THE DINOSAUR RISK
- A FABLE OF CIVILIZATIONAL COLLAPSE -

Kaushik Basu* and Jörgen Weibull†

June 6, 2019.‡

Abstract. This is a descriptive paper that shows how a game of life we are engaged in can slowly change, and even improve in the sense of the same actions resulting in higher payoffs as time progresses. However, while this society keeps growing, in terms of payoffs increasing over time, the slow drift of the underlying game takes it towards a collapse by changing the game's strategic character. This fable thus concerns our long-run survival and prompts us to do what dinosaurs were not capable of but we, in principle, are, which is to think about our own predicament and try to engineer changes in our own behavior and in market regulation to avert the risk of our own extinction.

JEL-codes: A13, D60, H0, H41, I31, O44, Q01, Q50.

Keywords: dinosaur risk, climate change, civilizational risk, Nash equilibrium, pollution tax.

*Department of Economics and SC Johnson College of Business, Cornell University.

†Stockholm School of Economics, and Institute for Advanced Study in Toulouse. Weibull's research was partly financed by the Knut and Alice Wallenberg Research Foundation, Sweden, and by ANR-Labex, Institute for Advanced Study in Toulouse, France, Chaire IDEX ANR-11-IDEX-0002-02.

‡The authors are grateful for comments and suggestions from Atahan Afsar, Geir Asheim, Avinash Dixit, Tore Ellingsen, Paul Seabright and Peter Wikman. Early ideas from this paper were presented at a student seminar at the Indian Institute of Technology, Delhi, on 23 February, 2019. We are grateful for the stimulating discourse that followed.

1. Affluence and civilizational collapse

Sixty-six million years ago the dinosaur lived peacefully enough on earth. It had evolved steadily, from the relatively small creature to the massive Tyrannosaurus rex. Then within a relatively short period—meaning a million years—it was gone, extinct. It may have been a massive extra-terrestrial asteroid hit, a bout of increasingly severe volcanos or a gradual shift in the climate, with changing sea-levels. Whatever the cause, the dinosaur had evolved into a creature ready for extinction. However, there were other creatures from that time—snails, birds and sea urchins—that survived and even flourished.

The Indus Valley civilization collapsed somewhere between 1800 and 1500 BCE. The original belief was that this sophisticated society buckled under repeated attacks by a nomadic group, the Aryans, who had one advantage—they had learned to tame and ride the horse. This has now given way to the theory of climate change. The slow eastward shift of the monsoon and also the plentiful use of water by the inhabitants of the civilization, unaware that this was affecting their own viability, caused water tables to drop, trees to perish and ultimately for the civilization to collapse and scatter.

This paper creates a fable, the dinosaur game, that society plays, with a slow drift beneath the surface, brought about by technological progress or some environmental drift. But this same process also moves the society towards its own possible doom. If the players continue the same way, each maximizing his or her own payoff, this same behavior becomes the source of the society's collapse. We call the dinosaur game a fable because it does not try to describe any reality pertaining to the dinosaur but simply illustrates how the slow shift of the ground beneath our feet, even when it looks like progress, could be laying the seeds of our own collapse. While this game is imaginary, it illustrates some of the real risks that human societies face and how we may want to analyze them to escape such a predicament.

One difference that we humans have with the dinosaur is that we are the first creatures who can think and analyze the risks of our own extinction, and maybe even take corrective actions. The people of Indus Valley, though they were capable of self-analysis, were not endowed with enough scientific knowledge to know that they were damaging the environment in which they lived and pushing it towards a non-viable state. As the Economist magazine recently pointed out while discussing contemporary technological advances in synthetic biology, that, unlike earlier advances when their importance and challenge were realized in retrospect, "This time there will be foresight. It will not be perfect. There will be unanticipated effects. . . . It will challenge the human capacity for wisdom and foresight. It might defeat it. But carefully nurtured, it might also help expand it." [The Economist, April 6, 2019, p. 11]

This is precisely what the fable in this paper tries to highlight. It illustrates a slow drift towards a possible calamity and urges us to analyze our own predicament and try to change our own behavior. Today, we have the capacity for self-analysis and also an increasing body of scientific knowledge. This capacity of human beings to step beyond their immediate self-interest is usually ignored in standard economics and game theory. Yet it is important in reality. As Dixit and Nalebuff (2008) observe (p. x), "Strategic thinking [...] is also the art of finding ways to cooperate, even when others are motivated by self-interest, not benevolence. It is the art of convincing others, and even yourself, to do what you say. It is the art of interpreting and revealing information. It is the art of putting yourself in others' shoes so as to predict and influence what they will do." This thinking beyond the immediate game may be hard to formalize but that is what we need to do to promote sustainability.

