

Catastrophe by Design: Destabilizing Wasteful Technologies & The Phase Transition from Proof of Work to Proof of Stake

STEFANOS LEONARDOS, IOSIF SAKOS, COSTAS COURCOUBETIS, and GEORGIOS PILIOURAS, Singapore University of Technology and Design

Cryptocurrency mining in Proof of Work (PoW) blockchains is notorious for its expansive environmental footprint. Environment-friendly alternatives such as Proof of Stake (PoS) protocols have been developed, however, adoption is hindered by entrenched economic interests and network effects. To make matters worse, the committed decentralized nature of these ecosystems is contrary to standard mechanism design approaches that rely on strong persistent centralized authorities with abundant resources that e.g., by using preferential subsidies can effectively dictate system outcomes. What other type of mechanisms are feasible?

We develop and analyze a mechanism to induce a transition from PoW to PoS with several desirable properties. The mechanism is transient and does not exogenously favor one technology over another. Instead the phase transition from PoW to PoS emerges endogenously by analyzing a standard evolutionary learning model, Q-learning, where agents trade off exploration and exploitation. Introducing short-term taxation, common for both technologies, encourages exploration and results in irreversible phase transitions and long-lasting stabilization of PoS. At the technical level, our work is based on bifurcation and catastrophe theory, a branch of mathematics that deals with changes in the number and stability properties of equilibria. Critically, our analysis is shown to be structurally robust to significant and even adversarially chosen perturbations of the parameters of both our game and our behavioral model.

CCS Concepts: • **Theory of computation** → **Algorithmic game theory and mechanism design**; **Cryptographic protocols**.

1 INTRODUCTION

Since the launch of Bitcoin (BTC) by the pseudonymous Satoshi Nakamoto, [Nakamoto, 2008], Proof of Work (PoW) blockchains and their applications – most notably cryptocurrencies – have taken the world by storm. Widely considered as a revolutionary technology, blockchains have attracted the attention of institutions, technology-corporations, investors and academics. However, among other concerns, blockchains face a major hurdle in their expansion and broad public adoption: the bottleneck of immense energy waste. Currently, one BTC transaction wastes as much energy – in terms of carbon footprint – as 775.818 VISA transactions [Digiconomist, 2020]. Even more alarming than its current levels – which rank the BTC network above Finland and Pakistan – is the consumption’s increasing trend: the electricity used by the BTC network approximately doubles every year [University of Cambridge, Judge Business School, 2020]. The total picture can only get worse, if one takes into account all other PoW blockchains – such as Ethereum [Buterin et al., 2019].

Authors’ address: Stefanos Leonardos; Iosif Sakos; Costas Courcoubetis; Georgios Piliouras Singapore University of Technology and Design.

These alarming figures call for mechanisms to accelerate the development and, more crucially, the adoption of alternative protocols – or *virtual mining*, [Bentov et al., 2016] – technologies such as Proof of Stake, see e.g., [Brown-Cohen et al., 2019, Hazari and Mahmoud, 2019]. With this in mind, academic research is focusing on the understanding of miners’ behavior and potential strategies, [Fiat et al., 2019a, Goren and Spiegelman, 2019]. Yet, an important barrier is that the value of a cryptocurrency – or in general the reliability of its applications – depends on the size of its mining network (with larger network implying higher safety). More mining power implies that it is more costly for a potential attacker to gather the required resources and compromise the functionality of the blockchain, [Brown-Cohen et al., 2019, Kiayias et al., 2016]. Hence, when the rest of the population mines a specific PoW cryptocurrency, then it is individually rational (preferable) for any single miner to also mine that cryptocurrency.

Using game theoretic terminology, the population state (or equilibrium) in which the PoW technology is used by everyone is evolutionary stable and small perturbations – adopters of alternative technologies – are doomed to fail. This creates a deadlock: a situation – among many known in social and economic sciences – in which selfish behavior stands at odds with the social good. On the other hand, the state in which everyone adapts the new technology is also a stable equilibrium. How can we facilitate the transition from one to the other?

This challenge is primarily not technological¹ but game-theoretical in nature. This tension between individual incentives (miners) and social welfare indicates that this is a setting where mechanism design should be applied, however, the inherently decentralized nature of blockchains raises new challenges that severely lessen the applicability of off-the-self solutions.

First, miners form loosely-organized, decentralized pseudonymous networks that are not governed by central authorities [Eyal and Sirer, 2018]. Even if we assume that a centralized scheme could provide some control over the network, any mechanism that consistently subsidizes socially beneficial behavior by offsetting potential losses would not be economically feasible and would be subject to gaming. Even more of a show-stopper is the fact that any top-down policy that treats differentially one technology versus another would be rather hard, if not outright impossible, to enforce in practice as disgruntled network users can easily splinter off forming new networks. Lastly, standard expected utility models, the bedrock of classic mechanism design, are arguably too simplistic to model miners behavior in practice for several reasons (volatility, risk attitudes, hedging, collusion, politics/governance, e.t.c.). It would be thus important to develop solutions that are robust to more complex behavioral assumptions [Bissias et al., 2019, Chen et al., 2019, Fiat et al., 2019b]. Due to the aforementioned complexities and despite the pressing nature of this problem, to our knowledge, no mechanism has been proposed to address it.

Our model: We introduce an evolutionary game theoretic model to capture miner behavior. The advantage of having both a game theoretic model (Section 2) as well as learning theoretic model (Section 3) is that it allows us to formally argue about the stability

¹Both from theoretical as well as practical perspective PoS technologies have shown to offer strong guarantees analogous to those of PoW [Buterin et al., 2019, Garay et al., 2015, Kiayias et al., 2017].

