Universität
Basel

Wirtschaftswissenschaftliche
Fakultät

WWZ forum

February 2018

Adaptation to catastrophic events with two layers uncertainty: Central planner perspective

WWZ Working Paper 2018/10

Anton Bondarev, Frank C. Krysiak

A publication of the Center of Business and Economics (WWZ), University of Basel.

Universität Basel
Peter Merian-Weg 6
4052 Basel, Switzerland
wwz.unibas.ch

Corresponding Author:
Dr. Anton Bondarev
Tel: +41 (0) 61 207 33 81
Mail: anton.bondarev@unibas.ch

Adaptation to catastrophic events with two layers uncertainty: Central planner perspective*

Anton Bondarev[†1] and Frank C. Krysiak[1]

[1]University of Basel, Department of Business and Economics

February 16, 2018

Abstract

We study the optimal adaptation to extreme climate events by the central government in a setup where events are dynamically uncertain and the government does not know the true probabilities of events. We analyze different policy decision rules minimizing expected welfare losses for sites with different expected damages from the catastrophic event. We find out under which conditions it is optimal to wait before implementation of a prevention measures to obtain more information about the underlying probabilistic process. This waiting time crucially depends on the information set of the planner and the implemented learning procedure. We study different learning procedures on behalf of the planner ranging from simple perfect learning to two-layers Bayesian updating in the form of Dirichlet mixture processes.

Keywords: climate change adaptation; catastrophic events; model uncertainty; Bayesian updating; Dirichlet mixtures

JEL codes: Q54, H41, C41

*This research is part of the activities of CCAdapt project (Climate Change Adaptation), which is financially supported by the Swiss National Science Fond (SNF).

†Corresponding author; email: anton.bondarev@unibas.ch

1 Introduction

In the past decades, economic studies on climate policy have focused on mitigation. Integrated climate assessment models have been developed, such as (Nordhaus, 2007), and game theory has been used to assess the potential for stable coalitions for mitigating climate change (Babonneau et al., 2013). However, so far actual climate policy has not stopped climate change but rather helped to slow it down. Thus the question of climate change adaptation is gaining in importance.

Indeed, it is widely accepted that global warming will lead to a higher frequency of natural disasters in the near future. Thus it is important to understand how countries may optimally adapt by choosing the right adaptation strategy and the right timing to implement that strategy.

The literature on adaptation strategies has rapidly increased in recent years. It has been specifically considered in the context of international cooperation in Buob and Stephan (2011) and Kane and Shogren (2000). De Bruin and Dellink (2011) focus on different constraints (including informational) for efficient adaptation.

However a particularly intriguing problem has received little attention so far: Given that there is still a high level of uncertainty regarding climate change impacts and the speed of climate change, what adaptation measures should be implemented and at which time. This problem is of particular interest, as there is a chance to learn by waiting but a large-scale disaster could cause high costs, if adaptation measures have been delayed too long.

In this paper, we use a stylized model to analyze this question. In our model, the likelihood of catastrophic events is influenced by climate change but this influence is uncertain. A social planner can decide between no adaptation, a small and a large adaptation project, where investment in the large project involves substantial sunk costs. Depending on the beliefs of the social planner regarding the rate and scale of the climate change and on the availability of information, different adaptation strategies (that is, a different timing of the different projects) can be optimal. Our results show that it is frequently optimal for the planner to wait for some time before implementing adaptation measures. Moreover, it might be optimal to start with minor measures and switch to major efforts only later on, when the effects of climate change become clearer.

The main contribution of this paper is that we account for a rich set of possible probabilistic structures, which include not only usual one-layered uncertainty over coming catastrophic events, but also multiple layers of such uncertainty, where the planner does not know the true probabilistic law governing the likelihood of future disasters. We explicitly derive decision rules for the implementation of different adaptation measures under full information concerning probability distribution (Section 3), under perfect and imperfect learning over time (Section 4) and under fully uncertain climate change, where the planner does not know and cannot learn the true probability law (Section 5).

2 The Model

We use a simple model of climate change adaptation to study different types of uncertainty. We assume that a central government minimizes expected total costs, which stem from damages from potential catastrophic events as well as from adaptation costs. There are two regions, indexed $\{1, 2\}$, which can suffer from natural disasters. Both regions can have either low or high vulnerability to catastrophic events, $t \in \{L, H\}$. At each point of time ($T \in \{0, 1, 2, .., \infty\}$), one event may happen, affecting both regions simultaneously. The event may ave low or high associated damages. The risk of a catastrophic event is uncertain, and there are three possible outcomes with associated probabilities.

$$E = \{0, l, h\},\ \forall T \in \{0, 1, 2, .., \infty\} : \ p_0(T) + p_l(T) + p_h(T) = 1, \tag{1}$$

that is, at every period there are 3 possible outcomes: no catastrophic event, an event with minor damage m, or an event with major damage h.

The probabilities of these events change with time. To model climate change, we assume the simple law of the probabilities' evolution:

$$\begin{aligned} &\vec{p}(T+1) = \vec{p}(T) + \vec{\alpha}(\vec{\rho} - \vec{p}(T)); \\ &\rho_0 + \rho_l + \rho_h = 1,\ \forall T \geq 0 : \ p_l(T) + p_h(T) + p_0(T) = 1, \end{aligned} \tag{2}$$

with ρ being the stationary value of probability for each of the events and $\vec{\alpha}$ governing the rate of change of probabilities of events.

In each period T, the governmental authority can implement a major or a minor abatement project to reduce the damages from potential catastrophic events, or it may choose to do nothing. We thus denote the action space of the government by:

$$\mathcal{A}_{1,2} = \{0, m, A\} \tag{3}$$

with 0 denoting the action of doing nothing, m being a minor prevention measure, and A a major project. The minor project takes effect in the same period, but has to be renewed every period. It can prevent only the low type event and does not reduce damages from the high type event. The major project requires time for completion, having an effect from the period next after construction is started, but prevents both types of events for the rest of the time. Projects are mutually exclusive, so only one of the actions from $\mathcal{A}$ can be undertaken in each period.

We assume that the costs of a major project are higher than the cost of a minor project:

$$C(A) > C(m) \tag{4}$$

We further assume that the high type and low type events may cause different damages for different types of sites, with

$$D_H(h) > D_H(l), D_L(h) > D_L(l),\ D_H(h) > D_L(h), D_H(l) > D_L(l) \tag{5}$$

where D_i denotes damages from event type $j \in \{l, h\}$ at site $i \in \{L, H\}$, that is, high type event causes more damage for every type of location than the low one, and every type of event causes higher damages to the more susceptible site than to the less susceptible one.

These assumptions characterize a setting that is typical for hydrological disasters, such as floodings. If a flood occurs, it causes costs at different locations that have different vulnerability (e.g., being more or less elevated compared to a river). A smaller flooding can be prevented by using simple measures, such as stapling sand bags, but such measures fail in case of more catastrophic disasters. A high-level disaster might be avoided by larger adaptation measures, such as building a dam, however, this takes considerable time (it cannot be done, once it is clear that a flood will occur).

In the next sections, we analyze the influence of the availability of information regarding the true probability distribution (2) on the optimal timing of the implementation of adaptation projects.

3 Benchmark case with known probabilities

As a benchmark, we first consider the case where the planner knows both initial probabilities and the rate of their increase. In this case, the problem reduces to a simple multi-period minimization of expected costs:

$$\min \mathbf{S} = \sum_{T=0}^{\infty} \frac{1}{(1+\delta)^T} \left(\mathbb{E}(D(T)) + C(T) \right), \tag{6}$$

where $\delta < 1$ is the social discount factor, $D(T)$ denotes total damages on both sites at period T (which in turn could vary in size and across sites) and $\forall T \in \mathbb{N}: \ C(T) = \{C(A), C(m), C(0)\}$ are adaptation costs (which are zero if no action is taken at T).