We run similar risks when we get ideologically committed to the belief that in everyday economic life, the invisible hand will always guide society to an optimal outcome. The invisible hand is indeed a beautiful idea and it can deliver a lot. We just have to look around and see what ordinary individuals have achieved without guidance from the state or heaven. However, there is no reason to believe that this will always be so. As our environment changes, and as the advance of technology continues, it is entirely plausible that our economic life, depicted by the underlying game of life, will also change. The invisible hand, which worked well for a long time, may cease to do so, beyond a certain point.

When early human beings lived in village societies with simple technologies there may have been no compelling reason to have conventions on climate change and regulatory restraints on behavior that has environmental externalities on other societies. That is no longer the case now. Without such conventions and restraints on individual behavior and the actions of nations, it is unlikely that human civilization as we know it will survive.

Likewise, the slow march of technology, globalization, and increasing returns to scale, may be changing the nature of the market, which needs new laws and regulations for us to survive. The dinosaur game is a fable to illustrate this, to show that the same behavior which promotes societal well-being and development may, beyond a critical point, become destructive, leading society towards civilizational collapse. As Harari (2011) reminds us so eloquently (p. 115): "Yet, it was at that moment of maximum affluence that the Roman political order collapsed into a series of deadly civil wars. Yugoslavia in 1991 had more than enough resources to feed all its inhabitants, and still disintegrated into a terrible bloodbath."

The aim of the fable is to contribute to our ability for self-analysis so that we can better understand if we are heading towards a collapse and, if so, to try to modify our behavior and to bring regulation where none may have been needed earlier, and, equally, remove regulations that may have outlived their purpose and are mere residues of past needs. In order to keep the analysis simple, we neglect distributional issues. Arguably, the poor will be more severely hit by pollution than the rich. The present simple model can be generalized to incorporate this aspect, albeit at a relatively high cost of analytical complexity, and is therefore left out. Instead, the analysis turns around technological change and the environment. In the words of Nordhaus (2019): "...the economics of technological change and the modelling of climate-change economics [...] might at first seem to live in separate universes. The truth is that they are manifestations of the same fundamental phenomenon, which is a global externality or global public good. Both involve science and technology, and both involve the inability of private markets to provide an efficient allocation of resources." (op.cit. p. 1991)

2. The dinosaur game

To keep the modeling simple and spare, we shall assume, as economists and game theorists usually do, that 'rationality' is defined as behavior meant to maximize one's own utility or payoff. Beyond this, there is no need for the subsequent analysis to assume that rationality is common knowledge. We shall, instead assume

> [**A1**] All individuals are rational and their preferences are represented by the game payoffs,
>
> [**A2**] A1 is known by all individuals.

The base-line dinosaur game is a finite and symmetric two-player game, in which each of the two players chooses a strategy from the same set $S = \{1, 2, 3\}$. The game is described in the payoff table given below, with player 1 choosing row and player 2 column, the number in each entry of the table being the payoff to player 1.

Table 1: The base-line dinosaur game.

	1	2	3
1	1	3	6
2	2	4	7
3	0	5	8

Given the above assumptions, A1 and A2, the outcome of this game is predictable. Clearly, no rational player will use strategy 1 because this is strictly dominated by strategy 2. No matter what the other player does you earn more by choosing 2 instead of 1. Since both players are rational and know that both are rational, they know that strategy 1 will not be chosen by anyone of them. That being so, it is rational not to choose strategy 2, since strategy 3 is better as long as the other player does not use strategy 1. In brief, the outcome of the above game is that both will play strategy 3.

As a consequence, they will earn payoff 8 each. In the language of game theory the strategy pair $(3,3)$ is a (strict) Nash equilibrium. But, not only that, it is the unique strategy pair that survives the iterated elimination of strictly dominated strategies. So this outcome is very robust.