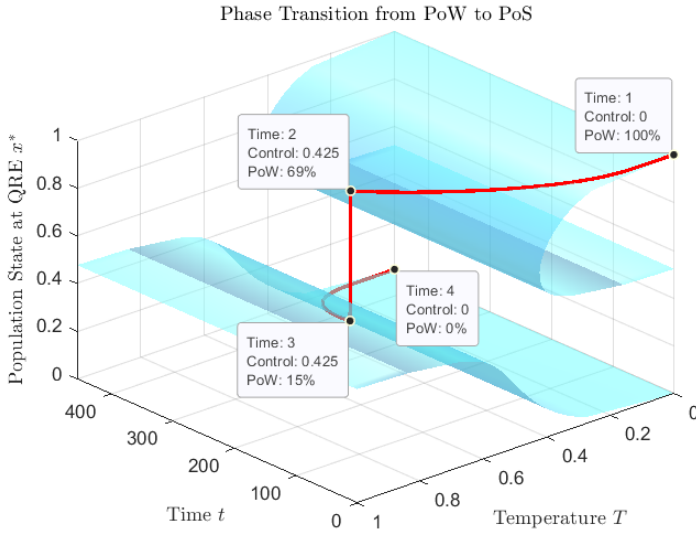


Fig. 1. Phase Transition from Proof of Work (PoW) to Proof of Stake (PoS). The wasteful but currently adopted technology (PoW) is destabilized by a controlled catastrophe and the population moves to a new equilibrium (PoS). At timepoint 1, the system starts with 100% PoW miners and the control parameter is at 0 (upper right datatip). As the control parameter increases, the population state (percentage of PoW miners) moves along the red line on the QRE surface. At timepoint 2 (time is indicative), the control parameter reaches the critical value or tipping point (upper middle datatip). At the next timepoint, 3, at which the parameter T is increased slightly beyond the critical level, the system undergoes an abrupt transition (bottom middle datatip). Between these two successive time points, the population state changes from 69% PoW miners right before the *catastrophe* to only 15% immediately after. After this point, the control parameter can be reset to 0 (here this is done gradually but since the population is now in the attracting region of the new equilibrium this is not necessary) and the system will converge to the new equilibrium in which everyone has adopted the new technology (red line to bottom right datatip at time 4.).

of the equilibria of the game. For example, in the simplest possible game theoretic model of PoW/PoS competition we have that the utility of using the PoW (resp. PoS strategy) increases linearly in the number of other agents that are using the same technology. This results in three type of fixed points, everyone using PoW, everyone using PoS, and a “mixed” population case at the exact split where both technologies are equally desirable/profitable. Intuitively, this mixed state is an unstable equilibrium as a slight increase of the fraction of PoW (resp. PoS) miners is enough to break ties and encourage convergence to a monomorphic state. However, to make this discussion concrete we need to formally describe how a mixed population state (i.e., the PoW/PoS split) evolves over time.

To model the adaptive behavior of the agents, we use one of the most well known models of evolutionary reinforcement learning, the Boltzmann Q-learning dynamics, [Tan, 1997, Watkins and Dayan, 1992]. The decision of each miner, or equivalently of each unit investment, is whether to adopt the PoW or the PoS technology given that $x \in [0, 1]$ fraction

of the population adopts the PoW technology. According to the Q-learning behavioral model, the miners update their actions by keeping track of the collective past performance – in particular, of a properly defined Q-score – of their available strategies (PoW or PoS). However, in a far-reaching twist, each miner’s utility function is enhanced by an entropy term, that is weighted by a parameter T – termed *temperature* or *rationality*. Informally, low temperatures capture cool-headed agents that focus primarily on strategies with good historical performance, whereas high temperatures favor hedging and exploration. Roughly, (see Section 2 for the formal specifications and notation), if we denote with $u(W, x)$ and $u(S, x)$ the utility of an investment unit from either of the two technologies PoW and PoS, when the state of the population is $(x, 1 - x)$ with $x \in [0, 1]$ denoting the fraction of PoW miners, then the Q-learning dynamics are given by the following scheme

$$\dot{x} = x \underbrace{[u(W, x) - u(S, x)]}_{\text{Replicator Dynamics}} - T \cdot \underbrace{[x \ln x + (1 - x) \ln (1 - x)]}_{\text{Entropy}} \quad (1)$$

When $T = 0$, the dynamics are precisely the replicator dynamics [Sandholm, 2010] and they recover the Nash equilibria of the game. Now, we can formally argue that the equilibrium $x = 1$ in which all miners use the PoW technology is *evolutionary stable* (Proposition 2.1). How do we escape from it and converge to $x = 0$, the PoS equilibrium?

Our solution: Catastrophe Design. Our approach is based on the combination of two observations: First, the number and stability of equilibria of Q-learning, known as Quantal Response Equilibria [McKelvey and Palfrey, 1995], is a function of T , i.e., the tradeoff level between exploration and exploitation. In Figure 1 for $T = 0$ there exist three QRE, the three Nash equilibria. For slightly larger $T > 0$ we still have three QRE but now due to the exploration term all three lie in the interior of the interval $(0, 1)$. Finally, beyond some critical temperature the number of QRE drops from three to one. Such phenomena are known as catastrophes, bifurcations [Kuznetsov, 2004].

The second observation is that we can effectively control temperature T (e.g., increase it by a multiplicative scale α), since it is mathematically equivalent to scaling down the utility of the agents by the same factor α (up to time repamaterization in equation 1). In policy terms, taxation of income (i.e. multiplicative decrease of payoffs for all actions) results in agent behavior which is less stringent about maximizing earnings at all costs. Informally and taking this idea to its logical extreme, a taxation level of 99% (i.e., a very large T) would effectively render the agents indifferent about payoffs and make them choose actions at random. More discussion on this connection can be found here [Kianercy and Galstyan, 2012, Wolpert et al., 2012, Yang et al., 2017].

Putting these two observations together, by controlling T , we can control the resulting QRE and thus the resulting state of the system over time. More critically, when exceeding critical temperatures (*tipping point*), a phase transition or catastrophe at which the state behavior changes abruptly is possible (Figure 1 from time 2 to time 3). Finally, we could (in principle at least) leverage these catastrophes to create irreversible phase transitions such that even when the controlling parameter T returns to its initial state $T = 0$, the state

of the systems is not the original undesirable stable state ($x = 1$), but the target PoS state ($x = 0$) (Figure 1 from time 1 to time 4).²

Our key contribution is to prove that there exists a simple, robust and transient catastrophe-based mechanism to destabilize the PoW equilibrium and enforce the PoS equilibrium. As a first step in the process, we provide a complete characterization of both the number as well as the stability of QREs given game theoretic models of PoW/PoS competition for all values of T (Theorem 4.4). This is a question of independent interest as it explores the question of what are the possible limit system behaviors in the lack of any controlling mechanism. Secondly, we describe how by simply raising taxation/temperature up to (slightly beyond) a critical value and then reducing it down to zero results into convergence to the socially optimal PoS equilibrium (Theorem 4.6).

