\theta^2}\right), \\ \theta &= \omega(t) - \frac{\nu(t)}{N} \frac{dN(t)}{dt}, \end{aligned} \quad (47)$$

where the parameter $\nu(t)$ is defined in the Appendix.

Proof: See Appendix.

The phase portrait of the ODE (47) is given in the Fig. 4.

Insert Fig.4 about here

The motion has one or two focal points for the negative and positive drift term $\alpha(t)$ and presented in Figures 4a and 4b, respectively. Making use of the expression for the drift (52) from the Appendix, we obtain a simple intuition for these two regimes, since the drift exponentially depends on the level of expected losses and positive or negative drift corresponds to the low and high level of expected losses, respectively. Clearly, all firms will leave the industry in case when a large shock arrives and expected losses become sufficiently high. On the other hand, the small expected losses (good news)

may attract more firms into the industry. In case of positive drift, (47) has a non-trivial stable equilibrium point

$$N_e(x, t) \approx \frac{4}{\gamma} \left[\frac{\omega(t)}{\rho(t) \sqrt{\sigma^2 t}} \right]^2 = \frac{16}{\gamma} \left[\frac{\int_0^t dt' \alpha(t')}{\sigma^2 t} \right]^2, \quad (48)$$

which gives an estimate $N_e(x, t) \approx \frac{16}{\gamma} \exp(-2x) \left[\frac{\eta}{k\sigma^4} \right]^2$. Clearly, the equilibrium number of firms decreases in the size of losses, as expected. One should note that in our model all optimal strategies are assumed to be smooth functions, implying that even in case of large expected losses when all firms tend to close their operations, it takes some time before all firms exit the industry. This is reflected in the dynamic nature of the equilibrium described by the ODE (47).

6 Conclusion

This paper suggests a model of price dynamics and of firm exit and reentry based on the concepts of real option theory to catastrophe insurance markets. We show that the decision of an insurance firm to take on catastrophe risks can be considered as an investment decision. Making use of this analogy, we show that the optimal policies of a catastrophe insurance firm can readily fit into the traditional real options model. Catastrophe insurance investments, however, may be reversible in the sense that an insurance firm can allow its current policies to mature, and then leave the industry. Thus, to apply the real option approach to catastrophe insurance, we have to amend the basic theory to cover the case of reversible investments.

The key benefit of applying real option theory to catastrophe insurance markets is that we obtain a unified theory of price dynamics and exit and entry decisions and one that fits the stylized facts of catastrophe insurance markets remarkably well.

Finally, we want to describe three topics on our own research agenda for extending the basic model of this paper.

1) In the current paper, the demand for insurance by policy holders is not directly a function of the loss variable x . It is sensible, however, to include an effect in which policy holder demand functions shift upward when new information suggests greater possible future losses. While we do not expect

this addition to change the general qualitative results of this paper, there may be special cases of policy holder demand which do create singular results. We hope to identify these cases as a result of this further research.

2) In the current paper, the loss variable x affects all firms in the same proportion, in the manner of an industry effect. Insurance firms, however, are heterogeneous not only in size (which is considered in the current paper), but also in terms of the specific geographic location of their risks. Thus, we plan to augment the basic model of the current paper with a second loss variable that is firm specific. We expect that this factor will provide a further basis for the heterogeneity that we actually observe across insurance firms.

3) In the current paper, a cost is imposed when firms reduce the level of their industry participation. Furthermore, this cost is specified as proportional to a firm's current premium income, in the manner of an opportunity cost of foregone income. It will be useful, however, to understand how our results might vary as a function of alternative specifications of the exit cost.

4) The current model does not take into account strategic behavior of the firms with respect to the "limited liability" assumption of the model when the firm with lower level of wealth (capital) will tend to take more risk than the otherwise identical firm with a higher level of capital. In this case, the level of capital plays the role of an extra degree of freedom in the model. This will be a subject of our future research.