However, this is not the full dinosaur game, but its start. The full dinosaur game involves a drift, that is, the game payoffs slowly shift over time. The full dinosaur game involves describing for each time $t \geq 0$ a game $G(t)$. The game captures the idea that danger of collapse lurks even when technology is progressing. So, here is the full dinosaur game. It is played by individuals in a large population, where the influence of each individual's action is negligible. It is as if individuals "play the field" in the sense that each individual at each moment in time chooses one of the three pure strategies, and the current payoff to each strategy depends on the population fraction using each of the three strategies. Aggregate population behavior may over time influence the natural environment as well as the rate of technological progress. At any time $t \geq 0$, the dinosaur game, $G(t)$, is defined by the following payoff table:

Table 2: The full dinosaur game.

	1	2	3
1	$1+2r-\alpha q^2$	$3+8r-\alpha q^2$	$6+19r-\alpha q^2$
2	$2+r-\alpha q^2$	$4+7r-\alpha q^2$	$7+20r-\alpha q^2$
3	$r-\alpha q^2$	$5+6r-\alpha q^2$	$8+19r-\alpha q^2$

Here $\alpha > 0$, $r = r(t)$ is the current *level of technology* and $q = q(t)$ the current *level of pollution*, both levels normalized to be zero at time $t = 0$. As seen in the table, each payoff increases linearly with the level of technology and decreases quadratically with the pollution level. These levels, in turn, evolve continuously over time, an evolution that may depend on the history of population play up to the time t in question. More precisely, we suppose that their evolution is given by the following pair of differential equations (with "$\cdot$" signifying time derivative):

$$\begin{cases} \dot{r}(t) = \phi\left[p(t), q(t), r(t)\right] \\ \dot{q}(t) = \psi\left[p(t), q(t), r(t)\right] \end{cases}, \tag{1}$$

where $p = p(t) = (p_1(t), p_2(t), p_3(t))$ is aggregate population play at time t, with $p_k(t) \geq 0$ denoting the population fraction using pure strategy $k \in S$ at time $t \geq 0$.[1] In order for this dynamic to be well-defined, we take both functions, ϕ and ψ, to be Lipschitz continuous. In addition, we assume ψ to be decreasing in p_1 and increasing

[1] As noted above, under assumptions A1 and A2, at time zero all individuals use strategy 3, so $p(0) = (0,0,1)$. We assume that thereafter individuals' strategy choices are made at discrete times, see Section 3.

in p_3, with $\psi(p,q,r) < 0$ if $p_1 = 1$, $\psi(p,r,q) = 0$ if $p_2 = 1$, and $\psi(p,r,q) > 0$ if $p_3 = 1$. In other words, the pollution level decreases if everybody uses strategy 1, stays constant if everybody uses strategy 2, and increases if everybody uses strategy 3. Strategy 1 may thus be called "environmentally friendly", strategy 2 "environmentally sustainable", and strategy 3 "environmentally harmful". Technology growth may hence depend on the current level of pollution and on how much each strategy is used, such endogenous-growth effects being reflected by the function ϕ. Moreover, the level of technology may influence the environmental consequences of the three strategies. Such dependence of pollution on technology is reflected by the third argument of the function ψ.

Looking back at the payoffs in Table 2, we note that if everyone else uses strategy 3 ("business as usual"), then all three strategies' payoffs increase strongly with the level of technology, at about the same rate, but with the environmentally sustainable strategy 2 at a slightly higher rate. If everyone else uses strategy 2, then payoffs to all three strategies still grow with technology, but less than when all others use strategy 3, and now the payoff of the environmentally friendly strategy 1 grows a bit faster than that of strategy 2, which in turn grows a bit faster than that to strategy 3. If, finally, all others use strategy 1, then the payoffs to all strategies still grow with technology, but now at slower rates, and now the payoff of strategy 1 grows the fastest. We assume that individuals can frequently change their strategies, so their time preferences do not matter for their strategy choice.[2]

We have intentionally made the structure of the payoff table in the dinosaur game in Table 2 simple, in order to facilitate the game-theoretic analysis. In particular, pollution enters as a constant term, added to the payoff to each strategy. The strategic nature of the game does thus not depend on the level of pollution. By contrast, the strategic nature of the game does depend on the level of technology.[3] Under assumptions A1 and A2, all individuals will initially, at time $t = 0$, use only strategy 3, and they will continue doing so as long as the level of technology, r, remains below unity. This can be verified by inspection of Table 2, where strategy 1 is strictly dominated by strategy 2 whenever $r < 1$, and, if nobody uses strategy 1, then also strategy 2 becomes strictly dominated.