Robustness of findings. Critically, we stress-test our findings and establish that they are robust to modelling uncertainty/misspecifications across different axes, cf. Section 5. In Section 5.1, we allow (possibly adversarial) uncertainty/perturbations in the dynamics and the population game theoretical model. In Section 5.2, we explore nonlinear utility functions which can capture e.g., superadditive effects on network valuation but introduce significant difficulties as these models lie outside the typical framework of evolutionary game theory with multi-linear utility functions. While *the number of equilibria – and the stability properties – of the dynamics may change, the resulting control mechanism can be applied without any significant changes*. In short, our baseline modelling assumptions are chosen with an eye towards simplicity primarily for expositional purposes. Our results, however, can be directly extended for a wide range of modelling designs and approaches.

2 MODEL: BLOCKCHAIN POPULATION GAME

We consider a society or population p of agents, investors or miners (physical or virtual)³ who form a continuum of mass $K > 0$. Here K denotes the total available capital or resources, e.g., money, hardware or electricity, all expressed in monetary terms, that the agents are willing to invest. There are two available strategies or technologies: a costly technology, W , and an innovative technology, S . Here W and S stand for Proof of Work and Proof of Stake, respectively, but the model applies to any similar setting. Accordingly, the *set of strategies or available technologies* is denoted by $A = \{W, S\}$. Investing one (infinitesimal) unit of resource (e.g., one dollar \$) in technology W incurs a cost of $\gamma > 0$ to the investor, while an investment in technology S incurs a cost of 0. This assumption implies that the investors disregard any third potential alternative and will invest all available resources on either of the two technologies (there is no loss from doing so for rational agents). Accordingly, the set of *population states* is $X = \{(x, 1 - x) : x \in [0, 1]\}$, where $x \in [0, 1]$ denotes the fraction of agents (in terms of capital or resources) in population p that are choosing technology W .

The *payoff function*, $u : X \rightarrow \mathbb{R}^2$, assigns to each population state a vector of payoffs, one for each strategy in A . In this context, we assume that the total value created by each

²Such phenomena where the state of a system depends on its history (path dependent) are ubiquitous in physical processes (e.g., the magnetization process used to record tapes) are known as hysteresis [Wolpert et al., 2012, Yang et al., 2017].

³The theory of population games and revision protocols that is presented here follows [Sandholm, 2010].

technology $\{W, S\}$ in A depends on the fraction x of capital that has been invested in that technology via a parameter $\alpha > 0$ and that the total value is distributed evenly among all invested units (fairness property of prevailing blockchain reward schemes, [Chen et al., 2019, Fiat et al., 2019a]). We assume that both technologies can generate an aggregate value $V > K$, if fully adopted by the population of investors. In particular, the value V_W created by technology W depends on the population state x via the relationship

$$V_W = V \cdot (xK)^\alpha \quad (2)$$

and similarly, for the value V_S created by technology S , $V_S = V \cdot ((1-x)K)^\alpha$. Here $\alpha > 0$ is a parameter that expresses the relationship between the value created by a technology and the degree of its adoption. Different values of α give rise to different problems as discussed in Remark 3. In particular, $\alpha < 1$ implies subadditive value (it is optimal for the population to split), $\alpha = 1$ implies linear value and $\alpha > 1$ implies superadditive value, i.e., the population is better off if it fully adopts either of the two technologies. Accordingly, the payoff of each strategy $\{W, S\} \in A$ is given by

$$u(W, x) = V_W \cdot \frac{1}{xK} - \gamma = V \cdot \frac{(xK)^\alpha}{xK} - \gamma = VK^{\alpha-1}x^{\alpha-1} - \gamma \quad (3a)$$

$$u(S, x) = V_S \cdot \frac{1}{(1-x)K} = V \cdot \frac{[(1-x)K]^\alpha}{xK} = VK^{\alpha-1}(1-x)^{\alpha-1} \quad (3b)$$

Hence, the *average payoff* obtained by the members of the mining population at population state x is equal to

$$\bar{u}(x) = xu(W, x) + (1-x)u(S, x) = VK^{\alpha-1} \left[x^\alpha - \frac{\gamma x}{VK^{\alpha-1}} + (1-x)^\alpha \right] \quad (4)$$

and the *aggregate payoff* that is achieved by the population as a whole is $u_A(x) = K\bar{u}(x)$. The cost γx is paid by the population as a whole and hence, captures the negative externality (or cost) of the undesirable technology.

2.1 Evolutionary Game and Nash Equilibria

To study instances in which a union is preferable than a split, our main focus will be the case $\alpha \geq 1$. For expositional purposes, we will restrict attention on the case $\alpha = 2$, but all arguments essentially carry over to any $\alpha > 1$ (and to the trivial case, $\alpha = 1$) as we show in Section 5.2. For $\alpha = 2$, eqs. (3) and (4) become linear in x ,

$$u(W, x) = VKx - \gamma, \quad u(S, x) = VK(1-x) \quad (5a)$$

This allows for an equivalent (yet more intuitive) interpretation of the agents' interaction as a single population evolutionary game, cf. [Hofbauer and Sigmund, 1998]. Hence, by substituting $x = 0$ and $x = 1$, we can represent the game by the following matrix

$$P = \begin{array}{c} \\ W \\ S \end{array} \begin{array}{cc} W & S \\ \left(\begin{array}{cc} VK - \gamma & -\gamma \\ 0 & VK \end{array} \right) \end{array} \quad (G1)$$

Since $V \gg K \gg \gamma > 0$, we can normalize VK to 1 and write $\gamma \rightarrow \gamma/VK$ with $\gamma \in (0, 1)$. The equilibria of the resulting game are characterized next. The proof is standard and is presented for completeness in the Appendix A.