Assume for certainty that one site is of the low type (small susceptibility to damage) and the other of the high type. If no project at any site is implemented in any period, expected damages are:

$$\mathbf{S}_{\mathcal{A}=\{0,0\}} = \sum_{T=0}^{\infty} \frac{1}{(1+\delta)^T} \left\{ p_h(T)(D_H(h) + D_L(h)) + p_l(T)(D_H(l) + D_L(l)) \right\} \tag{7}$$

If the planner implements project A at one of the sites at time 0, it prevents major and minor damages altogether starting from the next period for both sites:

$$\mathbf{S}_{\mathcal{A}=\{A,0\}} = p_h(0)(D_H(h) + D_L(h)) + p_l(0)(D_H(l) + D_L(l)) + C(A), \tag{8}$$

If the minor project is implemented on both sites, it prevents only minor event at period 1 and

$$\mathbf{S}_{\mathcal{A}=\{m,m\}} = 2C(m) + p_h(0)(D_H(h) + D_L(h)) + \\ \sum_{T=1}^{\infty} \frac{1}{(1+\delta)^T} \left\{ p_h(T)(D_H(h) + D_L(h)) + p_l(T)(D_H(l) + D_L(l)) \right\}. \tag{9}$$

Thus, it is always optimal for the social planner to implement project A in the first period, if the distribution of types is $\{H, L\}$ (or $\{H, H\}$) and the probability of a catastrophic event is positive from the beginning and increasing in time.

Now assume that both sites have low vulnerability, that is, we have $\{L, L\}$ as types. In this case, it is not optimal to implement any project at time 0. But as the probabilities change, it might be the case that a project (minor or major) becomes optimal later on. First, consider a minor project. In (2) probabilities changes are monotonic. Thus if it is better to implement a minor project at any of the sites than doing nothing at time $T = s$, it is also better in all subsequent periods $T > s$. The time at which it starts being worthwhile to implement a minor project, can be found on the per period basis, as the minor project has to be renewed every period:

$$s_{0,m}(i): \; C(m) = p_l(s_{0,m})D_i(l) \tag{10}$$

As by assumption the planner knows true probabilities states, the difference equations (2) may be solved to yield

$$\vec{p}(T) = \vec{p}(0)(1-\vec{\alpha})^T - \vec{\rho}(1-\vec{\alpha})^T + \vec{\rho} \tag{11}$$

and the time, when minor project is worth implementing at site i is:

$$[s^*_{0,m}(i)] = \frac{\ln\left(\frac{D_i(l)\rho_l - C(m)}{D_i(l)(\rho_l - p_l(0))}\right)}{\ln(1-\alpha_l)}, \tag{12}$$

where ρ_l, α_l are parameters of probability of the low damage event, superscript $*$ denotes the benchmark case solution and half-bold square bracket means taking the floor (minimal integer) from the resulting value, since the discrete time setting.

Similarly, a major project A may be delayed, if it does not pay off at period 0. The time when a major project starts to dominate the implementation of minor projects is given by[1]:

$$s_{m,A}: \; \frac{1}{(1+\delta)^s}\left(C(A) + p_h(s)(D_H(h) + D_L(h)) + p_l(s)(D_H(l) + D_L(l))\right) =$$
$$\sum_{T=s}^{\infty} \frac{1}{(1+\delta)^T}\left(2C(m) + p_h(T)(D_H(h) + D_L(h))\right). \tag{13}$$

[1]observe that $s_{m,A}$ is not a function of location i since the planner is indifferent og where to implement a major project

There is a unique solution to (13), given by

$$[s^*_{m,A}] = \frac{\ln\left(-\frac{(\alpha_h+\delta)((2\,D_h\,\rho_h+C(A)-2\,C(m))\delta+D_h\,\rho-2\,C(m))}{(p_h(0)-\rho_h)(\alpha_h+2\,\delta+1)D_h\,\delta}\right)}{\ln(1-\alpha_h)}, \tag{14}$$

where $D_h = D_H(h) + D_L(h)$ and where α_h, ρ_h are the parameters of the probability of the high damage event.

Finally, we have to assess when doing nothing is dominated by the major project. For this we compare expected damages from time s onwards:

$$s_{0,A} : \frac{1}{(1+\delta)^s}\left(C(A) + p_h(s)(D_H(h) + D_L(h)) + p_l(s)(D_H(l) + D_L(l))\right) =$$
$$\sum_{T=s}^{\infty} \frac{1}{(1+\delta)^T}\left(p_l(T)(D_H(l) + D_L(l)) + p_h(T)(D_H(h) + D_L(h))\right) \tag{15}$$

As this equation contains both minor and major events probabilities, the time $s_{0,A}$ cannot be found explicitly. However, for every set of parameters, the time $s_{0,A}$ is well defined by an implicit function. To fully define the socially optimal program under known probabilities evolution we need to know the values of $s_{0,A}, s_{m,A}, s_{0,m}(i)$, defined above.

Assume first that all of them are real, positive and finite. Then the plan of actions is defined by the ordering of these numbers: as soon as, for example, $s_{0,A} < s_{0,m} < s_{m,A}$, only the major project is implemented at one of the sites at period $s_{0,A}$, while if $s_{0,m} < s_{0,A}$ first the minor measures are implemented starting with period $s_{0,m}$ and the switch to the major project is done at $s_{m,A}$ which in this case is lower than the $s_{0,A}$ number.

If some of the values are real and negative, this implies the associated policy has to be implemented starting from period 0 without the delay. As a tie-breaking rule we assume that if both $s_{0,A}, s_{0,m}$ are negative, the major project is implemented immediately.

Finally, if some of the values are not real, there does not exist the time, starting from which the given policy is better than the other. If for example, $\nexists s_{m,A} : \Im(s_{m,A}) = 0$, there is no gain in switching from minor project to the major one at any time.

We summarize these observations in the following proposition.

Proposition 1 (Benchmark social planner solution)**.**
Under the full information about the probabilities with which disasters occur, we have:

1. *As long as* $s^*_{0,A}, s^*_{m,A}, s^*_{0,m}(i)$ *exist and are nonnegative:*

   - *At time* $s^*_{0,m}(H)$ *the minor prevention measures start at the higher risk site, and later, at* $s^*_{0,m}(L) \geq s^*_{0,m}(H)$*, at the low risk site;*
   - *If the ordering* $s^*_{0,A} \geq s^*_{m,A} \geq s^*_{0,m}(L) \geq s^*_{0,m}(H) \geq 0$ *holds, then at the times* $s^*_{0,m}(i)$*, only minor measures start and the major project is constructed at one of the sites at time* $s^*_{m,A}$*;*
   - *If* $0 \leq s^*_{0,A} < s^*_{0,m}(i)$*, no minor measures are undertaken and at time* $s^*_{0,A}$*, a major project is constructed at one of the sites;*

2. *If any of the quadruple* $s^*_{0,A}, s^*_{m,A}, s^*_{0,m}(i)$ *is real but negative, the associated project is implemented from period* 0*. If both* $s^*_{0,m}(i), s^*_{0,A}$ *are negative, the major project is implemented from period 0.*

3. *If any of the quadruple has a non-zero complex part, the associated project is never implemented.*

We thus see that even without any higher-level uncertainty, the introduction of dynamically changing probabilities is sufficient to yield rather complex optimal adaptation strategies. Depending on the rates of increase and stationary values of the probability vector, it can be optimal to wait with the implementation of a major project, or even to wait some time, then implement only minor measures at the higher risk site and switch to the major project even later on.

4 One layer uncertainty: Known initial and final probabilities

As a second step, we assume that the social planner does not know the true current state of the probability vector, but knows only initial and final probabilities and the structure of the law that governs the rate of probabilities changes.

This means, parameters $\vec{p}(0), \vec{\rho}$ are known, but $\vec{\alpha}$ is unknown to the planner in (2). The planner has a distribution of subjective beliefs over the values of $\vec{\alpha}$ from which the subjective probability of each event at time T is drawn. Then using (11) the subjective probability at time T is:

$$\vec{\theta}(T) = \vec{p}(0)(1-\vec{\xi})^T - \vec{\rho}(1-\vec{\xi})^T + \vec{\rho},$$
$$\theta_l(T) + \theta_h(T) + \theta_0(T) = 1,\ \vec{\theta}(0) = \vec{p}(0), \tag{16}$$

where $\vec{\theta}(T)$ is the set of subjective probabilities at time T and $\vec{\xi}$ is the planner's current set of beliefs upon the value of correction coefficients in the probability law (2).

To complete the model we also need the rule for the update of planner's beliefs. We assume that if neither minor nor major events occur before period T, the set of beliefs remains unchanged. As soon as a major or minor event occurs, this leads to a decrease in uncertainty over the probabilities increase coefficients $\vec{\alpha}$.

4.1 Perfect learning case

As the simplest case, we first consider the situation when the arrival of one event of type $\{l, h\}$ is sufficient for the planner to learn the true value of the associated $\vec{\alpha}$[2]. In this case the updating rule is particularly simple:

$$p(\xi = \alpha) = 1,\ p(\xi = \alpha | \{l, h\}) = 0, \tag{17}$$

that is, the initial belief is that the planner knows the true α, but this is updated to the true value after observation of a single event. In this case, as long as no events occurred, the policy follows the same scheme as in Proposition 1, but with the times of introduction of projects being computed according to probabilities $\vec{\theta}$ rather than $\vec{p}$.