7 Appendix

Proof of Theorem 1

Integrating both sides of the equation (41) over the distribution of shocks given by (14), we obtain the forward Kolmogorov equation for the marginal probability distribution of $\Psi(\psi, t)$ corresponding to the firms' size distribution, in the form

$$\frac{\partial \Psi(\psi, t)}{\partial t} = \frac{\partial}{\partial \psi} \left\{ \eta A(\psi, t) \Psi(\psi, t) + \frac{\sigma^2}{2} B(\psi, t) \frac{\partial \Psi(\psi, t)}{\partial \psi} \right\}, \quad (49)$$

where

$$\begin{aligned} \eta &= \mu + \lambda (e^\Delta - 1), \\ A &= E[\varphi(\psi, x) x] = \\ &= x e^{\eta t} \left\{ \frac{1}{k} \Phi \left[\frac{z_0}{\sigma \sqrt{t}} \right] - \frac{1}{k_1} \left(1 - \Phi \left[\frac{z_1}{\sigma \sqrt{t}} \right] \right) \right\}, \\ B &= E[\varphi(\psi, x) x^2] = \\ &= x^2 e^{2\eta t} \left\{ \frac{1}{k^2} \Phi \left[\frac{\tilde{z}_1}{\sigma \sqrt{t}} \right] + \frac{1}{k_1^2} \left(1 - \Phi \left[\frac{\tilde{z}_0}{\sigma \sqrt{t}} \right] \right) \right\}, \end{aligned} \quad (50)$$

the expectation conditioned on ψ and t $E[.]$, the Poisson distribution $P_n(t) = (\lambda t)^n / n! \exp(-\lambda t)$ and the standard error function $\Phi[.]$, with

$$\begin{aligned} z_0 &= \log \psi - \log x + \log k - (\eta + \sigma^2/2) t, \\ z_1 &= \log \psi - \log x + \log k_1 - (\eta + \sigma^2/2) t, \\ \tilde{z}_0 &= \log \psi - \log x + \log k - \eta t, \\ \tilde{z}_1 &= \log \psi - \log x + \log k_1 - \eta t. \end{aligned} \quad (51)$$

From equation (50), it follows that the effective distributions have "cuts" for both sufficiently large or small sizes of the shocks corresponding to the investment and disinvestment boundaries, respectively. This feature should be expected in the general equilibrium type models of investment (insurance) involving the trigger-type strategies. Indeed, the "cuts" correspond to the

trigger strategy being "switched on". This is taken into account in a self-consistent way at equilibrium by the equation (49). In the range of sufficiently long evolution times, we approximately have

$$\begin{aligned} A(t, x) &\approx \frac{1}{2} \exp\left(\frac{1}{2}\sigma^2 t\right) \left[\frac{1}{k} \exp(-\bar{x}) - \frac{1}{k_1} \exp(\bar{x}) \right], \\ B(t, x) &\approx \frac{1}{2} \exp(2\sigma^2 t) \left[\frac{1}{k^2} \exp(-2\bar{x}) + \frac{1}{k_1^2} \exp(2\bar{x}) \right], \\ \bar{x} &= x + \eta t. \end{aligned} \quad (52)$$

Assuming that the initial firms size distribution was normal $N(\psi_a, \sigma_0^2)$, applying the method of Green's functions for (49) and making use of (52), we obtain a solution of the evolution equation (49) in the form

$$\begin{aligned} \Psi(\psi, x, t) &= J(t) \Gamma(t) \exp\left[-\frac{(\psi - \bar{\psi}(t))^2}{2\Delta(t)}\right] \exp\left[\frac{(\psi_m - \bar{\psi}(t))^2}{2\Delta(t)}\right], \\ \bar{\psi}(t) &= \int_0^t dt' \alpha(t') + \psi_a, \quad \alpha(t) = \eta A(t), \\ \Delta(t) &= \int_0^t dt' \beta(t') + \sigma_0^2, \quad \beta(t) = \sigma^2 B(t), \\ \Gamma(t) &= \frac{2\Delta(t)}{\omega(t) - \rho(t) \left(\frac{\psi_m - \bar{\psi}(t)}{\sqrt{\sigma^2 t}}\right)}, \\ \omega(t) &= 2\alpha(t) \sigma_0^2 + 2\beta \int_0^t dt' \alpha(t'), \\ \rho(t) &= \beta(t) \sqrt{\sigma^2 t}, \end{aligned} \quad (53)$$