PROPOSITION 2.1. *For $VK = 1$, the population game described by the payoff function in (5) is a single population evolutionary game with three Nash equilibria: $x_1 = 0$ with payoffs $\bar{u}(x_1) = 1$, $x_2 = \frac{1+\gamma}{2}$ with payoffs $\bar{u}(x_2) = \left(\frac{1-\gamma}{2}\right)$ and $x_3 = 1$ with payoffs $\bar{u}(x_3) = 1 - \gamma$. The two pure equilibria, $x_1 = 0$ and $x_3 = 1$ are evolutionary stable, whereas the fully mixed equilibrium, x_2 , is not. Equilibrium x_1 – in which the desired technology is fully adopted – is strictly payoff and risk dominant.*

Accordingly, the undesirable (W, W) equilibrium in which everybody adopts the costly technology, while payoff and risk dominated by the desirable environmental friendly equilibrium (S, S) , is evolutionary stable. This is precisely the reason that the introduction of the new, cost efficient technology may not succeed. In particular, evolutionary stability implies that even if a small part of the population adopts the new technology, it is unlikely that the system will be able to move away from the current equilibrium.

This motivates to look for an alternative approach to destabilize the system and obtain the desired outcome. The solution that we pursue in this case is via the hysteresis and optimal control design mechanisms that leverage the effects of bifurcations that emerge in the population dynamics when the game is played repeatedly in a online setting (as is the case in real applications), [Wolpert et al., 2012, Yang et al., 2017]. The framework to develop these mechanisms is described next.

3 BEHAVIORAL MODEL

Predictions that are based on equilibration notions are difficult to justify in the context of large number of agents, [Sandholm, 2010]. Moreover, the adoption of a new technology, is a gradual process at which agents gradually *learn* their strategies via repeated interaction with their environment. In particular, each agent’s payoff depends on the fractions $x, 1 - x$ of the population that adopt either of the two technologies, cf. equation (3), which is constantly changing. Using the term *revision protocols*, [Sandholm, 2010] provides several alternative microfoundations of the standard *replicator dynamics* in this context. However, although successful in many ways, the best response dynamics characterize greedy and myopic behavior with its own limitations, [Fiat et al., 2019b].

As argued in Section 1, a unified approach, termed *Q-learning dynamics*, that combines the most useful elements from the two approaches – exploitation of the best strategy by the replicator dynamics and exploration of the strategy space in online learning – has been recently developed [Tuyls et al., 2003, Wolpert et al., 2012] and [Kianercy and Galstyan, 2012]. In a more abstract setting, [Yang et al., 2017] among others, demonstrate how these dynamics can be used to provide improved predictions of the agents’ behaviors in large populations and more importantly, how the control parameter $T \geq 0$ can be utilized to steer their behavior of the system from a mechanism design perspective. The *Q-learning dynamics* are based on the principle of *Q-learning*, [Watkins and Dayan, 1992], and the

framework that is presented here closely follows the aforementioned works without further reference.

3.1 Population States with Q-learning Dynamics

We consider a setting in which each agent of the population is adapting their strategy according to repeated interaction with their environment. This allows us to focus on the standpoint of a single agent and describe the system via the Q -learning dynamics that are derived as follows:

Q -values: At each time t , the learning agent assigns a value $Q_t(j)$ to each strategy $j \in A = \{W, S\}$ via the update rule

$$Q_{t+1}(j) = Q_t(j) + \delta [u_t(j, x_t) - Q_t(j)] \quad (6)$$

where $\delta > 0$ is the learning rate and $u_t(j, x_t)$ for $j = W, S$ is the reward from selecting strategy $j = W, S$ (as given by eq. (3)) provided that the distribution of the population is $x_t \in [0, 1]$ at time point $t > 0$.

Strategies & Population States: Using the Q -values, the learning agent's critical decision is the update of their strategy. To avoid suboptimal results derived by a greedy updating, i.e., selecting the strategy with the highest Q -value, the agents incorporate in their maximization decision an entropy term to reward *exploration* of the whole strategy space. In particular, the learning agent selects their strategy $(x_t) \in [0, 1]$, at time point $t \geq 0$ as the (unique) solution to the convex optimization problem⁴

$$x_t = \arg \max_{x \in (0,1)} \{xQ_t(W) + (1-x)Q_t(S) - T[x \ln x + (1-x) \ln(1-x)]\} \quad (S1)$$

The decision rule in (S1) results in the following mixed strategy (probability distribution)

$$x_t = \frac{e^{Q_t(W)/T}}{e^{Q_t(W)/T} + e^{Q_t(S)/T}} \quad (7)$$

which is known as the *Boltzmann distribution*. Here, we allow the agent to use a mixed strategy $x \in (0, 1)$ with a slight abuse of notation, since x or x_t denotes both the learning agents' strategy and the distribution – state – of the population. Both these notations are equivalent under the assumption that all agents are symmetric and that they are learning concurrently.

Continuous time dynamics: If we take the time interval to be infinitely small, this sequential joint learning process can be approximated as a continuous-time model, [Kianercy and Galstyan, 2012, Tuyls et al., 2003]. After rescaling the time horizon to $\delta t/T$, the continuous-time dynamics of the *population state* are

$$\dot{x} = x \left[u(W, x) - \bar{u}(x) + T \sum_{j=W,S} x_j \ln(x_j/x) \right] \quad (8)$$

which is the desired expression of the dynamics in terms of strategies (rather than Q -values).

⁴For an intuitive explanation of the objective function see Appendix A.

Quantal Response Equilibria (QRE): The steady states of the system in (8) are the roots the right side. In turn, as shown in [Kianercy and Galstyan, 2012], these are precisely the solutions (if they exist) of the maximization problem

$$x^* \in \arg \max_x \left\{ u(W, x) - \bar{u}(x) - T \sum_{j=W,S} x_j \ln x_j \right\}$$

for any given value of *temperatures*, $T > 0$ (the case $T = 0$ corresponds to the replicator dynamics and is best treated separately). A direct calculation yields the solutions

$$x^* = \frac{e^{u(W, x^*)/T}}{e^{u(W, x^*)/T} + e^{u(S, x^*)/T}} \quad (9)$$

which is known as the Gibbs distribution. Points $x \in (0, 1)$ that satisfy equation (9) are known as *Quantal Response Equilibria (QRE)*, cf. [McKelvey and Palfrey, 1995, Yang et al., 2017]. Starting from any interior point⁵ $x \in (0, 1)$, the Q -learning dynamics are known to converge to interior rest points for any $T > 0$.

