Efficient avoidance of tipping points

A viable cost of carbon

Pierce Donovan

Department of Economics, Colgate University

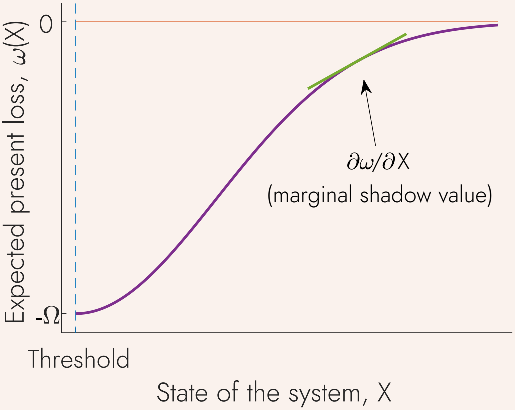
UC Davis NatuRE Policy Lab, January 2021

Threshold-driven management strategies are useful when we don't know the explicit benefits of avoiding disaster.

When we voluntarily impose constraints on ourselves, we are signaling that we want to protect something valuable.

1. Thresholds implicate danger
2. Threshold crossing events are irreversible/cannot be avoided with certainty
3. The damages associated with crossing the threshold are difficult to value with precision

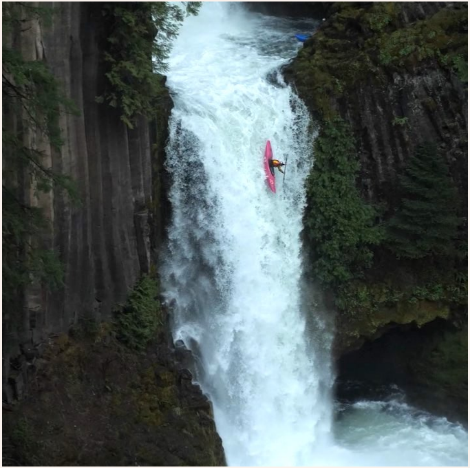
What motivates a disaster-avoider's actions in the day-to-day? Economic thinking would suggest translating the implied value of avoiding danger to the benefits of reducing risk on the margin.



(Ω : loss due to disaster)

What is the “viability problem?”

Imagine you are kayaking on a river. You have become stuck in a section above a massive waterfall. Going over would be terrible.



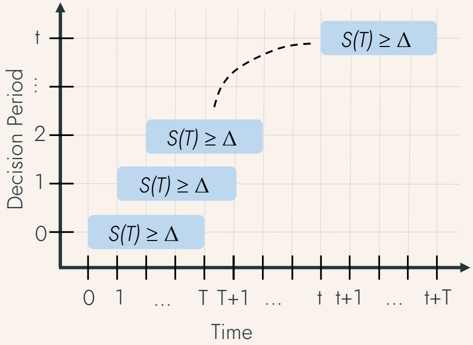
† Dane Jackson, *Toketee Falls* (Oregon, 2019)

The current is fickle—because it sometimes overwhelms even maximum paddling, long-run survival is not guaranteed.

†Dmitry Naumov, *Waikato River* (2019)

Viability goal: ensure avoidance of a threshold over window T w/ confidence Δ. Being continually concerned with a rolling window over an infinite horizon violates a useful Markov assumption.

$$S(T) = \Pr\left\{\bigcap_{\tau=t}^{t+T} (X_\tau > 0)\right\} \geq \Delta$$



Our standard dynamic programming problems without the viability constraint will often provide us no reason to ensure safety.

$$\underbrace{V(X_t)}_{\text{value function}} = \max_{A_t} \overbrace{\pi(A_t, X_t)}^{\text{flow value}} + \underbrace{\beta}_{\text{discount rate}} \cdot \overbrace{\mathbb{E}_\varepsilon \left[V(X_{t+1}) \mid A_t, X_t \right]}^{\text{continuation value}}$$

$$\text{s.t. } X_{t+1} = \overbrace{G(A_t, X_t, \varepsilon_t)}^{\text{growth } G, \text{ action } A, \text{ state } X, \text{ shock } \varepsilon} \qquad \text{(stochastic } (\varepsilon) \text{ dynamics)}$$