That the level of pollution does not influence strategy choice may first appear strange. However, this follows from three assumptions in combination: (i) individuals' preferences are represented by their individual payoffs, (ii) pollution is unaffected by any single individual's strategy choice, and (iii) the payoffs to all strategies are equally

[2] See Section 3 for a more precise specification.

[3] A more general class of dinosaur games is obtained by letting its payffs be of the form $\pi_{ij} + \gamma_{ij} r - \alpha_{ij} q^2$, where π_{ij} are the payoffs in Table 1 and γ_{ij} and α_{ij} are positive coefficients that represent the sensitivity to technology and pollution of the payoff to strategy i when matched against strategy j.

affected by the current level of pollution. Assumption (iii) is arguably reasonable in some choice situations, such as when choosing among alternative energy systems for one's house. However, there clearly are other situations in which pollution does affect strategy payoffs differentially. This is, for example, the case when choosing between recreational activities such as lake swimming versus reading a book, where the payoff to swimming is evidently sensitive to the level of water pollution but reading is not. By letting the pollution sensitivity parameter α differ across strategies, the dinosaur game in Table 2 can be generalized to permit such differential payoff effects.

Suppose that the level of technology rises gradually over time. During an initial time interval, everybody will continue using strategy 3. However, when the level of technology at some point in time crosses over the value 1, then the strategic character of the game suddenly undergoes a drastic transformation. The iterated elimination of dominated strategies now instead leads everybody to use the environmentally friendly strategy 1. Indeed, if $r > 1$, then the dinosaur game, defined in Table 2, flips over to become a Traveler's Dilemma (Basu, 1994). This is a dominance solvable game with strategy 1 as its unique solution; for any value $r > 1$, strategy 2 strictly dominates strategy 3, and, once strategy 3 has been removed, strategy 1 strictly dominates strategy 2. Hence, for all technology levels r above unity, the dinosaur game is dominance solvable with the strategy pair $(1, 1)$ as its unique solution. As a consequence, even though everybody's payoffs may initially increase continuously over time (although at a decreasing rate since the pollution level gradually increases), standard individual rationality will lead the population to a potential disaster at the critical time when the level of technology hits the value 1. At that time, all individuals will switch to strategy 1, and thus their payoffs will immediately fall from $27 - \alpha q^2$ to $3 - \alpha q^2$. This amounts to a drastic fall in everyone's payoff, and may thus lead to civilization collapse, a possibility that we illustrate and discuss within a numerical example in the next section.[4]

However, let us first step back and reflect on the big picture that we here try to capture by means of a simple mathematical model. Because although in this paper we tell a story in terms of technology and pollution, the underlying message is much broader. Ideologues in economics often take sides on whether human decision-making should be left entirely to each individual or whether there should be some intervention at the level of the collective or the state. Unfortunately, reality does not lend itself easily to such universal rules. The slow shifting of the ground beneath our feet can change the strategic environment in such a way that the invisible hand may cease to do the job.

[4]While in absolute terms the drop is 24 payoff units, in relative terms the size of the drop depends on the environmental sensitivity, α, and on the current level, q, of pollution.

This civilizational risk and the challenge of climatic disasters, and the need to go beyond business as usual, is beginning to receive a lot of attention. As the environmentalist Bill McKibben's (2019) book title, *Falter: Has the Human Game Begun to Play Itself Out?* itself suggests, human beings face the risk of end game. Diamond (2019) drives this point home in his recent review of this book, when he writes (p. 13): "If environmentalists refuse to engage with big companies, in order to push them to do more good things and fewer bad ones, we could well end up in McKibben's worst-case scenario: human extinction." This means that what could once have been left to the free market may now need intervention.

What this paper stresses is the need to have this openness to altering policy needs. Further, the solution need not always lie in government intervention. A modicum of Kantian ethics can play a critical role in preventing the collapse of markets and society. That individual selfishness can lead to social good was such a surprising result that the economics profession was transfixed by it. But the more commonplace idea that individual morality is also important for the success of society fell by the wayside.[5] This is analyzed at length in Alger and Weibull (2019).[6] That paper shows that morality can potentially play a role in turning society around from the brink of collapse.