As soon as the information arrives (the event happens), the consequent problem is reduced to the benchmark case above, but starting from the period when the event has occurred. Provided no actions have been undertaken before this time e, the problem is fully equivalent to the one from Proposition 1.

We list the timing of projects' implementation here without derivations as those are fully equivalent to the benchmark case. The timing of the implementation of the minor

[2]such form of perfect learning is an oversimplification but is widely used in flood protection literature, see e. g. van der Pol et al. (2014)

project at site i is:

$$[s^{\theta}_{0,m}(i)] = \frac{\ln\left(\frac{D_i(l)\rho_l - C(m)}{D_i(l)(\rho_l - p_l(0))}\right)}{\ln(1-\xi_l)} \tag{18}$$

differing from (10) by subjective rate ξ_l only. The superscript θ denotes timing of projects under the subjective probabilities $\vec{\theta}$. The switch from minor measures to the major project is given by:

$$[s^{\theta}_{m,A}] = \frac{\ln\left(-\frac{(\xi_h+\delta)((2\,D_h\,\rho_h + C(A) - 2\,C(m))\delta + D_h\,\rho - 2\,C(m))}{(p_h(0)-\rho_h)(\xi_h + 2\,\delta + 1)D_h\,\delta}\right)}{\ln(1-\xi_h)} \tag{19}$$

and the time of major project implementation without prior minor prevention measures is given by the equation alike (15) but with $\xi_{l,h}$ instead $\alpha_{l,h}$.

Proposition 2 (Solution with perfect learning)**.**
In the perfect learning case the solution to the social planner's problem is:

1. *As long as no event occurs, the implementation of minor or major projects follows the scheme from Proposition 1 with subjective values $\vec{\theta}, \vec{\xi}$. Timing rules are denoted $s^{\theta}_{j,j}$;*

2. *As soon as one of events $\{l, h\}$ occurs at time $e \geq 0$, the policy is switched to the equivalent of Proposition 1, provided $e < s^{\theta}_{j,j}$ and timing rules are $s^{e}_{j,j} = e + s^{*}_{j,j}$;*

3. *If the event occurs after major project is implemented, the policy rule is the same as in 1.;*

4. *If the event occurs after only one or two minor projects are implemented, the implementation of major project follows the timing rule $s^{e}_{m,A} = e + s^{*}_{m,A}$.*

4.2 Imperfect learning case

As a contrasting case, we assume that the observation of one catastrophic event is not enough to achieve the full certainty over the $\vec{\alpha}$. On the contrary, the planner can never be sure that her estimate is giving the true vector, but every observation of a catastrophic event reduces the risk of an estimation error.

We assume the planner uses the following Bayesian updating mechanics: As soon as an event arrives, the planner learns the probability of the event at that time, $p(e)$, but not the rate of change of the probability, α. Then the subjective probability evolution will always be different from the true one, no matter how much events will arrive, but with every event it comes closer to the true one. The concept is illustrated by Figure 1.

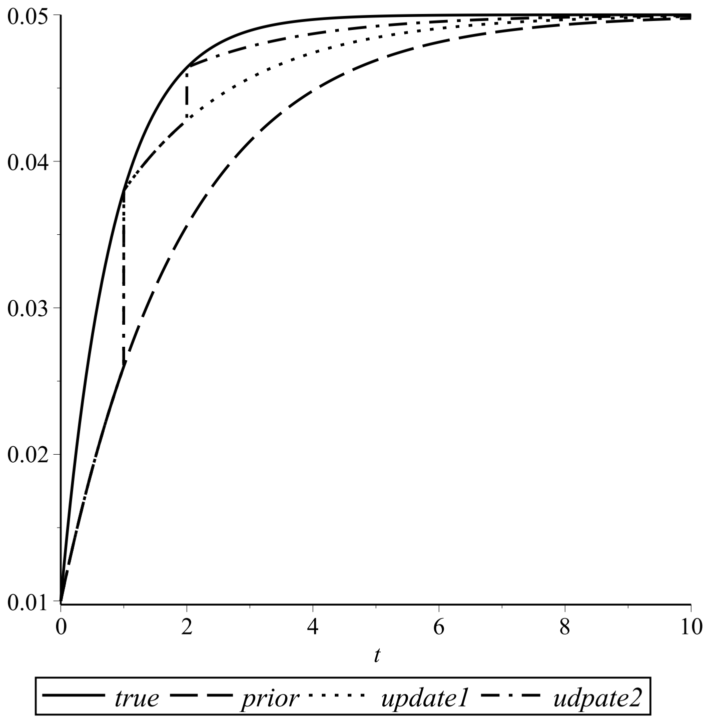


Figure 1: Convergence of subjective probabilities to the true one

To obtain a solution of this decision problem, we start at the point of Proposition 2: as soon as no event occurs, the implementation follows the same scheme as under perfect learning with subjective probabilities $\vec{\theta}$. As soon as one of the events occurs, the probability rule is updated, and error costs go down. In contrast to the perfect learning case, there is an option to wait until the next event to gain more information. Still qualitatively the solution and timing rules are similar to the perfect learning case.

In particular, if k is the number of events observed before the implementation and T_k is the period when the last of catastrophic events has been observed, the subjective probability vector is given by

$$\vec{\theta}_k(T) = \vec{p}(T_k)(1 - \vec{\xi})^{T-T_k} - \vec{\rho}(1 - \vec{\xi})^{T-T_k} + \vec{\rho} \tag{20}$$

and timing rules are obtained in the same way as before. We denote them by $s_{j,j}^{\theta_k}$.

Proposition 3 (Solution with imperfect learning)**.**
In case the planner is able to update his/her beliefs at every catastrophic event, the following timing rules of projects implementation hold:

1. *As long as no event (major or minor) occurs, the timing is given by* $s_{j,j}^{\theta}$*;*
2. *As soon as any number of events occur before the implementation, the timing is updated to* $s_{j,j}^{\theta_k}$ *after time* T_k*. As long as* $s_{j,j}^{\theta} > T_1$*, it follows that* $s_{j,j}^{\theta_k} < s_{j,j}^{\theta}$*.*

Thus the inclusion of one-parameter uncertainty regarding the probabilities of disasters into the baseline model does not change the results qualitatively.

5 Two-layered uncertainty

So far, we have used simple types of uncertainty, where the social planner still has substantial knowledge regarding the likelihood of disasters and how it changes over time, due to climate change. In these cases, learning has only limited value.

In this section, we consider more sophisticated cases, where the social planner has to learn the probabilities to a substantial extent from observing disasters. Given that climate change alters the distribution of extreme events, this is a plausible scenario, as old (pre-climate change) knowledge regarding the likelihood of disasters becomes unreliable.

5.1 Bayesian updating with unique static distribution of probabilities

First assume that the social planner does not know the probability distribution of p_h, p_l, p_0 but believes that it is static. This is a plausible future scenario, where some climate change has occurred, altering the likelihood of catastrophic events, but these impacts have been stabilized by mitigation measures.