In the remaining part, our main task is to understand the behavior of the system that is described via the dynamics in equation (8) and explain how a central designer can influence this behavior through the control parameter $T \geq 0$.

4 ANALYSIS

We consider the Q -learning dynamics in a homogeneous population⁶. The critical parameter that we need to track (and update) is the fraction $x \in [0, 1]$ of the population that invests on the costly (and currently prevailing) technology W . By equations (3) and (8) and after normalizing VK to 1, the population state, i.e., the fraction $x \in [0, 1]$ of the total available investment capital, changes according to the following dynamics

$$\dot{x} = x \left[u(W, x) - \bar{u}(x) + T \left(x \ln \left(\frac{x}{1-x} \right) + (1-x) \ln \left(\frac{1-x}{x} \right) \right) \right] \quad (10)$$

$$= x \left[(1-x) [u(W, x) - u(S, x)] + T (1-x) \ln \left(\frac{1-x}{x} \right) \right]$$

$$= x(1-x) \left[u(W, x) - u(S, x) + T \ln \left(\frac{1-x}{x} \right) \right] \quad (11)$$

where $T \geq 0$ is the control parameter, $\gamma \in (0, 1)$ is the cost per unit of investment for the costly technology, here, Proof of Work (PoW). Using, the explicit payoffs in equation (3), equation (10) becomes

$$\dot{x} = x(1-x) \left[2x - (1+\gamma) - T \ln \left(\frac{x}{1-x} \right) \right] \quad (12)$$

Keeping in mind that T can be leveraged as a control variable to influence the evolution of the population, in the remaining part, we will study equation (12) and determine its

⁵Note that the introduction of the exploitation term renders the choices $x = 0$ and $x = 1$ not admissible, since they are not in the domain of $\ln(x/(1-x))$ for any $T > 0$. This is not a problem in any realistic application.

⁶See, e.g., [Arnosti and Weinberg, 2018], who demonstrate that largely heterogeneous players are unlikely to survive in the same competitive environment.

steady states (QRE) and their stability properties for all $T \geq 0$. It is more intuitive to treat the instance $T = 0$ separately. In this case, the system reduces to the well known replicator dynamics and its steady states are precisely the Nash equilibria of the respective evolutionary game, cf. Section 2.1.

LEMMA 4.1. *For $T = 0$, the steady states of the Q -learning dynamics, equation (12), are $x_1 = 0$, $x_2 = \frac{1}{2}(1 + \gamma)$, and $x_3 = 1$. The steady states on the boundary, i.e., x_1 and x_3 , are stable, whereas x_2 is unstable.*

PROOF. For $T = 0$, the dynamics become $\dot{x} = x(1-x)(2x-1-\gamma)$ and the first claim follows trivially. Concerning their stability, observe that $\dot{x} < 0$ for any $x \in (0, \frac{1}{2}(1+\gamma))$ and $\dot{x} > 0$ for $x \in (\frac{1}{2}(1+\gamma), 1)$. Hence, starting from any point other than $x = \frac{1}{2}(1+\gamma)$, the system will converge to the boundary steady states, to $x_1 = 0$ for any initial starting point $x < x_2$ and to $x_3 = 1$ for any initial starting point $x > x_2$, which completes the proof. \square

The steady states and their stability properties are illustrated in Figure 4. To treat the case $T > 0$, we restrict attention to $x \in (0, 1)$ (otherwise $\ln(\frac{x}{1-x})$ is undefined). The key intuition from the instance $T = 0$ is that the stability of the steady states can be fully determined by the sign of the dynamics \dot{x} , due to the fact that the dynamics described by equation (12) are one-dimensional.

For $x \in (0, 1)$, the term $x(1-x)$ is always strictly positive and hence, the sign and the roots of \dot{x} in equation (12), i.e., the steady states or equivalently the Quantal Response Equilibria (QRE) and the direction of movement of the dynamics, are fully determined by the last term of the dynamics in equation (12). Hence, for a given cost parameter $\gamma \in (0, 1)$, and a temperature $T \geq 0$ it will be convenient to define $f(x; T, \gamma)$ for $x \in (0, 1)$ as follows.

Definition 4.2. Given parameters $T \geq 0$ and $\gamma \in (0, 1)$, let

$$f(x; T, \gamma) := 2x - (1 + \gamma) - T \ln\left(\frac{x}{1-x}\right), \quad \forall x \in (0, 1) \quad (13)$$

Whenever obvious from the context, we will simply write $f(x)$.

The main result of this section is that for each $\gamma \in (0, 1)$, there exists a unique critical temperature, $T_c(\gamma)$ (or simply T_c), such that the number of steady states depends on the whether T is less than, equal to or larger than T_c . This is illustrated in Figures 2 and 3 and formally proved in Theorem 4.4. In particular, for $T < T_c(\gamma)$, there exist three QRE, which for $T = 0$ are precisely the Nash equilibria of the underlying evolutionary game. At the transition point, i.e., when $T = T_c(\gamma)$, the initial pure equilibrium, $x = 1$, and the unstable mixed equilibrium merge into an unstable equilibrium. For $T > T_c(\gamma)$, there is only one equilibrium, which is stable and which lies in the attracting region of desirable new equilibrium, $x = 0$. Additionally, as can be seen from Figure 2, the critical temperature T_c is always between $(0, 1/2)$ for the three depicted cases. In fact, for each $\gamma \in (0, 1)$, the critical temperature $T_c(\gamma) \in (0, 1/2)$ can be determined analytically. This is the statement of Lemma 4.3.