In discussing the issue that we are concerned with in this paper, namely, solving what is essentially a collective action problem, we face some deep conceptual challenges which are not easy to tackle formally but must not be brushed aside for that reason. First, note that many collective action problems in life are solved by social norms which may have evolved over a long time because of a persisting collective action challenge (see Ostrom, 1990). This phenomenon has been discussed recently by Sunstein (2019), who points out (p. 44), "Most social norms solve collective action problems. Some of these problems involve coordination; others involve prisoner's dilemmas. Norms solve such problems by imposing social sanctions on defectors. When defection violate norms, defectors might well feel guilt or shame, important motivational sources. The community may enforce norms through informal punishments, the most extreme form of which is ostracism."[7]

Our game with the slow drift through time forces us to ask a second-order question. If the need for a norm arises gradually and imperceptibly slowly, or suddenly because the game's strategic character undergoes a shift, society may be caught without such a norm. Therein lies the heart of the collective action problem that arises with the dinosaur game. We may have to collectively anticipate the need for such a norm

[5]Exceptions are Arrow (1973), Laffont (1975) and Sen (1977).

[6]See also Bergstrom (2009) and Roemer (2019).

[7]See also Ellickson (1991), Lindbeck, Nyberg and Weibull (1999). For several perspectives on how psychology and social reciprocators can result in moral behavior and cooperation, see Gintis, Bowles, Boyd and Fehr (2005).

in advance and try to bring it in consciously, through collective deliberation. But this in turn gives rise to a philosophically troublesome question that arises, which is the following. If it is possible for the players to decide that it is time to play the game differently or to do something collectively that they did not do earlier, then that decision possibility should have been part of the description of the game of life (Mailath, Morris and Postlewaite, 2017; Basu, 2018). After all, the game of life is meant to include all the things the players can do. Hence, if the game we described above was truly the game of life, then it is impossible for players to propose that they play the game differently (since the scope for making such a proposal was not a part of the game). In that case, the civilizational collapse is inevitable.

One way to counter this is to take the line that there is really no such thing as the "game of life" (Basu, 2018). And this leads to a related way to handle this problem. This entails bringing in the idea of tenable strategy blocks (Myerson and Weibull, 2015), which suggests that human beings often have many strategies at their disposal but get used to considering only a subset of strategies. They then behave as if that subset is all the strategies they have. By this approach, players in the above game, viewed as a block game within a larger game, can really do many things, such as talk to one another and propose collective actions and changes; but for long years they may not have needed to consider any of these things and their only relevant decision was to choose one strategy from the set $\{1, 2, 3\}$. So these other options were dormant—but can potentially be brought back to attention.[8]

Where we have to go further, and this is beyond what is done in any previous work, including our own, is to recognize that the 'set' of actions open to each individual is not really a well-defined set. During conventional game play they get used to choosing from within some well-defined subset of the things that they can actually do and so for all practical purposes that is the set of strategies a player has. But civilizations do reach times when players have to think beyond this set and marshal new ideas for new actions, which were always there but had gone dormant. That is what happens in the Dinosaur game. As r approaches 1 we reach such a situation where we have to think of new actions. Formalizing this is not easy but, on the other hand, ignoring this because we cannot formalize it is not right.

3. Example

Let us briefly consider a simple parametric example of the above model. Suppose, first, that technological progress is constant over time and independent of the level of pollution under "business as usual". Formally: $\phi(p, q, r) = \lambda$ if $p_3 = 1$, where $\lambda > 0$ is a fixed constant. Secondly, suppose that pollution increases at a constant rate under "business as usual", and diminishes exponentially over time towards zero

[8]See Alger and Weibull (2017) for a model of such a dynamic process.

if all use the environmentally friendly strategy 1. To be more specific, assume that $\psi(p,q,r) = -q$ if $p_1 = 1$, $\psi(p,q,r) = 0$ if $p_2 = 1$, and $\psi(p,q,r) = 1$ if $p_3 = 1$. Third, let $\alpha = 1$.