The dynamic programming problem of interest is intractable...

$$\underbrace{V(X_t)}_{\text{value function}} = \max_{A_t} \overbrace{\pi(A_t, X_t)}^{\text{flow value}} + \underbrace{\beta}_{\text{discount rate}} \cdot \overbrace{\mathbb{E}_\varepsilon \left[V(X_{t+1}) \mid A_t, X_t \right]}^{\text{continuation value}}$$

$$\text{s.t. } X_{t+1} = \overbrace{G(A_t, X_t, \varepsilon_t)}^{\text{growth } G\text{, action } A\text{, state } X\text{, shock } \varepsilon} \qquad \text{(stochastic } (\varepsilon) \text{ dynamics)}$$

$$\text{and } S(\underbrace{T}_{\text{rolling window}} \mid A_t, X_t) \geq \underbrace{\Delta}_{\text{confidence}} \qquad \text{(viability (survival) constraint)}$$

Could something else stand in for the viability constraint during optimization?

The dynamic programming problem of interest is intractable... *until* we add an additional constraint on the value function.

$$V^{\Omega}(X_t) = \max_{A_t} \pi(A_t, X_t) + \beta \cdot \mathbb{E}_{\varepsilon}\left[V^{\Omega}(X_{t+1}) \mid A_t, X_t\right]$$

s.t. $X_{t+1} = G(A_t, X_t, \varepsilon_t)$ (stochastic dynamics)

and $S(T \mid A_t, X_t) \geq \Delta$ (viability constraint)

and $V^{\Omega}(0) = -\Omega$ (extinction loss)

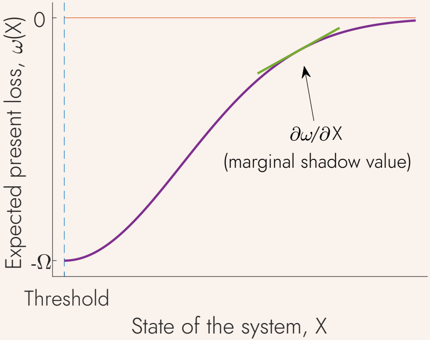
In each period we either don't encounter disaster (and receive flow π), or we do (and make a perpetuity loss payment).

$$V^{\Omega}(X_t) = \overbrace{\mathbb{E}_{T_f}\left[\sum_{\tau=t}^{t+T_f-1} \beta^{\tau-t} \cdot \pi(A^{\Omega}, X_{\tau}) \;\middle|\; X_t\right]}^{\text{typical stream of benefits/costs}}$$

$$\underbrace{- \mathbb{E}_{T_f}\left[\sum_{\tau=t+T_f}^{\infty} \beta^{\tau-t} \cdot ((1-\beta)\cdot\Omega) \;\middle|\; A^{\Omega}, X_t\right]}_{\text{expected present value of } \Omega \text{, I call this } \omega(X_t)}$$

$\omega(X_t)$ discounts the loss Ω by the expected time to/likelihood of disaster, conditional on starting state X_t and optimal policy.

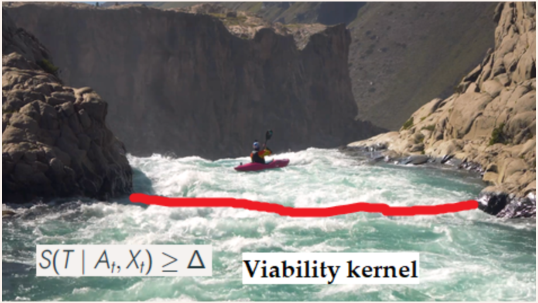
$$\omega(X_t) = \Omega \cdot \sum_{\tau=t}^{\infty} \beta^{\tau-t} \cdot \Pr\left(T_f = \tau \mid A^{\Omega}, X_t\right)$$



Jointly estimate Ω and the policy function $A^{\Omega}(X_t)$, and uncover $\omega(X_t)$.