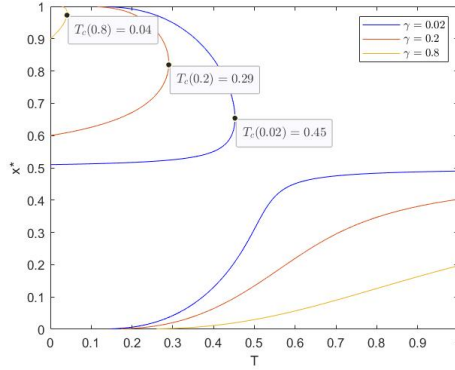


Fig. 2. The QRE correspondence for $\gamma = 0.02$ (blue line), $\gamma = 0.2$ (red line) and $\gamma = 0.8$ (green line). For $T < T_c(\gamma)$, and in particular for $T = 0$, there are three steady states. Starting from the top-left corner – population state $x = 1$ at which PoW is fully adopted – and as T increases, there is a critical point, $T_c(\gamma)$ at which the two upper states merge. After this point, i.e., as T increases above $T_c(\gamma)$, the system is subjected to an abrupt change: the two upper QRE disappear and the system retains only one steady state (branches at bottom right). From QRE with $x > 1/2$ when $T \leq T_c(\gamma)$, the population equilibrates at states with $x < 1/2$, for all $T > T_c(\gamma)$. The transitional point at which the critical mass is reached, is precisely $T_c(\gamma)$.

LEMMA 4.3. *Let $T \in (0, 1/2]$. Then, for any $\gamma \in (0, 1)$, the equation*

$$\sqrt{1 - 2T} - \gamma - T \cdot \ln \left(\frac{1 + \sqrt{1 - 2T}}{1 - \sqrt{1 - 2T}} \right) = 0 \tag{14}$$

has a unique solution $T_c(\gamma)$ or simply $T_c \in (0, 1/2)$.

PROOF. See Appendix A.1. □

Given T_c , we can now state and prove the main result of this section in Theorem 4.4. We will also use the notation

$$x_{l,u}(T) := \frac{1}{2} \left(1 \pm \sqrt{1 - 2T} \right), \quad \text{for } T \in (0, 1/2]. \tag{15}$$

Also, whenever obvious from the context, we will omit the dependence of $x_{l,u}(T)$ and of the steady states $x(T)$ on T and write $x_{l,u}$ and x , respectively

THEOREM 4.4. *Let $\gamma \in (0, 1)$ and $T > 0$ and let $T_c(\gamma)$ denote the critical temperature as given by (14). Then, the steady states of the population or equivalently the Quantal Response Equilibria (QRE) of the Q-learning dynamical system*

$$\dot{x} = x(1 - x) \left[2x - (1 + \gamma) - T \ln \left(\frac{x}{1 - x} \right) \right]$$

are given as follows

- $T < T_c(\gamma)$: 3 steady states x_1, x_2, x_3 , with $x_1 \in (0, x_l)$, $x_2 \in (1/2, x_u)$ and $x_3 \in (x_u, 1)$.
- $T = T_c(\gamma)$: 2 steady states x_1, x_2 , with $x_1 \in (0, x_l)$ and $x_2 = x_u$.
- $T > T_c(\gamma)$: 1 steady state x_1 , with $x_1 \in (0, x_l)$ when $T < 1/2$ and $x_1 \in (0, 1/2)$ when $T \geq 1/2$.

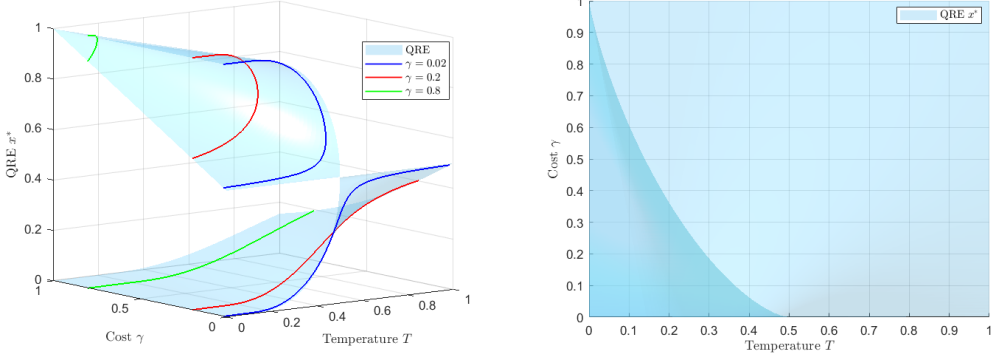


Fig. 3. The QRE correspondence for all possible values of $\gamma \in (0, 1)$ (left panel) and its projection on the (γ, T) plane (right panel). To understand the left panel, the (T, x^*) -slices at the $\gamma = 0.02, 0.2$ and 0.8 levels correspond precisely to the blue, red and green line, respectively, in Figure 2. In the right panel, darker areas correspond to the locations of multiple QRE. Abrupt transitions (bifurcations) occur at the middle boundary between the dark and light region.

where $x_{l,u}$ are given by (15). In all cases, the steady state $x_1 \in (0, x_l)$ is stable. For $T < T_c(\gamma)$, steady states $x_1 \in (0, x_l)$ and $x_3 \in (x_u, 1)$ are stable whereas x_2 is not. In particular, for $T = T_c(\gamma)$, steady state $x_2 = x_u$ is unstable.

The proof of Theorem 4.4 is split into two steps that are presented next. Intuitively, as mentioned above, the steady states of the Q-learning dynamics and their stability properties are determined by the roots and the sign of function $f(x; T, \gamma) = 2x - (1 + \gamma) - T \ln\left(\frac{x}{1-x}\right)$ for any starting point $x \in (0, 1)$ that is given in (13). These are resolved next.

LEMMA 4.5. Let $\gamma \in (0, 1)$ and let $T_c(\gamma) \in (0, 1/2)$ as given by Lemma 4.3. Then, for $x_{l,u} := \frac{1}{2} \left(1 \pm \sqrt{1 - 2T}\right)$, cf. equation (15), it holds that $0 < x_l < 1/2 < x_u < 1$, and the number of the solutions of the equation $f(x; T, \gamma) = 0$ with

$$f(x; T, \gamma) = 2x - (1 + \gamma) - T \ln\left(\frac{x}{1-x}\right)$$

depends on the value of $T > 0$ as follows

- $0 < T < T_c(\gamma)$: 3 solutions x_1, x_2, x_3 , with $x_1 \in (0, x_l)$, $x_2 \in (x_l, 1/2)$ and $x_3 \in (x_u, 1)$.
- $T = T_c(\gamma)$: 2 solutions x_1, x_2 , with $x_1 \in (0, x_l)$ and $x_2 = x_u$.
- $T > T_c(\gamma)$: 1 solution x_1 , with $x_1 \in (0, x_l)$ when $T < 1/2$ and $x_1 \in (0, 1/2)$ when $T \geq 1/2$.