Under this specification, all individuals will choose strategy 3 at all times $t < t^*$, where $t^* = 1/\lambda$ is the time at which the technology level reaches unity.[9] Hence, until t^*, both technology and pollution increase linearly over time, according to the equations $r(t) = \lambda t$ and $q(t) = t$. In this time interval, each individual's expected payoff will change continuously over time according to $\pi(3,t) = 8 + 19\lambda t - t^2$. Hence, everyone's payoff initially increases over time, but at a decreasing rate, and reaches its maximum at time $\hat{t} = 19\lambda/2$, from which time on it declines at an increasing rate. This "payoff bliss" time $\hat{t}$ occurs before the strategy switching time t^* if and only if $\lambda < \sqrt{2/19} \approx 0.324$. Irrespective of the speed λ of technology growth, everyone's payoff will suddenly drop at time t^* from $27 - 1/\lambda^2$ to $3 - 1/\lambda^2$. Such a drastic drop may cause the civilization to collapse. There may not be sufficient food and shelter for everyone, or some environmental calamity may strike.

See the diagram below, drawn for $\lambda = 0.3$ (for which $\hat{t} \approx 2.85$ and $t^* \approx 3.33$). The discontinuous (red) solid curve is each individual's payoff. It first increases, then reaches its maximum at $\hat{t}$, then gradually decreases, and, at time t^* suddenly drops from 16 to -8. The sharp hilltop-shaped (blue) dashed curve is the pollution level (scaled up by a factor 3 for increased visibility in the diagram), showing how this level first increases linearly until the critical time t^*, and thereafter declines exponentially towards zero.

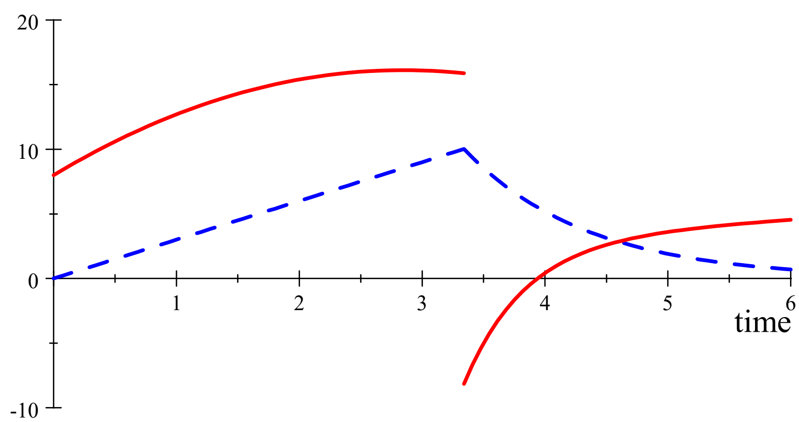

Figure 1: Payoff (solid) and pollution (dashed) under laissez-faire.

[9] We here imagine that strategy choices are made simultaneously by all individuals at exogenously given discrete times with short intervals inbetween. Hence, the strategy change will occur at the first such decision time after t^*. For generic parameter values, t^* does not coincide with such a decision time.

The smooth evolution of payoffs and pollution after the crisis time t^* shown in the diagram is premised on the assumption that the population continues to play the dinosaur game and that the rate of technological progress is maintained. However, if society collapses at or around time t^*, then the "rules of the game" may radically change and the rate of technological progress fall, in which case the post-crisis future may be much more irregular and grim, in terms of payoffs and pollution, than depicted.

The scenario in Figure 1 unfolds under *laissez-faire* when all individuals act rationally. And this is true independently of their time preferences, and is also true even if they care about pollution. The reason is simple: having only an infinitesimal influence on pollution, myopic maximization of one's own material payoff is the unique optimal solution to a forward-looking individual's dynamic choice problem, if he or she is a consequentialist who cares about his or her own material well-being and potentially also about the environment.

Suppose, first, that the *laissez-faire* policy at some time t_0 before t^* (the time of collapse under laissez-faire) is replaced by a governmental regulation that from time t_0 on forbids the use of the environmentally harmful strategy 3. Everybody will then switch to strategy 2 at time t_0, and—if the rate of technical progress does not change much due to the population's change of strategy—at around time t^* switch from strategy 2 to strategy 1 (because then strategy 1 starts to strictly dominate strategy 2). In the intermediate time interval, between t_0 and t^*, the pollution level will remain constant while the rate of technical progress may change somewhat (depending on its degree of endogeneity as reflected by the function ϕ). In any case the fall in payoff levels at time t^* will not be as drastic as under *laissez-faire*.