with the "inflow" of the firms' distribution defined as $J(t) = \frac{d}{dt} N(t)$ and given by

$$J(t) = \eta A(\psi_m, t) \Psi(\psi_m, t) + \frac{\sigma^2}{2} B(\psi_m, t) \frac{\partial \Psi(\psi, t)}{\partial \psi} \Big|_{\psi=\psi_m}. \quad (54)$$

■

Proof of Corollary 1

The total number of active firms $N(x, t)$ and the expectation $E[\psi^{-\gamma}]$ are directly obtained from the distribution (53) as

$$N(x, t) = J(t) \sqrt{\frac{\pi \Delta(t)}{2}} \Gamma(t) \left[1 + \frac{\psi_m - \bar{\psi}(t)}{\sqrt{(\psi_m - \bar{\psi}(t))^2 + \sigma^2 t}} \right], \quad (55)$$

and

$$E[\psi^{-\gamma}] = \left\{ \frac{\bar{\psi}(t)}{2} \left[1 + \frac{\psi_m - \bar{\psi}(t)}{\sqrt{(\psi_m - \bar{\psi}(t))^2 + \sigma^2 t}} \right] \right\}^{-\gamma}, \quad (56)$$

respectively. Combining (53), (53) and (45), we obtain

$$\begin{aligned} \left(1 - \frac{1}{N}\right)^{-1/\gamma} &= \left(1 + \frac{1}{\rho \bar{\psi}} \theta\right) \left(1 + \frac{\rho^2 \sigma^2 t}{4\theta^2}\right), \\ \theta &= \omega(t) - \frac{\nu(t)}{N} J(t), \quad \nu(t) = \sqrt{\pi} [\Delta(t)]^{3/2}. \end{aligned} \quad (57)$$

After the substitution $J(t) = \frac{d}{dt} N(t)$, we obtain the ODE (47).

■

References

- [1] A. Boulatov and D. Hackbarth [2002], "A Model of Liquidity Risk in Dynamic Economies", Working paper, Haas School of Business, UC Berkeley.
- [2] Dixit, A.K., and R.S. Pindyck [1994], "Investment Under Uncertainty", Princeton University Press, Princeton, New Jersey.
- [3] Duffie, Darrell [2001], "Dynamic Asset Pricing Theory", (Third edition), Princeton University Press, Princeton, New Jersey.
- [4] Kenneth A. Froot, Steven E. Posner [2001], "The pricing of event risks with parameter uncertainty", NBER working paper #8106, NBER.
- [5] Steven R. Grenadier [2001], "Option Exercise Games: An Application to the Equilibrium Strategies of Firms", *Review of Financial Studies*, forthcoming.
- [6] Xin Guo, Juianjun Pan and Erwan Morellec [2002], "Irreversible Investment with regime Shifts", Working paper, Simon School of Business, University of Rochester.
- [7] Jaffee, Dwight and Thomas Russell [2002], "Markets Under Stress: The case of Extreme Events Insurance," forthcoming in Festschrift volume for Joseph Stiglitz.
- [8] Jaffee, Dwight and Thomas Russell [1997], "Catastrophe Insurance, Capital Markets, and Uninsurable Risks," *Journal of Risk and Insurance*, Vol 64 No 2, pp 205-230.
- [9] Jun Liu, Jun Pan and Tan Wang [2002], "An Equilibrium Model of Rare Event Premia", Working paper, Sloan School of Management, MIT.
- [10] R. Novy-Marx [2002], "An Equilibrium Model of Investment Model of Liquidity Risk in Dynamic Economies", Working paper, Haas School of Business, UC Berkeley.