How do we calibrate Ω?

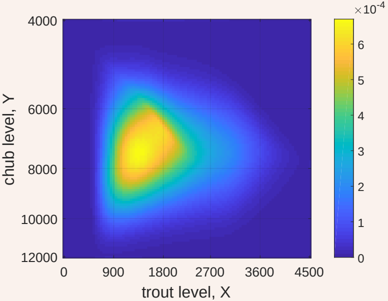
First we determine the *viability kernel*, the set $\{X\}_k$ where the viability constraint can be satisfied. Then we find the smallest loss Ω that motivates us to meet the constraint where it is possible.



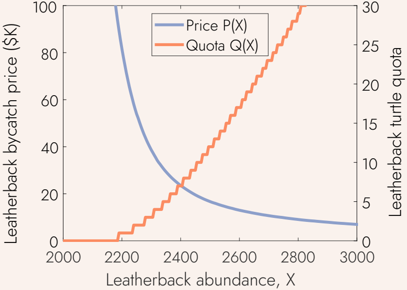
$$\min |\Omega| \text{ s.t. } S(T \mid \overbrace{A^\Omega}^{A^\Omega \text{ maximizes the value function } V^\Omega}, \{X\}_k) \geq \Delta$$

† Dane Jackson/NYT, *Salto Del Maule, Chile* (2020)
(Annotations by Michael Springborn)

“Past” applications of the SVV approach involve conservation on public lands and decentralized fisheries management.



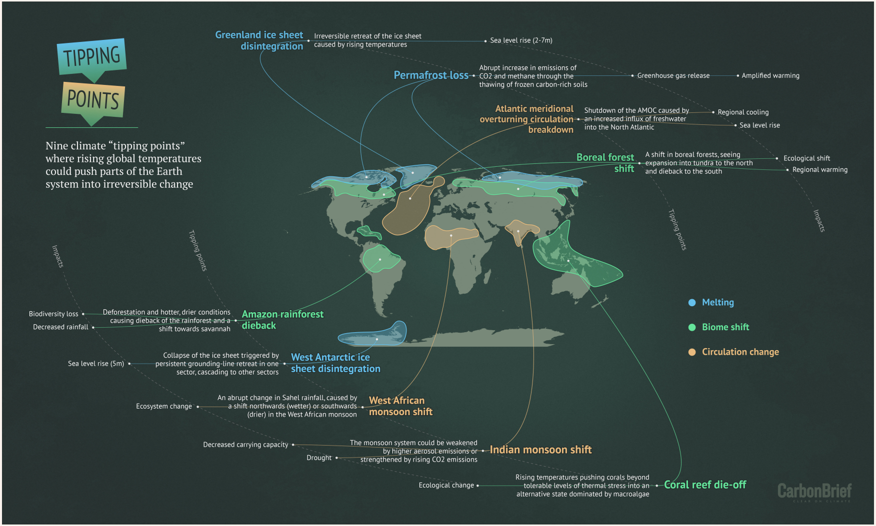
Managing endangered and non-native species in the Grand Canyon*



Setting the strength of regulatory instruments for bycatch viability†

*Donovan, Bair, Yackulic, and Springborn, *Safety in numbers* (2019)
†Donovan and Springborn, *Balancing conservation and commerce* (Working Paper)

Tipping points: thresholds with large, irreversible damages



†Rosamund Pearce/Tom Prater, *Tipping Points Explainer* (2020)

A threshold-driven policy would center importance on avoiding tipping points (the spirit of goals in the Paris Agreement).

But the climate change literature focuses instead on *maximizing the present value of the flow of societal welfare by controlling all global economic activity.*

Maybe they're right to do this:

- it's what we know
- we have to take into account opportunity costs of abatement
- can throw a temperature constraint on top

Integrated assessment models provide clear and coherent linkages between key dynamic variables of climate and the economy.