In this case, the Bayesian updating rule with the use of conjugate prior can be described by the categorical distribution $\mathbf{Cat}(k, \vec{p})$ (generalized Bernoulli distribution for k possible outcomes):

$$f(x = i|p) = p_i, \sum_{i=1}^{k} p_i = 1. \tag{21}$$

Using more compact notation, the probability mass function is

$$f(x|p) = \prod_{i=1}^{k} p_i^{[x=i]}, \tag{22}$$

where $[x = i]$ is the Iverson bracket evaluating to 1 if $x = i$ and 0 otherwise. The conjugate prior distribution for the categorical one is the Dirichlet distribution $\mathbf{Dir}(k, \vec{\beta})$ with probability mass defined as:

$$f(x_i, \beta_i) = \frac{1}{\mathbf{B}(\vec{\beta})} \prod_{i=1}^{k} x_i^{\beta_i - 1}, \tag{23}$$

where $\mathbf{B}(\vec{\beta})$ is the normalizing constant given by the mutlinomial Beta-function:

$$\mathbf{B}(\vec{\beta}) = \frac{\prod_{i=1}^{k} \Gamma(\beta_i)}{\Gamma(\sum_{i=1}^{k} \beta_i)}, \Gamma(n) = (n-1)!\,. \tag{24}$$

The Bayesian updating rule for probabilities vector θ is

$$p(\theta|\mathbb{X}, \vec{\beta}) = \frac{p(\mathbb{X}|\theta) p(\theta|\vec{\beta})}{p(\mathbb{X}|\vec{\beta})}, \tag{25}$$

with $p(x)$ given by categorical distribution $\mathbf{Cat}(k, \vec{p})$ (22), $p(\theta|\vec{\beta})$, the prior distribution, given by our conjugate prior $\mathbf{Dir}(k, \vec{\beta})$ in (23), $p(\mathbb{X}|\theta)$ being the likelihood function, defined by the choice of the prior distribution and $p(\mathbb{X}|\vec{\beta})$ being the marginal likelihood function given by the Dirichlet-multinomial distribution

$$p(\mathbb{X}|\vec{\beta}) = \frac{\Gamma(\sum_{i=1}^{k} \beta_i)}{\Gamma(N + \sum_{i=1}^{k} \beta_i)} \prod_{i=1}^{k} \frac{\Gamma(c_i + \beta_i)}{\Gamma(\beta_i)}. \tag{26}$$

This choice of a prior distribution yields particularly simple posterior predictive probabilities of the form:

$$p(\tilde{x} = i|\mathbb{X}, \vec{\beta}) = \frac{c_i + \beta_i}{N + \sum_{i=1}^{k} \beta_i} \tag{27}$$

In our case, we have $k = 3$, as only 3 outcomes are possible. The prior probability vector is

$$\vec{p} = \begin{pmatrix} p_h^B \\ p_l^B \\ p_0^B \end{pmatrix}, \text{ and } p_h^B + p_l^B + p_0^B = 1. \tag{28}$$

The observations sample is updated every period depending on how many events have been observed so far. We denote the number of events observed up to time T as $\#_T^h, \#_T^l, \#_T^0$: $\#_T^h + \#_T^l + \#_T^0 = T$ for high, low and no event types respectively. The predictive probability at period T:

$$\vec{p}^B(T+1) = \begin{pmatrix} \frac{\#_T^h + \beta_h(0)}{T + \sum_{i=1}^{3} \beta_i(0)} \\ \frac{\#_T^l + \beta_l(0)}{T + \sum_{i=1}^{3} \beta_i(0)} \\ \frac{\#_T^0 + \beta_0(0)}{T + \sum_{i=1}^{3} \beta_i(0)} \end{pmatrix}. \tag{29}$$

The update of hyperparameters value is:

$$\vec{\beta}(T) = \begin{pmatrix} \beta_h(T-1) + [x = h] \\ \beta_l(T-1) + [x = l] \\ \beta_0(T-1) + [x = 0] \end{pmatrix}, \tag{30}$$

where we again use the Iverson bracket notation.

Note that hyperparameters $\vec{\beta}$ are estimations of probabilities of events at time T (up to a normalization), dependant on the sampled data.

The vector $\vec{\beta}(0)$ represents the prior estimation of occurrences of all types of events. In our case, this can be set with usage of knowledge of previous occurrences of catastrophic events and not just in a uniform way, as in the usual procedure. Observe also, that the vector (29) gives the predictive probability for the next period only, while the probability to observe exactly n events in the category k till time T is given by a more complicated multinomial law.

Now we can calculate the optimal timing of implementation under Bayesian updating with a single static prior. For that, we use the same concept of expected loss minimization as before. The subjective probabilities of a social planner are different, depending on how much events have been observed. Thus we can estimate the optimal time to construct the project (minor or major).

For a minor project (to be renewed every period), the question is: what is the threshold value of subjective probability of a minor event that makes the implementation of a minor project at site i optimal? The answer is simple (using relationship (10)):

$$\hat{p}_l^B = \frac{C(m)}{D_i(l)} \implies s_{0,m}^B(i) = f(\hat{p}_l^B) \tag{31}$$

We do not explicitly derive the timing as a function of probability estimate, since it is of the very same form as in (12) and (18). However, under Bayesian updating, the probabilities are no longer monotonic and thus it cannot be stated that once it is optimal to implement project m at site i at time s, it is always optimal onwards.

The decision rule for the social planner under Bayesian updating is the following. Given the observed frequency of events up to time T, the planner chooses one of five possible actions $\{0, m(1), m(2), m(1)m(2), A\}$ in the following manner:

- As soon as the observed frequency of minor event at time T exceeds

$$\#_l(T) \geq (T + \sum_{i=1}^{3} \beta_i(0)) \frac{D_i(l)}{C(m)} - \beta_l(0), \tag{32}$$

  the minor project is implemented at site i.

- As soon as (32) holds and the observed frequency of the major event at time T exceeds

$$\begin{aligned} &\#_h(T) \geq \\ &\frac{D_L(l)\left((C(A)\delta - C(m))(T + \sum_{i=1}^{3} \beta_i(0)) - \beta_h(0)(D_L(h) + D_H(h))\right)}{D_L(l)(D_L(h) + D_H(h))} - \\ &\frac{D_H(l)C(m)(T + \sum_{i=1}^{3} \beta_i(0))}{D_L(l)(D_L(h) + D_H(h))}, \end{aligned} \tag{33}$$

  the major project A is implemented at one of the sites.

- If both (32) and (33) do not hold, the planner chooses to wait at least one more period and repeats the procedure in $T+1$.

5.2 Bayesian updating with competing scenarios

We now modify the above approach to allow the planner to have several competing scenarios over the true state of nature. However, we still assume that each of these scenarios is static.

We assume that the planner believes that observations of disasters could result from several different distributions and does not know which of these distributions is correct. Over time, observations will show which of the distributions is the most likely one; however, waiting for this knowledge could cause substantial costs.

In our model, we assume each event may have both low and high probabilities (this can be easily extended to any finite number of priors):

$$X_T \overset{H_0}{\sim} \mathbf{Cat}(p^0, k),\ X_T \overset{H_1}{\sim} \mathbf{Cat}(p^1, k). \tag{34}$$

As a prior distribution of beliefs, we assume an *uninformed* social planner, that is, both hypotheses have the same probability, $P(H_i) = 1/2$. After observing an event type $\{0, l, h\}$, those probabilities are updated according to the Bayes rule:

$$P(H_i|X_T) = \frac{p^i P(H_i|X_{T-1})}{\sum_{i=0}^{N} p^i P(H_i|X_{T_1})}, \tag{35}$$

with N being the number of hypotheses tested.

Depending on the sequence of realized events before T, the likelihood of one or another scenario increases. The exact values can be defined by event trees. We condense this with the following notation:

$$P_T(H_i|\#_h(T), \#_l(T), \#_0(T)) = \frac{(p_h^i)^{\#_h(T)}\,(p_l^i)^{\#_l(T)}\,(p_0^i)^{\#_0(T)}\,P_0(H_i)}{\sum_{i=0}^{N}(p_h^i)^{\#_h(T)}\,(p_l^i)^{\#_l(T)}\,(p_0^i)^{\#_0(T)}\,P_0(H_i)}, \tag{36}$$

with N denoting the number of hypotheses (scenarios) over the true probabilities of events. In the limit, it is possible to define the continuous set of such scenarios for all possible probabilities' values from 0 to 1.

At each period the social planner thus observes the event and modifies her believes in each of the scenarios. The expected social losses at time s are then given by a linear combination of losses under different scenarios with weights associated with probabilities of hypotheses:

$$\mathbf{S}(H_0, H_1, .., H_N) \stackrel{def}{=} \sum_{i=0}^{N} P_s(H_i)\mathbf{S}(\vec{p}^i), \tag{37}$$

where the weight of each scenario is computed via (36).

The predictive probability of the catastrophic event in this case depends on the weights assigned to different hypotheses and always lies in between the extreme values, giving for predictive probability of event k

$$\hat{p}_k^S \in [\min_{i \in N}\{p_k^i\}, \max_{i \in N}\{p_k^i\}]. \tag{38}$$

The efficiency of this decision rule is fully defined by the initial choice of probabilities of events in competing scenarios. As long as these are sufficiently diverse, the scenario-based decision rule would cover the probabilities estimates provided by Bayesian updating above. However the speed of convergence of the estimate will be different.