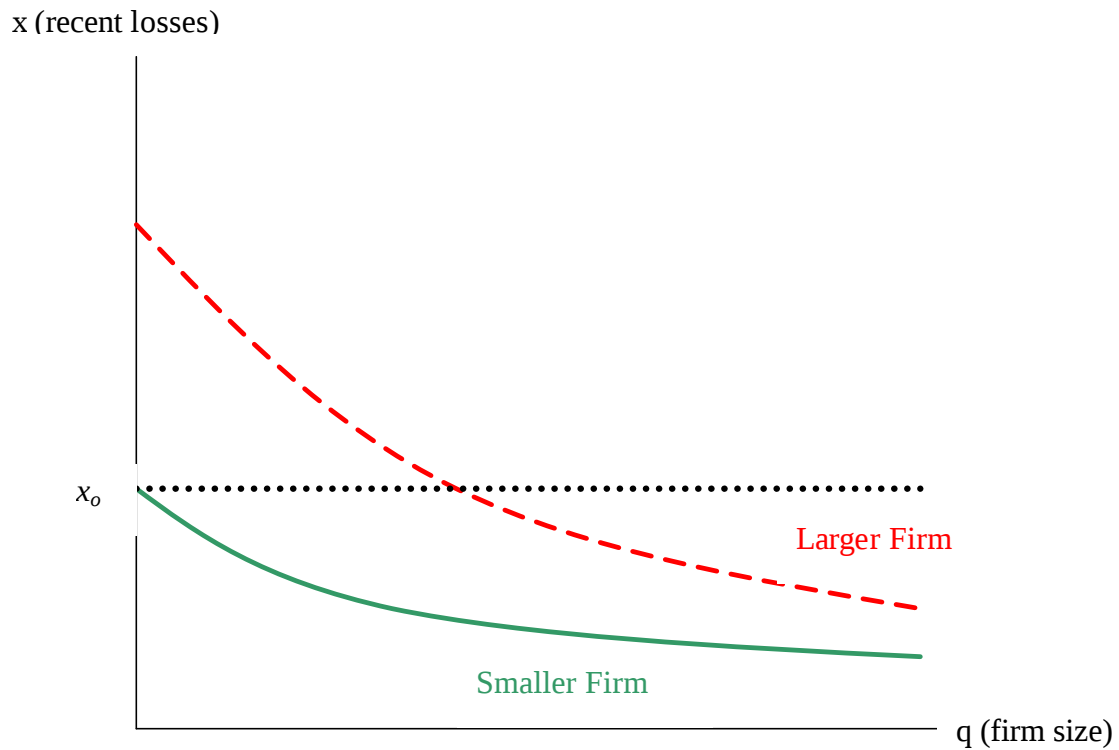


Figure 1. The figure shows the trigger functions $x(q)$ for two firms. A trigger function indicates a firm's desired size q as a function of the level of recent losses x . The smaller firm's curve is flatter and lower than the curve for the larger firm. Thus, for any level of recent losses, the larger firm desires a larger capacity. At sufficiently large losses a firm will choose a size of 0 (shown as x_0 for the small firm)