Assumes that future damages are measurable and predictable:

- welfare discounting (CC investments should fair as well as others)
- damage functions (some damages are much harder to put a number on than others)
- size and scope of industries/people considered (consequence: care less about areas with lower economic potential)

Keynes: *Better to be vaguely right than precisely wrong.* (???)

I'd like to provide another complementary tool [to IAMs] for pathing emissions/setting a carbon price.

Currently in a purely-physical framework: *maximize a metric of [reversible] global warming potential (energy absorbed by the Earth) while constraining the likelihood of irreversible climate change to some maximal tolerable level.*

This is social planning without attempting to prescribe the benefits of abatement—setting the cap on emissions for each year without a model of the economy.

The full model follows the SVV pattern.

$$V^{\Omega}(S_t) = \max_{E_t} \left\{ \pi(E_t, S_t) + \beta \cdot \mathbb{E}_{\varepsilon} \left[V^{\Omega}(S_{t+1}) \mid E_t, S_t \right] \right\}$$

$$\begin{aligned} \text{s.t.} \quad & S_{t+1} = G(E_t, S_t, \varepsilon_t) \\ & V^{\Omega}(Tip_t = 1) = -\Omega \\ & \Pr\{Tip_{t+T} = 0 \mid A^{\Omega}(S_t), \{S\}_k\} \geq \Delta \\ & S_t = \{Tip_t, CE_t, X_{A,t}, X_{O,t}, R_{n,t}\}, \ (n = 1...4) \end{aligned}$$

$\pi(\cdot)$ (global warming potential): total amount of energy absorbed by the Earth over some time horizon b/c of an increase in emissions today.

What does discounting (β) mean? (increased ability to mitigate?)

$\{T, \Delta\}$ reflects a social preference (assumed).

Determining the kernel $\{S\}_k$ is hard with complicated model.

One last piece—we don't exactly know where these tipping points are, or to what level we should limit temperature increases.

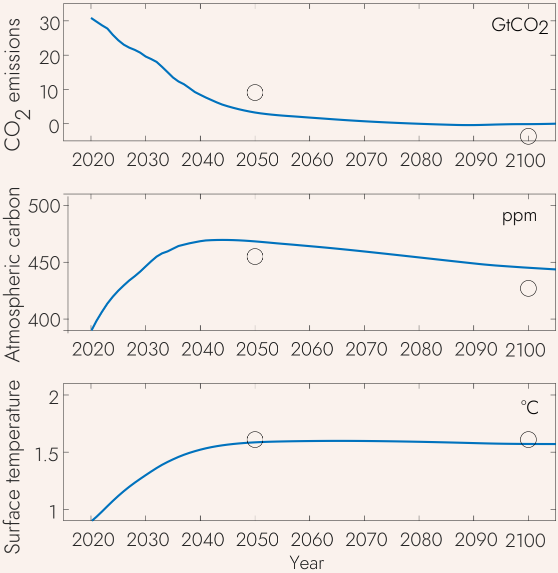
Subjective hazard function (probability of tipping):

$$H(Tip_{t+1}|X_{t+1}, X_t) = \max\left\{0, \frac{\min\{X_{t+1}, \overline{X}\} - X_t}{\overline{X} - X_t}\right\}$$

If we observe a higher temperature without tipping, then we know the true threshold must be higher than the current state of the world and the probability density function squeezes into a smaller, warmer domain.

Choosing $\overline{X}$ is a little strange... in paper, 5°C gives 21% chance of tipping by 2°C, given current state of the world.

Expected emissions path, given the system hasn't tipped:



Additional questions...

1. Is this pursuit pointless without considering abatement costs?
2. What other policy/value function stuff should I display? (lots of hidden outputs, like the size of the loss Ω)
3. How do I move from an emissions path to a carbon price? (SCC)

Thanks to global warming, the temperature will soon go through the roof, and we'll have an endless summer!

—Mr. Krabs (2005)

† Encyclopedia SpongeBobia, *Mr. Krabs* (2005)