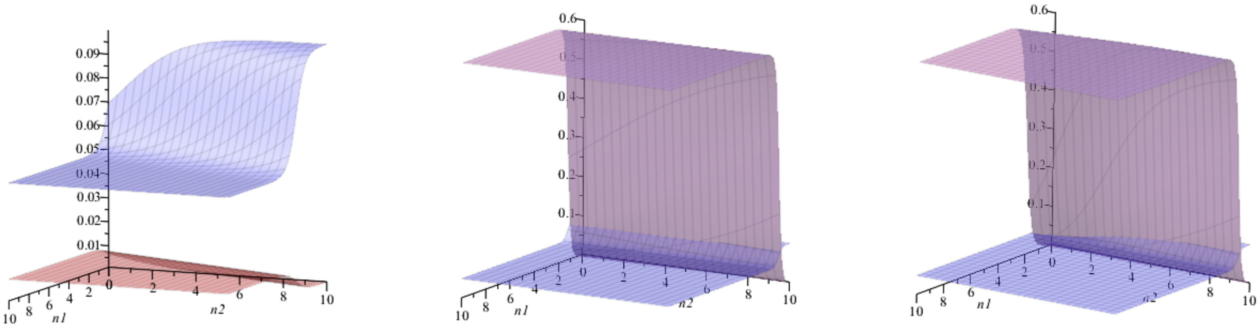
To illustrate this, consider a numerical example. Assume there are only two competing scenarios. We set the initial data as given in Table 1.

Probability	Hypotheses 0	Hypotheses 1	Hypotheses A
p_h	0.001	0.01	0.6
p_l	0.1	0.05	0.05
p_0	0.899	0.94	0.35

Table 1: Probabilities of events for competing scenarios

In the setting H_0, H_1, both scenarios are rather close in the chosen probability vectors, while choosing scenarios H_0, H_A gives rather diverse setup. We can also use all three scenarios simultaneously.

Figure 2 shows that the more diverse scenarios are taken into account, the more dynamic predictive probabilities can be. At the same time, there is only marginal value in



(a) Hypotheses 0&1 at T=10 (b) Hypotheses 0&A at T=10 (c) Hypotheses 0,1,A at T=10

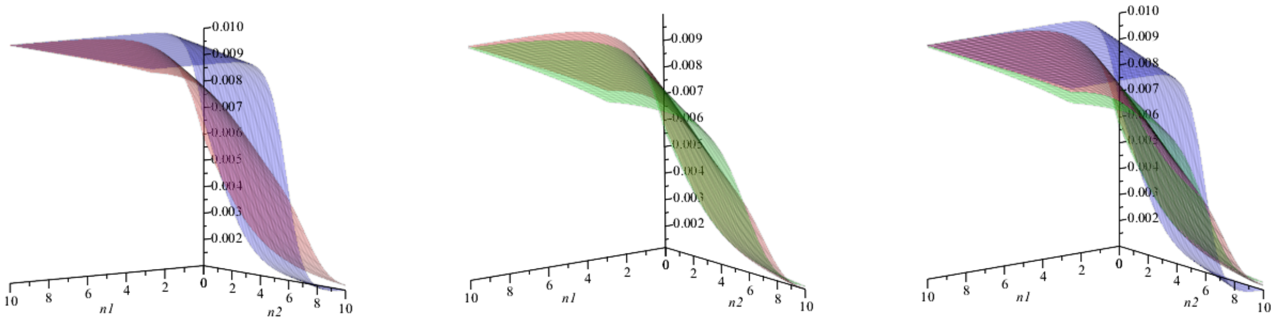
Figure 2: Probabilities of high (red) and low (blue) events as functions of number of events $(n1, n2)$ at $T = 10$

adding more scenarios. Consider the example where the additional scenarios are assigned probabilities in between the maximal and minimal values. In this case, the additional scenario will smooth the probabilities estimates and each next additional scenario will do so to a lesser extent. Figure 3 illustrates the concept for the transition from 2 to 3 and to 4 scenarios.

Formally the information gain from an additional scenario can be expressed in terms of prior probabilities assumed under existing scenarios. This can be positive or negative depending on the position of the new scenario within the existing spectrum. However the changes in resulting predictive probabilities follow the diminishing return: each next scenario adds a smaller learning boost to the predictive probabilities than all preceding ones, as this change is a linear combination of differences with existing scenarios.

Now consider the comparison with the unique Bayesian updating from Subsec. 5.1. Assume the initial prior probabilities for Bayesian update are set at the low level of H_0. The Table 2 gives predictive probabilities after 10 periods with different number of observed minor and major events.

From this table, we can see that the predictive probability under scenarios quickly converges to one of the scenarios, but cannot (by construction) exceed that level. Under



(a) Scenarios 0&1 at T=10 (b) Scenarios 1&2 at T=10 (c) Scenarios 0&1&2 at T=10

Figure 3: p_h under 2 (blue), 3 (red) and 4 (green) competing scenarios at $T = 10$

$\#_h(10)$	$\#_l(10)$	$\{\hat{p}_h, \hat{p}_l, \hat{p}_0\}$ under Bayes	$\{\hat{p}_h, \hat{p}_l, \hat{p}_0\}$ under H_0, H_1	$\{\hat{p}_h, \hat{p}_l, \hat{p}_0\}$ under H_0, H_A
1	1	$\{0.091, 0.1, 0.809\}$	$\{0.008, 0.056, 0.934\}$	$\{0.082, 0.093, 0.824\}$
2	2	$\{0.182, 0.19, 0.627\}$	$\{0.009, 0.051, 0.938\}$	$\{0.598, 0.05, 0.351\}$
3	3	$\{0.273, 0.282, 0.445\}$	$\{0.01, 0.05, 0.94\}$	$\{0.599, 0.05, 0.35\}$
1	4	$\{0.091, 0.373, 0.536\}$	$\{0.005, 0.08, 0.915\}$	$\{0.15, 0.087, 0.761\}$
4	1	$\{0.364, 0.1, 0.536\}$	$\{0.01, 0.05, 0.94\}$	$\{0.599, 0.05, 0.35\}$

Table 2: Predictive probabilities at T=10

Bayesian updating the probability is updated faster, but at the cost of a higher risk of overestimation.

This numerical example illustrates that, for given observed events up to time T, the procedure of Subsec. 5.1 yields a more dynamic update of the probabilities. This is not particular to this example, but a general feature of the procedure.

To see this, note that the scenario-based updating rule yields probability estimates that cannot possibly exceed the extremes set by the prior beliefs regarding the scenarios. In contrast, the Bayesian updating procedure is not limiting the posteriori predictions by initial beliefs. However, if the set of scenarios is rich enough (as in the example of H_0, H_A), the scenario procedure can induce higher estimates of disaster probabilities.

As the decision on implementation of both projects is based on the threshold levels of probabilities, the usual Bayesian updating can thus result in a faster or a slower implementation of a project, depending on the diversity of scenarios included into the (35) decision rule. Denote the difference in priors for scenario rule by the vector Δ:

$$\vec{\Delta} \stackrel{def}{=} \max_{i,j \in S} \begin{pmatrix} p_h^i - p_h^j \\ p_l^i - p_l^j \\ p_0^i - p_0^j \end{pmatrix}, \tag{39}$$

with S being the set of included scenarios. The speed of the probability updates in the scenario rule depends on the level of this Δ in the following way:

$$\hat{p}_k^S(T+1) - \hat{p}_k^S(T) = \sum_{i=1}^{N} p_k^i(0) \left(P_{T+1}^i(\Delta) - P_T^i(\Delta) \right), \tag{40}$$

where $P_T^i(\Delta)$ is the weight (probability) of scenario i at T, given by (35).

In contrast, the speed of probability updates in the unique Bayes rule does not depend on Δ, but only on number of observed events:

$$\hat{p}_k^B(T+1) - \hat{p}_k^B(T) = \frac{(T+1)\#_k(T+1) - (T+2)\#_k(T) - p_k(0)}{(T+1)(T+2)} \tag{41}$$

Comparing (40) and (41) yields the following result:

Proposition 4. *For any $N \leq \infty$ there exist an increasing sequence of Δ_s such that:*

1. *For $\Delta < \Delta_1$, the Bayesian procedure* (29) *yields a faster update of the predictive probabilities with regard to the number of observed events;*

2. *For $\Delta_2 > \Delta > \Delta_1$, the scenario-based rule* (35) *yields a faster update of probabilities;*

3. *For each next $\Delta_{s+1} > \Delta > \Delta_s$, the relationship reverts in sign*

The sequence Δ_s is finite and is given by the roots of the equation

$$\hat{p}_k^S(T+1) - \hat{p}_k^S(T) = \hat{p}_k^B(T+1) - \hat{p}_k^B(T). \tag{42}$$

Proof. First observe that the equation (42) is a polynomial in Δ and contains only powers $T, \#_k(T)$ that are natural numbers. Thus it has finitely many real roots.