PROOF. See Appendix A.1. □

Using Lemma 4.5, it is now immediate to determine the convergence properties of the Q-learning dynamical system \dot{x} and hence, prove Theorem 4.4.

PROOF OF THEOREM 4.4. The existence of the steady states in the three cases has been established in Lemma 4.5. Hence, it remains to prove the claims about their stability. The dynamics defined by \dot{x} are 1-dimensional and hence their convergence properties and the

stability of their steady states can be fully determined by the sign of \dot{x} . Since $x(1-x) > 0$ for any $x \in (0, 1)$, the sign of \dot{x} fully depends on $f(x; T, \gamma)$. In turn, the sign of $f(x; T, \gamma)$ for any $x \in (0, 1)$ has already been determined by the calculations in the proof of Lemma 4.5 and in particular by equation (19). Formally, we have the following cases

- $T < T_c(\gamma)$. In this case, the dynamics \dot{x} have three steady states $x_1, x_2, x_3 \in (0, 1)$ that correspond to the respective roots of function $f(x; T, \gamma)$. In particular, it holds that $0 < x_1 < x_l < 0.5 < x_2 < x_u < x_3 < 1$ and the sign of \dot{x} starts positive and alternates accordingly. This gives the stability results in Figure 5.
- $T = T_c(\gamma)$. At this point, the two roots that are larger than $1/2$, namely x_2 and x_3 , merge to one root precisely at x_u . The critical observation is that this new steady state is unstable, since the dynamics \dot{x} have a negative sign at both sides of the root x_u . This is shown in Figure 6.
- $T > T_c(\gamma)$. In this case, $f'(x; T, \gamma) < 0$ which implies that the dynamics \dot{x} are decreasing for any $x \in (0, 1)$. Since $f(x; T, \gamma)$ starts positive and ends up negative, there remains only one root (steady state), x_1 of $f(x; T, \gamma)$, which is strictly less than $1/2$. An illustration is given in Figure 7.

□

Note that the case $T = 0$ which was separately treated in Lemma 4.1, can also be derived as a special case of $T < T_c(\gamma)$ in Theorem 4.4 in which $x_1 = 0$, $x_2 = \frac{1}{2}(1 + \gamma)$ and $x_3 = 1$. The stability considerations remain the same, i.e., x_1 and x_3 are stable, whereas x_2 is unstable. Summing up, the statements of Theorem 4.4 and Lemma 4.1 are illustrated in Figure 8.

4.1 Catastrophe by Design and Hysteresis Effects

Theorem 4.4 implies that at the critical level or *tipping point*, $T_c(\gamma)$, the system undergoes an abrupt change, also called a *fold bifurcation*, [Strogatz, 2000]. Starting from $T = 0$ (no control) and $x = 1$ (at which PoW is adopted by all miners) and gradually increasing T , the population of agents (miners) follows the stable path up to $T_c(\gamma)$, at which the stable and unstable paths meet and cancel each other out. At this tipping point, the population of miners makes a sudden transition from a stable state with $x > (1 + \gamma)/2$ to a new stable state with $x < 1/2$. This abrupt is also catastrophe, but in this case it can be leveraged to design the destabilization of an undesirable state in favor of a desirable one. With T still in the control of the designer, it remains to explore what will happen to the population dynamics when T is reset gradually or even abruptly back to initial level, $T = 0$.

The answer is that the population will now converge to the new equilibrium at $x = 0$, i.e., it will fully adopt the new technology and is formally stated in Theorem 4.6. This is due to the *hysteresis mechanism* that is created due to the bifurcation: the dynamics have a *memory*, the current state, and their convergence properties differ depending on whether they start in the attracting region of $x_1 = 0$ or $x_3 = 1$, the two stable equilibria for $T = 0$. As illustrated in Figure 8, starting at any point $x < 1/2$, the dynamics will converge to $x_1 = 0$, as T goes to 0. From the perspective of mechanism design, this implies that if the designer can move the system to any such point, e.g., by temporarily increasing the temperature above the critical level $T_c(\gamma)$ and allowing the dynamics to stabilize at the unique remaining QRE, then they can directly (or gradually) reset the control back to 0

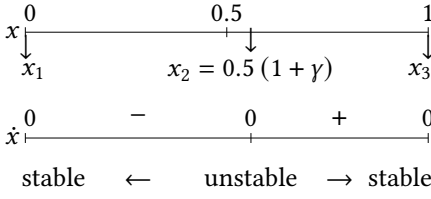
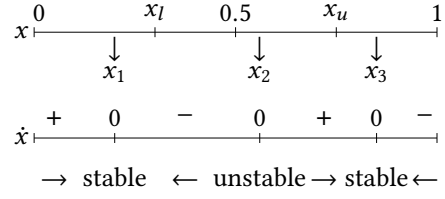
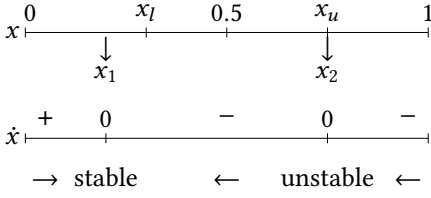
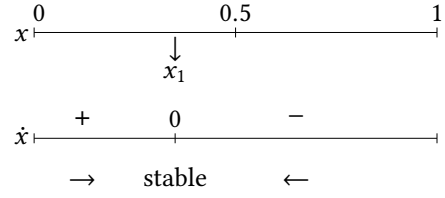
Fig. 4. $T = 0$.Fig. 5. $0 < T < T_c(\gamma)$.Fig. 6. $T = T_c(\gamma)$.Fig. 7. $T > T_c(\gamma)$.

Fig. 8. Stability properties of the Q-learning dynamics in the population for all values of $T \geq 0$. The stability of a steady state (QRE) x , (upper axis), is determined by the sign of \dot{x} , (lower axis). The stability properties change before, at, and after the critical temperature $T_c(\gamma)$, at which the bifurcation occurs. The case $T = 0$ (upper left panel) can be also treated as a special case of $0 < T < T_c(\gamma)$ (upper right panel), for any $\gamma \in (0, 1)$.

and still have theoretically provable guarantees that the population state will converge to $x_1 = 0$.