Second, suppose that instead of forbidding the use of strategy 3 from time t_0 on, a fixed (*per capita* lump-sum) tax τ is imposed on all users of the environmentally harmful strategy 3. If the tax revenues are instantly returned to all individuals as a (*per capita* lump-sum) transfer β, the same for all (even those who use strategy 3), then governmental budget balance condition requires $\beta(t) = \tau \cdot p_3(t)$ at all times t. The transfer will not affect individuals' choice of strategy. Only the tax has an incentive effect, making strategy 3 less attractive. The resulting scenario is similar to that under prohibition of strategy 3.

As a third scenario, suppose that a benevolent planner, striving to maximize the present value of the future payoff-stream to the citizens, could decide what strategies people are to use at all times. This is an optimal-control problem with (q, r) as the state variable and p as the control variable. The optimal policy, $\langle \hat{p}(t) \rangle_{t \geq 0}$, will depend in part on the sensitivity α of payoffs to pollution, on the dynamics of technical progress, as well as on the discount factor, δ, that the planner attaches to future payoffs.[10] The lower values $\alpha > 0$ and $\delta > 0$ have, the less urgent is the switch away from strategy 3, "business as usual," to strategies 1 or 2. Particularly critical is the time discount factor δ, just as it is in the current debate about climate change.

In closing the section, it is worth pointing to some parallels with the extensive literature on tipping and threshold effects, which followed from the writings of Thomas Schelling (Schelling, 1971). Like in our story, the tipping models draw attention to narrow ledges, beyond which behavior or outcomes undergo dramatic changes. The original models dealt with large masses of individuals and showed how when a critical mass of people behaved in a certain way, a huge mass of people follow and cause mass behavioral change. This can happen when a few people of one ethnic group move out of an area, causing a cascade, as in Schelling's analysis or in models of "spirals of silence," where no one speaks out on some matter, till a critical mass of people, for whatever be the reason, do, and then everyone does (Granovetter and Soong, 1988). In our analysis, the game itself shifts slowly and, as long as the shift occurs within a region, there is no significant change in behavior. But beyond a threshold, the behavior changes drastically and so do the outcomes, causing a calamitous collapse. What our paper tries to do is to alert us against the comfort zone we, including prominent political leaders, can easily slip into by taking the attitude that these small changes have occurred in the past, and we have had no calamity.

4. Morals and more

The above analysis was based on the assumption that all individuals are consequentialists, that they base their strategy choices solely upon the material consequences of their actions for themselves; the material consequences for others and the environment being negligible since the population is very large. Arguably, moral principles also play a role in human decision-making. One moral principle is Immanuel Kant's categorical imperative, "Act only according to that maxim whereby you can, at the same time, will that it should become a universal law." In decisions with externalities, such deontological moral principles may affect decisions even when an individual's action has negligible influence.

In the present context, no individual can influence the level of pollution. Hence, even if individuals do care about future pollution levels and others' welfare, their individual choice of strategy will, under standard economics assumptions, still be determined entirely by their current pure self-interest. However, suppose some individuals have a Kantian moral motivation when choosing strategy. If these individuals make up a positive share of the total population, then, as shown in Alger and Weibull (2019), their behavior will affect the level of pollution. In the extreme case when they

[10] If p can be changed only at discrete times $t = 0, \Delta, 2\Delta, ...$,then $\delta = e^{\rho\Delta}$, where $\rho > 0$ is the planner's (subjective) discount rate per time unit.

are all Kantian moralists, they will all "do the right thing", that is, choose the strategy that would be optimal if everyone in society used it. In more realistic cases, an individual of the *Homo moralis* variety with intermediate *degree of morality* κ (see Alger and Weibull, 2013, 2016, 2019) will choose the strategy that would be optimal if the population share κ would use it.[11] While it is rational for Homo oeconomicus to act myopically in the dinosaur game, the strategy choice of Homo moralis is necessarily dynamic and forward-looking. It thus depends on time preference and the dynamics of technological change and pollution. A Homo moralis may thus switch to an environmentally more friendly strategy even if this runs against her immediate personal material interest. Indeed, this is arguably what we currently see happen in the world, especially among the young. A hopeful sign indeed. But also an analytical challenge for economists, since (dynamic) choice by partly deontologically motivated agents falls outside mainstream economics analysis (see discussion in Alger and Weibull, 2017).