Ordering these (real-valued) roots in an increasing sequence yields Δ_s. Comparing (40) with (41) for every difference in scenarios' priors shows that, for a small diversity of scenarios, the update is faster under the unique Bayes rule. By the property of polynomial roots, it changes its sign at every interval between roots, hence claims 2. and 3. follow. □

This Proposition tells us that there is no single optimal choice of the diversity of initial priors and quantity of scenarios. Rather, there always exist a range of initial beliefs, where unique Bayes rule yields faster update and vice versa.

As an illustration, consider the case of two scenarios with initial priors as of H_0 above and H_A with $\Delta = \delta$ being the difference in initial probability of high type event. Figure 4 illustrates two threshold lines, δ_1, δ_2 such that the probabilities update is faster under Baeysian rule for $\delta < \delta_1$ and $\delta > \delta_2$.

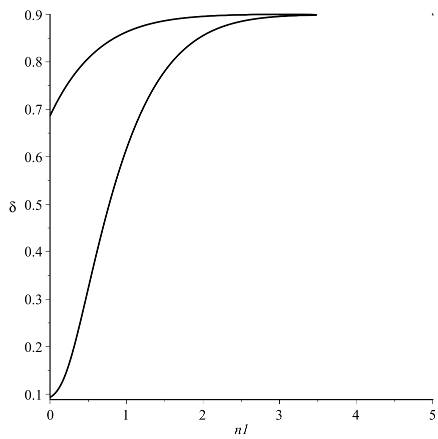


Figure 4: Thresholds of initial priors difference for two scenarios

At some number of observed major events, those two threshold coincide, yielding faster updating for the unique Bayes rule after this number of observations.

5.3 Two-layers Bayesian updating

Finally, we assume that the social planner uses the information about events both to update her belief in scenarios H_0, H_1 and to update the estimated probability vectors.

This leads to a combination of the decision rules defined in two previous subsection and is done rather easily. At some point probabilities under both priors would converge to a single vector, and the scenario which describes this vector is chosen as a governing one, yielding the scenario of Subsec. 5.1. We are interested in how much better the social planner performs with this more complex updating structure versus the simple one described above.

The updating procedure now is described by the Dirichlet mixture of priors[3]. In what follows, we closely follow the approach of Ye et al. (2011) and related papers:

$$P_T(H_i|\#_h(T), \#_l(T), \#_0(T)) = \frac{(p_h^i(T))^{\#_h(T)}(p_l^i(T))^{\#_l(T)}(p_0^i(T))^{\#_0(T)} P_0(H_i)}{\sum_{i=1}^N (p_h^i(T))^{\#_h(T)}(p_l^i(T))^{\#_l(T)}(p_0^i(T))^{\#_0(T)} P_0(H_i)}, \tag{43}$$

with

$$p_j^i(T) = \frac{\#_T^j + p_j^i(0)}{T+1}. \tag{44}$$

The predictive probability (taking major event for certainty):

$$\hat{p}_h^D(T) = P_T(H_0)\hat{p}_h^0(T) + P_T(H_1)\hat{p}_h^1(T) = \frac{(p_h^0(T))^{\#_h(T)+1}(p_l^0(T))^{\#_l(T)}(p_0^0(T))^{\#_0(T)} P_0(H_0)}{(p_h^0(T))^{\#_h(T)}(p_l^0(T))^{\#_l(T)}(p_0^0(T))^{\#_0(T)} P_0(H_0) + (p_h^1(T))^{\#_h(T)}(p_l^1(T))^{\#_l(T)}(p_0^1(T))^{\#_0(T)} P_0(H_1)} + \frac{(p_h^1(T))^{\#_h(T)+1}(p_l^1(T))^{\#_l(T)}(p_0^1(T))^{\#_0(T)} P_0(H_1)}{(p_h^0(T))^{\#_h(T)}(p_l^0(T))^{\#_l(T)}(p_0^0(T))^{\#_0(T)} P_0(H_0) + (p_h^1(T))^{\#_h(T)}(p_l^1(T))^{\#_l(T)}(p_0^1(T))^{\#_0(T)} P_0(H_1)} \tag{45}$$

where superscript D denotes Dirichlet mixture computation procedure.

The number of events, necessary to make major project implementation under the more complicated rule is (for the case of $N = 2$):

$$\#_h(T) \geq P_T(H_0)\#_h^0(T) + P_T(H_1)\#_h^1(T), \tag{46}$$

[3]Dirichlet mixtures is the tool extensively used in Mathematical Biology and related areas. Some overview may be found in Marin and Robert (2007)

where superscripts $0, 1$ indicate the threshold number of major events observed that is necessary for a major project implementation under scenarios $0, 1$. As the sum of probabilities of both scenarios cannot exceed one, the threshold probability giving the decision to implement a project is always a linear combination of probabilities under the two different priors.

The changes in predictive probabilities over time:

$$\hat{p}_h^D(T+1) - \hat{p}_h^D(T) = \big(P_{T+1}(H_0)\hat{p}_h^0(T+1) - P_T(H_0)\hat{p}_h^0(T)\big) + \big(P_{T+1}(H_1)\hat{p}_h^1(T+1) - P_T(H_1)\hat{p}_h^1(T)\big) \tag{47}$$

combine changes due to unique Bayes updating, (29) and scenario-based rule, (35).

It thus follows that the speed of such an update lies in between the two previous rules. In particular, it comes closer to unique Bayes updating, once this is changing faster than the scenario procedure and vice versa. This Dirichlet mixture procedure thus helps to improve upon the shortcomings of both previous procedures: it allows for predictive probabilities to change over time without scenario spectrum limitations, and at the same time it allows for the diversity of priors.

In case of close initial priors (small Δ) this procedure grants faster probabilities update, since initial priors do not play that much role as in the scenario-based rule. In case of extremely diverse priors (high Δ value) it yields more conservative estimate, since weights of scenarios are not updated that fast, and approaches unique Bayes updating. We thus observe:

Corollary 1. *The Dirichlet mixture procedure defined by* (43) *grants predictive probabilities convergence speed in between the unique Bayes rule* (29) *and scenario-based rule* (35)*:*

$$\hat{p}_k^D(T+1) - \hat{p}_k^D(T) \in [\hat{p}_k^S(T+1) - \hat{p}_k^S(T), \hat{p}_k^B(T+1) - \hat{p}_k^B(T)] \tag{48}$$

In particular, it holds for any Δ_s in the sequence defined by Prop. 4:

1. *For $\Delta > \Delta_s$: $\hat{p}_k^S(T+1) - \hat{p}_k^S(T) > \hat{p}_k^D(T+1) - \hat{p}_k^D(T) \to \hat{p}_k^B(T+1) - \hat{p}_k^B(T)$*
2. *For $\Delta < \Delta_s$: $\hat{p}_k^B(T+1) - \hat{p}_k^B(T) > \hat{p}_k^D(T+1) - \hat{p}_k^D(T) \to \hat{p}_k^S(T+1) - \hat{p}_k^S(T)$*

Proof. The first claim follows from the comparison of (40), (41) and (47): the latter combines changes of the former two and thus lies in between them.

Two other claims follows from Prop. 4: once an increasing sequence of Δ_s exists, it always holds either 1. or 2. The equivalent to (42) for Dirichlet mixture is again a polynomial equation with finitely many real roots which lie within intervals of changes of the polynomial (42). □

This last result demonstrates that once the planner is faced with a two-layered uncertainty, there is no need to choose either unique but dynamically updating belief or the set of pre-defined scenarios. In fact the planner may combine the two at relatively small additional computational costs.

Once we assume risk-neutral authority, it follows that the Dirichlet procedure would be expected welfare-improving over both unique Bayes and scenario rule, since it provides ample flexibility to account for arbitrary number of dynamically changing scenarios. However, once the costs of research are taken into account, it might be the case that the simplest Bayes rule (29) may be preferred. We postpone this discussion to future extensions of the model.

5.4 Value of climate research

By this we denote the gain in reduction of damages achieved by the change in the prior probability distributions. Assume the additional climate research may yield better estimates of planner's priors $\vec{\beta}$ in unique updating case first. Assume this research costs ν for each step in the direction of "true" values. The distance from subjective prior to the true state may be defined as

$$\mathcal{D} = \frac{|\vec{\beta}(0) - \vec{p}(0)|}{|\vec{\beta}(0)|} \tag{49}$$

then the cost of climate research might be given by $\nu\mathcal{D}$ and the benefits are given by the change in timing from (32), (33) resulting from the update of priors.