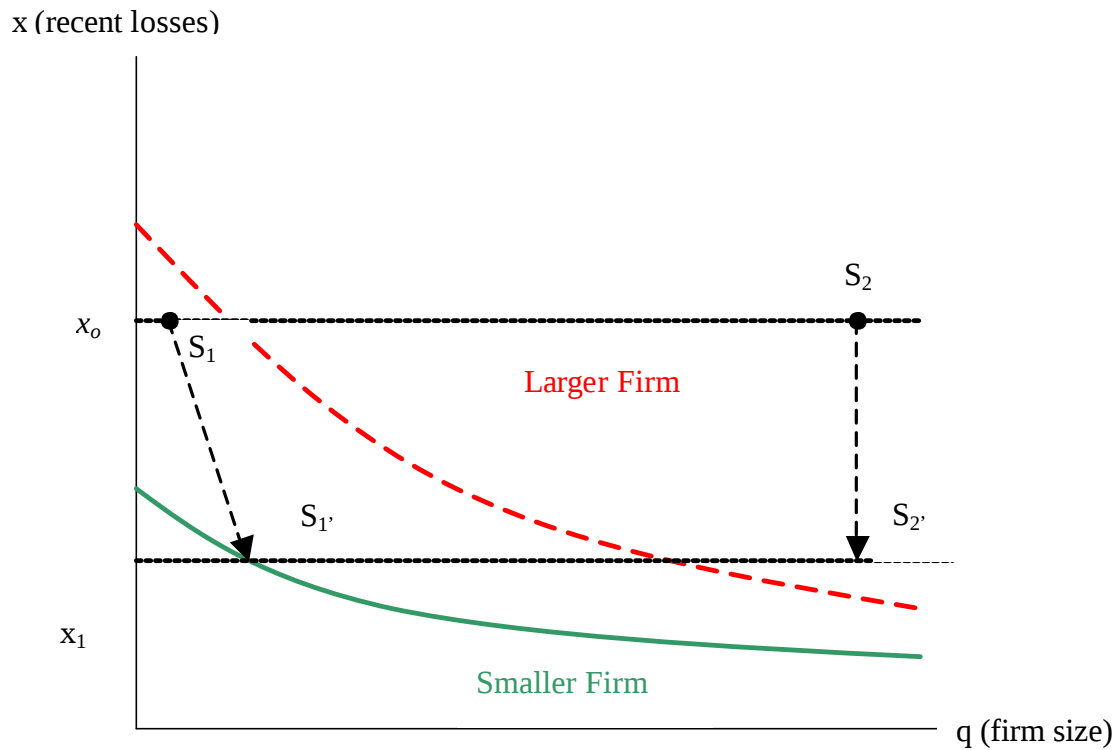


Figure 2. The figure shows two firms in arbitrary initial positions labeled as S_1 and S_2 respectively at an initial level of losses x_0 , and assuming firms can increase, but not decrease their capacity. Upon receiving new information on losses, x_1 , the smaller firm expands its capacity to S_1' . The larger firm would prefer to decrease its capacity, but cannot, so its capacity is unchanged at S_2 . This illustrates the basic pressure forcing firms to the same ultimate capacity, that is toward symmetry.