We have here developed a parable, a simple "game of life", in order to get a better grip on the strategic issues involved. However, for all actions described as parts of a game, one can always go beyond this and think of more actions. What may be argued is that we typically go through life by confining our attention to some artificially constructed set of strategies. But, in fact, it is always possible to go beyond any such set. Confronted with major civilizational problems, as raised in this paper, we, as citizens, must force ourselves to think of actions beyond those that conventionally have been described as feasible.

References

Alger, Ingela, and Jorgen Weibull. 2013. "Homo moralis: preference evolution under incomplete information and assortative matching," *Econometrica* 81, 2269-2302.

Alger, Ingela, and Jorgen Weibull. 2016. "Evolution and Kantian morality," *Games and Economic Behavior* 98, 56-67.

Alger, Ingela and Jorgen Weibull. 2017. "Strategic behavior of moralists and altruists," *Games* 8, 1-21.

Alger, Ingela, and Jorgen Weibull. 2019. "Morality: evolutionary foundations and policy implications," in Basu, K., D. Rosenblatt, and C. Sepulveda. (eds.), *The State of the World, The State of Economics*. Cambridge MA: MIT Press.

Arrow, Kenneth. 1973. “Social responsibility and economic efficiency,” *Public Policy*, 21, 303-317.

[11] The degree of morality is a number between zero (*Homo oeconomicus*) and one ("Homo kantientis").

Basu, Kaushik. 1994. "The traveler's dilemma: paradoxes of rationality in game theory," *American Economic Review, Papers and Proceedings* 84, 391-395.

Basu, Kaushik. 2018. *The Republic of Beliefs*. Princeton: Princeton University Press.

Bergstrom, Theodor. 2009. "Ethics, evolution, and games among neighbors," mimeo. University of California, Santa Barbara.

Diamond, Jarred. 2019. "How it Ends," *New York Times Book Review*, 21 April, 2019.

Dixit, Avinash, and Barry Nalebuff. 2008. *The Art of Strategy.* New York: W. W. Norton & Co.

Ellickson, Robert C. 1991. *Order Without Law: How Neighbors Settle Disputes*. Cambridge MA: Harvard University Press.

Gintis, Herbert, Samuel Bowles, Robert Boyd, and Ernst Fehr. 2005. *Moral Sentiments and Material Interests: The Foundations of Cooperation in Economic Life*. Cambridge, MA: MIT Press.

Granovetter, Mark, and Roland Soong. 1988. "Threshold models of diversity: Chinese restaurants, residential segregation, and the spiral of silence," *Sociological Methodology* 18, 69-104.

Harari, Yuval N. 2011: *Sapiens: A Brief History of Humankind*. London: Harper.

Laffont, Jean-J.acques. 1975. "Macroeconomic constraints, economic efficiency and ethics: an introduction to Kantian economics," *Economica*, 42, 430-437.

Lindbeck, Assar, Sten Nyberg and Jorgen Weibull. 1999. "Social norms and economic incentives in the welfare state," *Quarterly Journal of Economics* 114, 1-35.

Mailath, George, Steven Morris, and Adrew Postlewaite. 2017. "Laws and authority," *Research in Economics* 71, 32-42.

McKibben, Bill. 2019. *Falter: Has the Human Game Begun to Play Itself Out?*. New York: Henry Holt & Co.

Myerson, Rorger, and Jorgen Weibullas. 2015. "Tenable strategy blocks and settled equilibria," *Econometrica* 83, 943-976.

Nordhaus, William. 2019. "Climate change: the ultimate challenge of economics," *American Economic Review* 109, 1991-2014.

Ostrom, Elinor. 1990. *Governing the Commons*. Cambridge, UK: Cambridge University Press.

Roemer, John. 2019. *How we Cooperate: A Theory of Kantian Cooperation*. New Haven, CT: Yale University Press.

Schelling, Thomas C. 1971. "Dynamic models of segregation," *Journal of Mathematical Sociology* 1, 143-180.

Sen, Amartya K. 1977. "Rational fools: A critique of the behavioral foundations of economic theory," *Philosophy & Public Affairs* 6, 317-344.

Sunstein, Cass R. 2019: *How Change Happens*. Cambridge, MA: MIT Press.