The ultimate problem is, that within the Bayesian framework there is no true probability, so one cannot define $\vec{p}(0)$ other than as the long-run resulting estimate. This last is given by

$$p_i^{\infty} = \frac{\#_T^i}{T} \tag{50}$$

with $T \to \infty$ and i denoting the type of event. Given the categorical nature of the event at hand, this long-run estimate need not to be constant and is updated every period in a non-monotonic fashion, but with lower increments.

If the initial prior is chosen at exactly this value, no new information may be gained from observing future events. However, there is no way for the planner to know, whether the chosen prior is close to this value or not, since the long-run probability estimate is defined by future events. Still, if the prior is close to this estimate, the changes in probabilities from period to period will be small, indicating that new information has less value.

To see that, insert as a prior into (29) the vector (50). This will immediately yield the predictive probabilities

$$\vec{p}^{\infty}(T+1) = \begin{pmatrix} \frac{\#_T^h}{T} \\ \frac{\#_T^l}{T} \\ \frac{\#_T^0}{T} \end{pmatrix} \tag{51}$$

valid for any T, and this prediction cannot be further improved. Still, probabilities will change every period, since the number of observed events will change. The change in priors monotonically and linearly affects the resulting predictive probabilities, as the Figure 5 illustrates for the case of probability of major catastrophic event.

We thus infer that to reach the same probability level as given by the long-run estimate (50), the planner has to overestimate the probability of a catastrophic event. In case of underestimation the long-run predictive estimate cannot be achieved at all. In both other cases there is exactly one point, when the predictive probability based on subjective prior crosses the long-run estimate and diverges from it afterwards. Looking at Eqs. (32),(33) we can easily conclude that the more pessimistic is the initial prior, the lower number of observed events are required to implement the project (major or minor) resulting in sooner implementation time. Overly pessimistic prior may even result in immediate implementation. In this case the planner runs a risk of implementing too soon,

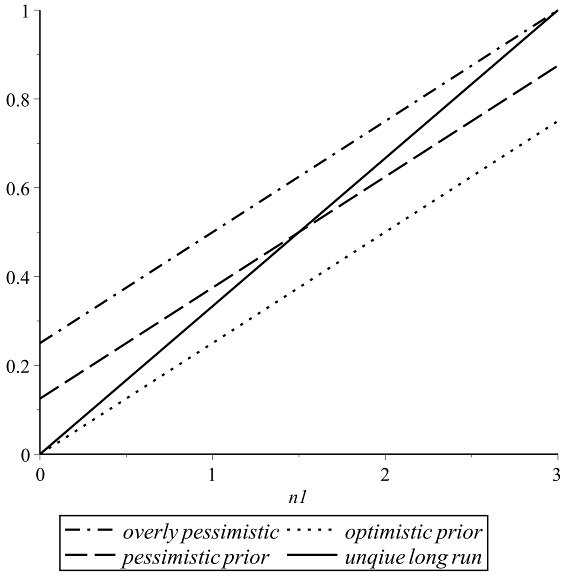


Figure 5: The effect of change in priors on unique Bayesian updates

since the prior probability estimate may decrease with time, rendering the immediate implementation suboptimal.

We conclude that additional climate research has the more value, the more extreme (optimistic or pessimistic) is the initial prior belief of the planner.

Observe that the same line of reasoning holds for *any* Bayesian updating rule from those discussed above. The value of climate research is the same for even the most complicated updating rule and is given by the difference between solid and dotted lines on the Figure 5 times expected damages both for minor and major events. As it was mentioned, the competing scenarios rule diversifies risks, putting weight to different levels of optimism, but the convergence of predictive probabilities to the long-run estimate (50) is slower, than for unique Bayesian updating. Thus the prior climate research may be seen as a substitute to more complicated updating rules, and the choice of which procedure to follow would depend on the comparison of climate research costs and costs of wrong decision (soon or late) upon implementation.

Corollary 2. *The value of additional climate research is the greater, the further is initial prior form the long-run estimate given by* (50)*. If time preferences of the planner are such* *that $\delta << 1$, the updating rule with diversified scenarios is preferable. If, on the other hand, $\delta \to 1$, the unique updating rule with additional climate research is preferable.*

5.5 Value of information

At last we concentrate on the question of irreversibility. The minor project has to be renewed every period, so as soon as the threshold predictive probability (31) is reached, the minor project is implemented under any Bayesian rule. However the decision to implement major project is irreversible. At the same time under Bayesian updating it is no longer the case that probabilities estimates are always increasing. It thus could be the case that once the threshold probability for major project is reached at T, it decreases below that level at $T + 1$ making the decision to implement project A at T suboptimal.

To study this situation we employ the Arrow-Fisher-Henry-Hanemann (AFHH) quasi-option approach, which measures the value of new potential information in irreversible setting. By definition the quasi-option value is the difference in value of an action given prior and posterior predictive probabilities. In the original simple setting this would mean the difference

$$\mathbb{E}(\min S_{\mathcal{A},0}) - \min \mathbb{E}(S_{\mathcal{A},0}) > 0 \tag{52}$$

and this is always positive. However in Bayesian setting the new information consists not in the learning the state of nature, but in the update of probabilities' values over periods.

In the model with Bayesian updating described throughout this section there are exactly three possible events at each period T. Thus the posterior predictive probability of a major event may have one of two values. For the case of unique Bayesian updating (29) it is given by:

$$\hat{p}_h(T+1) = \begin{cases} \frac{\beta_h(0)+\#_T^h}{\sum_{i=1}^3 \beta_i(0)+1+T}, [x \neq h]; \\ \frac{\beta_h(0)+\#_T^h+1}{\sum_{i=1}^3 \beta_i(0)+1+T}, [x = h] \end{cases} \tag{53}$$

The probability of high type event predicted at period T is given by $\hat{p}_h(T)$ in (29). Thus the expected change of predictive probability is

$$\hat{p}_h(T+1) - \hat{p}_h(T) = \begin{cases} -\frac{\beta_h(0)+\#_T^h}{(\sum_{i=1}^3 \beta_i(0)+T)(\sum_{i=1}^3 \beta_i(0)+1+T)}, [x \neq h]; \\ \frac{\sum_{i=1}^3 \beta_i(0)+T-\beta_h(0)-\#_T^h}{(\sum_{i=1}^3 \beta_i(0)+T)(\sum_{i=1}^3 \beta_i(0)+1+T)}, [x = h] \end{cases} \tag{54}$$

The value of new information in period $T+1$ given no project is implemented in period T is:

$$V^{AFHH}(T) = \mathbb{E}_T \left(S_{\mathcal{A}=\{A,0\}}(T+1)|\vec{\hat{p}}(T) \right) - \mathbb{E}_T \left(S_{\mathcal{A}=\{A,0\}}(T+1)|\vec{\hat{p}}(T+1) \right) \tag{55}$$

since S measures expected costs rather than value. There are three possible events arriving at $T+1$ and changing predictive probabilities in a way described by (53), thus there are three possible values of $V^{AFHH}(T)$.

Given the expression (8) it is straightforward to compute:

$$\begin{aligned} &\mathbb{E}_T \left(S_{\mathcal{A}=\{A,0\}}(T+1)|\vec{\hat{p}}(T) \right) - \mathbb{E}_T \left(S_{\mathcal{A}=\{A,0\}}(T+1)|\vec{\hat{p}}(T+1) \right) = \\ &\frac{1}{1+\delta} \left(\hat{p}_h(T)(D_H(h) + D_L(h)) + \hat{p}_l(T)(D_H(l) + D_L(l)) + C(A) \right) - \\ &\frac{1}{1+\delta} \left(\hat{p}_h(T+1)(D_H(h) + D_L(h)) + \hat{p}_l(T+1)(D_H(l) + D_L(l)) + C(A) \right) = \\ &\frac{1}{1+\delta} \left((D_H(h) + D_L(h))(\hat{p}_h(T) - \hat{p}_h(T+1)) + (D_H(l) + D_L(l))(\hat{p}_l(T) - \hat{p}_l(T+1)) \right) \end{aligned} \tag{56}$$