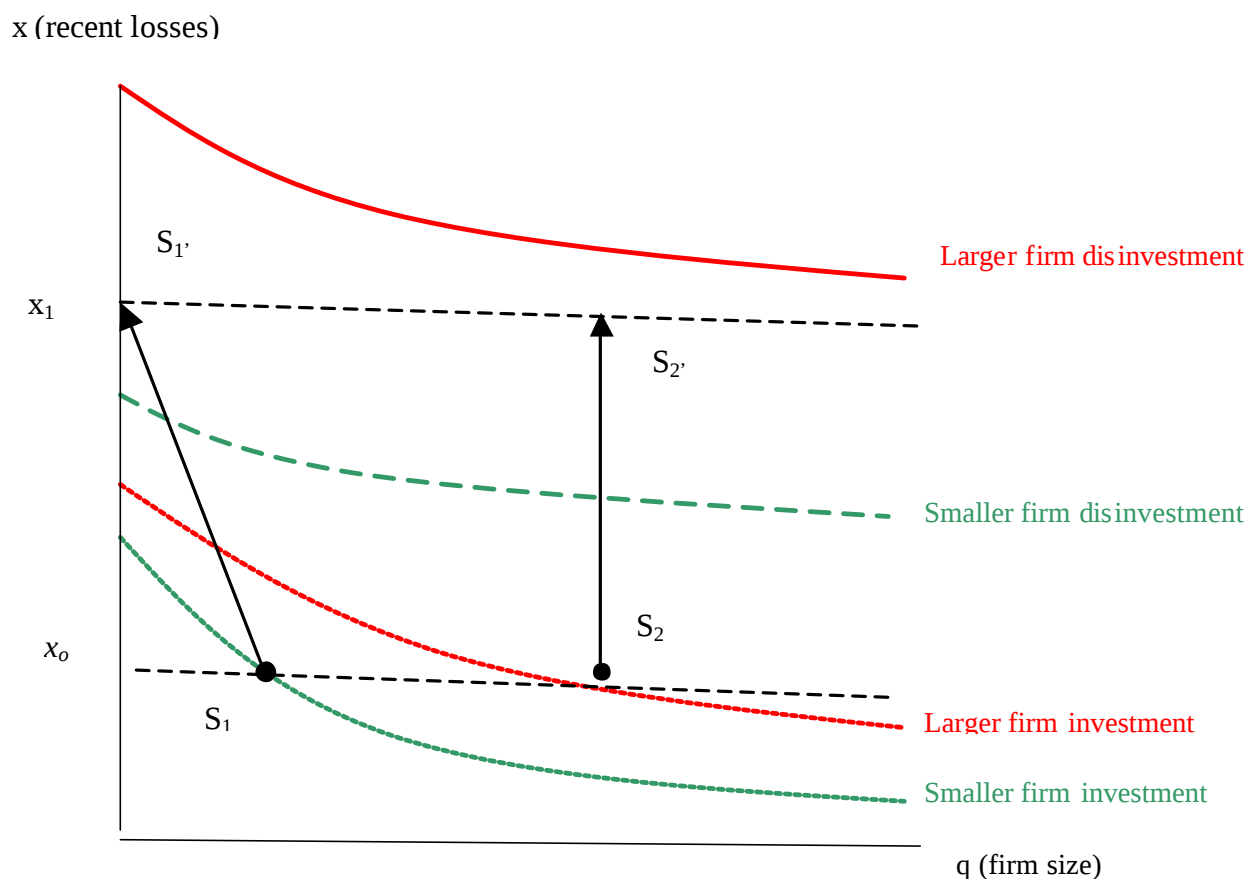


Figure 3. The figure shows the smaller firm and the larger firm both starting on their investment curve at points S_1 and S_2 respectively, based on losses x_0 . When the losses rise to x_1 , the smaller firm leaves the industry (S_1'), while the larger firm does not even reach its disinvestment boundary (S_2). This illustrates how large CAT shocks can lead to heterogeneity across firm size.