However values of posterior predictive probabilities are different given which of the events $0, l, h$ realize from T to $T+1$. Thus we arrive to three different values:

$$\mathbb{E}_T\left(S_{\mathcal{A}=\{A,0\}}(T+1)|\vec{\hat{p}}(T)\right) - \mathbb{E}_T\left(S_{\mathcal{A}=\{A,0\}}(T+1)|\vec{\hat{p}}(T+1)\right) \stackrel{E_{T+1}=h}{=}$$
$$\frac{1}{1+\delta}(D_H(h)+D_L(h))\frac{\beta_h(0)+\#_T^h-\sum_{i=1}^3\beta_i(0)-T}{(\sum_{i=1}^3\beta_i(0)+T)(\sum_{i=1}^3\beta_i(0)+1+T)}+$$
$$\frac{1}{1+\delta}(D_H(l)+D_L(l))\frac{\beta_l(0)+\#_T^l}{(\sum_{i=1}^3\beta_i(0)+T)(\sum_{i=1}^3\beta_i(0)+1+T)} \tag{57a}$$
$$\mathbb{E}_T\left(S_{\mathcal{A}=\{A,0\}}(T+1)|\vec{\hat{p}}(T)\right) - \mathbb{E}_T\left(S_{\mathcal{A}=\{A,0\}}(T+1)|\vec{\hat{p}}(T+1)\right) \stackrel{E_{T+1}=l}{=}$$
$$\frac{1}{1+\delta}(D_H(h)+D_L(h))\frac{\beta_h(0)+\#_T^h}{(\sum_{i=1}^3\beta_i(0)+T)(\sum_{i=1}^3\beta_i(0)+1+T)}+$$
$$\frac{1}{1+\delta}(D_H(l)+D_L(l))\frac{\beta_l(0)+\#_T^l-\sum_{i=1}^3\beta_i(0)-T}{(\sum_{i=1}^3\beta_i(0)+T)(\sum_{i=1}^3\beta_i(0)+1+T)} \tag{57b}$$
$$\mathbb{E}_T\left(S_{\mathcal{A}=\{A,0\}}(T+1)|\vec{\hat{p}}(T)\right) - \mathbb{E}_T\left(S_{\mathcal{A}=\{A,0\}}(T+1)|\vec{\hat{p}}(T+1)\right) \stackrel{E_{T+1}=0}{=}$$
$$\frac{1}{1+\delta}(D_H(h)+D_L(h))\frac{\beta_h(0)+\#_T^h}{(\sum_{i=1}^3\beta_i(0)+T)(\sum_{i=1}^3\beta_i(0)+1+T)}+$$
$$\frac{1}{1+\delta}(D_H(l)+D_L(l))\frac{\beta_l(0)+\#_T^l}{(\sum_{i=1}^3\beta_i(0)+T)(\sum_{i=1}^3\beta_i(0)+1+T)} \tag{57c}$$

The probability of each of the outcomes estimated at period T is given by $\vec{\hat{p}}(T)$, so the expected $V^{AFHH}(T)$ value is strictly zero since the linearity of damage function functional S. This comes with no surprise, since Bayesian updating rule uses all available information at any T, and AFHH value of waiting one more period amounts to *some* update of probability vector. Thus if the decision upon implementation of A is optimal at T, it is not optimal to wait longer.

The value to postpone investments into major project till next period is positive only if no major nor minor event occurs at T, since this drives predictive probabilities of catastrophic events down. Otherwise the postponement value is negative: if the probability $\hat{p}_h$ reaches a threshold level at T, the observation of one more event will drive this probability up, increasing the value of investing in project A at T rather than at $T+1$.

Now take into account the expected costs of information update and the reduction in costs of the project, $\hat{p}_h(T)(D_H(h)+D_L(h))+\hat{p}_l(T)(D_H(l)+D_L(hl))-C(A)\frac{\delta}{1+\delta}$. This gives together with Arrow-Fisher value the total postponement value, as defined in Mensink and Requate (2005). As soon as the threshold probability is reached the last expression is zero (it is optimal to invest now). Thus we conclude that the irreversibility of project A does not increase the time of learning per se. This does not mean the immediate implementation is optimal, since certain number of events have to be observed for the probability to reach the threshold value. At last observe that the same procedure as above is valid for scenarios-based updating and two-layers updating, but with more complicated derivations. We thus conclude with the proposition:

Proposition 5 (Effect of irreversibility of major project)**.**
Under Bayesian updating procedures (29)*,* (36)*,* (43) *the Arrow-Fisher quasi-option value implies the following:*

1. *As long as the planner is risk-neutral the irreversibility of a project is included into Bayesian estimates.*
2. *As soon as there is some degree of risk-aversion, the threshold probability decreases because of potential risks.*
3. *At the threshold level the expected costs of updating the information are exceeding the potential benefits even if no catastrophic event is expected in the near future.*

The total postponement value is zero once the threshold probability is reached and positive before that.

So we have demonstrated that optimal management under Bayesian updating rules includes the value of irreversibility. Observe that this is not the case with one layer uncertainty described in Sec.4.

6 Conclusion

In this paper, we have advanced and studied a simple model of adaptation to catastrophic events under uncertainty regarding the impact of climate change on the probability of these events. In particular, we have investigated different assumptions regarding the uncertainty of the impacts of climate change; do we know how fast probabilities of catastrophic events change, to which value they will adjust, or do we have to learn which of a set of climate impact scenarios is the true one by observing impacts.

Our model is inspired by the case of flood protection, where climate change is likely to require substantial adaptation measures. However, the general setup can be transferred easily to other cases of climate change impacts.

In the benchmark case of our model, where the social planner knows the true probabilities and their evolution, there is an optimal (and usually non-zero) delay for major adaptation projects, that is, it is usually optimal to wait for some time (depending on the rate of probabilities' increase) before starting with major adaptation measures.

If we allow for some uncertainty over the climate change probabilities, the outcome is not so clear. Under the perfect and imperfect learning described in Sec. 4, the timing of both minor and major projects can be defined in the same way as for the benchmark case. Depending on the initial beliefs of the planner over the evolution of nature, ex-post social damages may be higher or lower than in the benchmark case. Our results imply that (in this sense) it is optimal for the planner to err on the pessimistic scenario, as this involves a lower risk of investing too late (i.e., after the major catastrophic event already happened).

Finally, we investigated the optimal decision rule of the social planner for two-layered uncertainty. Under this assumption, the planner does not know at all the true probability and cannot learn it. The key factor determining the timing of projects is thus the number of events that have to be observed before adaptation becomes optimal. We demonstrate, that using a unique prior is not the optimal decision rule for the planner, and work out the procedure which combines initial priors of several scenarios and active learning over weights of each scenario and priors updating.

Altogether, our analysis shows that under the (realistic) assumption that climate change impacts cannot be foreseen with certainty, adaptation is a rather complex problem. In particular, the way how we handle climate change uncertainty has a substantial impact on optimal abatement strategies. Should we account for different scenarios, as the IPCC does, and strive to learn which of these scenarios holds true? Or should we rather assume that there is a single probability of climate impacts, which we can pin down more exactly over time? The results of this paper suggest that each of these approaches leads to different optimal adaptation strategies, as they imply different speeds of learning and different propensities to invest too early or too late.

References

Babonneau, F., A. Haurie, and M. Vielle (2013). A robust meta-game for climate negotiations. *Computational Management Science 10*(4), 299–329.

Buob, S. and G. Stephan (2011). To mitigate or to adapt: How to confront global climate change. *European Journal of Political Economy 27*(1), 1–16.

De Bruin, K. C. and R. B. Dellink (2011). How harmful are restrictions on adapting to climate change? *Global Environmental Change 21*(1), 34–45.

Kane, S. and J. F. Shogren (2000). Linking adaptation and mitigation in climate change policy. *Climatic Change 45*(1), 75–102.

Marin, J.-M. and C. P. Robert (2007). *Bayesian Core: A Practical Approach to Computational Bayesian Statistics*. Springer.

Mensink, P. and T. Requate (2005). The Dixit-Pindyck and the Arrow-Fisher-Hanemann-Henry option values are not equivalent: a note on Fisher (2000). *Resource and Energy Economics 27*(1), 83–88.

Nordhaus, W. (2007). *The challenge of global warming: economic models and environmental policy*, Volume 4. Yale University.

van der Pol, T., E. van Ierland, and H.-P. Weikard (2014). Optimal dike investments under uncertainty and learning about increasing water levels. *Journal of Flood Risk Management 7*(4), 308–318.

Ye, X., Y.-K. Yu, and S. F. Altschul (2011). On the inference of dirichlet mixture priors for protein sequence comparison. *Journal of Computational Biology 18*(8), 941–954.