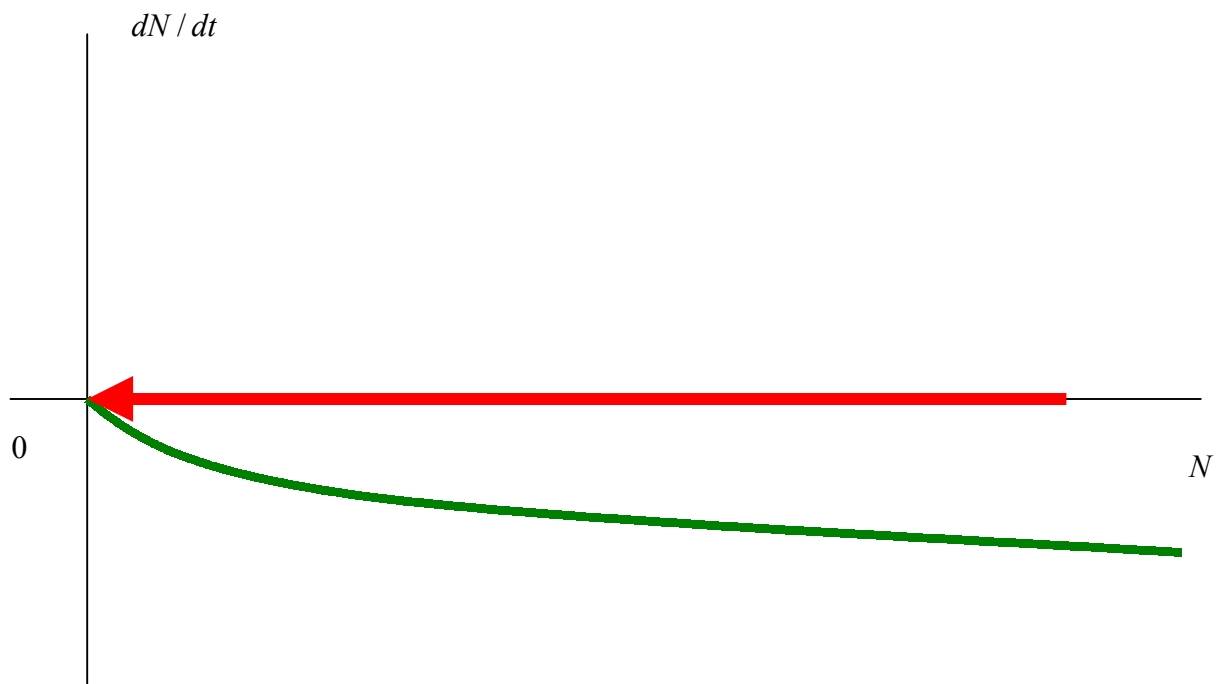


Fig. 4a

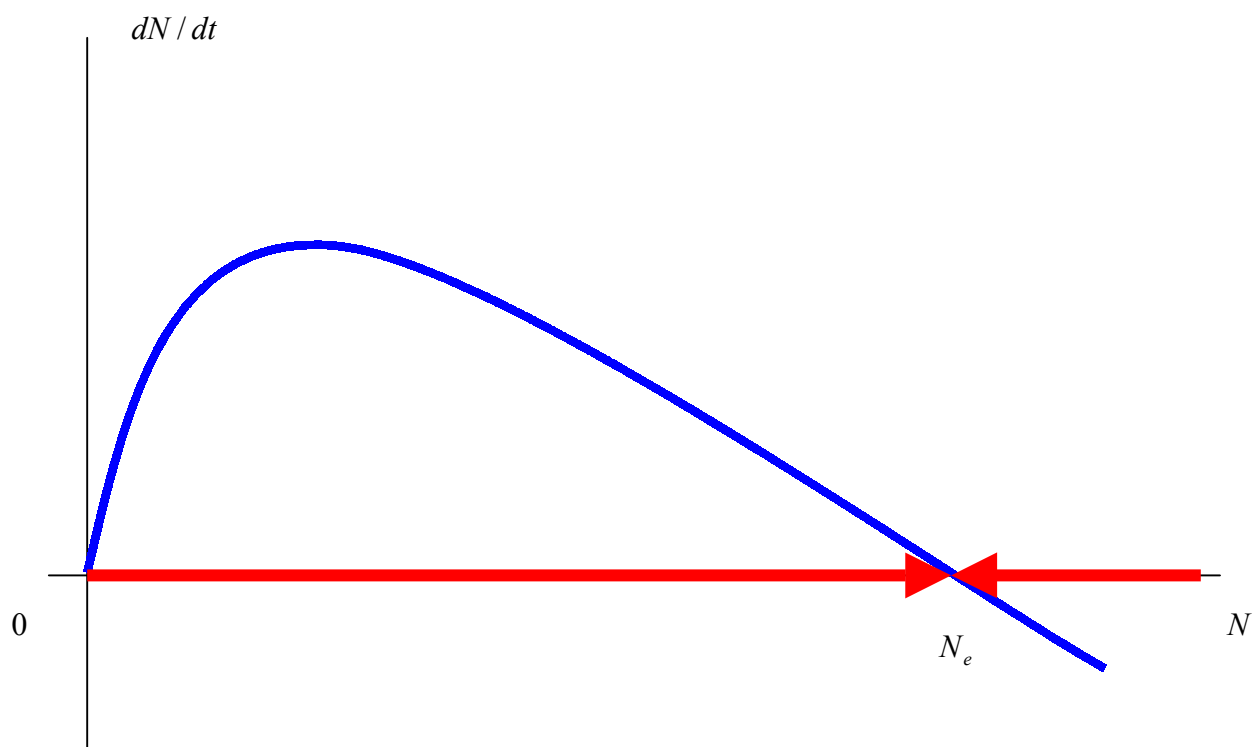


Fig. 4b