Precisely this hysteresis effect that emerges in the aftermath of the fold bifurcation is leveraged here to design a controlled mechanism – or a *controlled catastrophe* – that will disrupt an undesirable (yet stable, and currently prevailing) equilibrium to a desired outcome. It is immediate from the above, that the control needs to be only exercised temporarily. Importantly, this limits the expenses – assuming that it is costly to control T – to implement such a mechanism in practice. The above processes are intuitively illustrated in Figure 9 and Figure 1.

THEOREM 4.6. *Let $\gamma \in (0, 1)$ be fixed and for any $T \geq 0$, let $x_1(T)$ denote the stable steady state of the Q-learning dynamics \dot{x} in $[0, 1/2)$. Then, $x_1(T)$ is continuous and decreasing in T with $\lim_{T \rightarrow 0^+} x_1(T) = 0$. In particular, starting from any initial point $x_0 < 1/2$, and letting $T \rightarrow 0$, the Q-learning dynamics \dot{x} will stabilize at the steady state $x_1(0) = 0$, which corresponds to less costly and socially desirable technology.*

PROOF. See Appendix A.1. □

To conclude, note that the initial normalization of VK to 1, cf. equation (12), is equivalent to dividing equation (8) with VK . Hence, for practical purposes, the derived thresholds for γ should be interpreted as thresholds for γ/VK and similarly, the thresholds for T as thresholds for T/VK . For instance, $T_c(\gamma) < 1/2$ for any $\gamma \in (0, 1)$ implies that in applications with $VK \neq 1$, typically $VK \gg 1$, the tipping point satisfies $T_c < VK/2$. This provides an

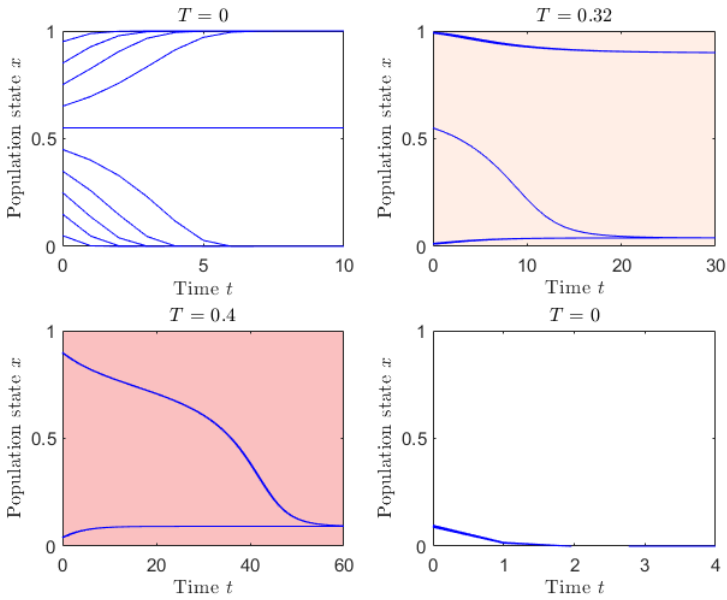


Fig. 9. The process of controlled heating a system then cooling it back to its initial temperature to destabilize certain (undesired) equilibria via *hysteresis mechanisms*. The upper left panel shows the evolution of the population state for 10 initial starting points and temperature $T = 0$. The QRE dynamics – in this case, equivalent to the replicator dynamics – converge to the two stable equilibria of the system. There is a third equilibrium which is unstable: the only trajectory that converges to it, is the one that precisely starts from it (horizontal line slightly above 0.5). In the upper right panel, the system temperature has been increased (but remains below the critical level). The starting points are the possible steady states for $T = 0$ (upper left panel). There are two stable equilibria, now in the interior of the admissible region, $x \in (0, 1)$. In the bottom left panel, the temperature is further increased, above the critical level. Starting again from the endpoints of the previous panel, the stability of the system changes abruptly: there is only remaining equilibrium and the population converges to it independently of the starting point. This lies in the attracting region of the desired equilibrium, which can be recovered by cooling the system back to its initial temperature, $T = 0$.

upper bound on the extent of control that needs to be exercised to the system in monetary terms. For instance, if T is interpreted as taxation, this implies that the critical level of taxation will never exceed – in fact, it will be much less than – half of the total investment K times the generated from the investment, aggregate value V .

5 STRUCTURAL ROBUSTNESS OF THE MODEL

The proposed mechanism is, critically, not tied to the current modelling and working assumptions. Thus, even if we err about our understanding of the system – which is very much likely in the particular blockchain setting that we examine or more generally in systems with constantly changing characteristics – our results can still be appealing. The goal of this section is to test and prove this claim in two directions.

First, we show that the proposed catastrophe and hysteresis mechanism can absorb state dependent perturbations in the Q-learning dynamics without significant impact in its actual implementation, cf. Section 5.1. This holds for any reasonably – so that the model parameters remain in their admissible regions – bounded perturbations.

Second, we study the robustness of the model for different values of the technology parameter α , cf. Equation (2). Different values of α capture different relationships between the degree of adoption of each technology and the total generated value for the population. In this direction, we obtain a (conservative, yet tight for extreme values of the parameters) bound on the control – critical value of T – that needs to be exercised to the system to trigger the desired disruption of its initial state. The mechanics of the proposed mechanism essentially remain unaffected. In practical terms, as α increases, a network split becomes more damaging and the population reaches the tipping point for increasingly lower values of T_c . In this respect, the case $\alpha = 2$ that we treated in the previous part is the most costly to destabilize.

5.1 State Dependent Perturbations

In general, small perturbations in a dynamical system can have major effects in its outcome, cf. [Puu, 1991, Strogatz, 2000] or [Palaiopanos et al., 2017] among many others. To study the behavior of the current dynamical system in equation (8), we consider the dynamics in (12), and add a noise term $\epsilon(x)$ that can be state dependent

$$\dot{x} = x(1-x) \left[u(W, x) - u(S, x) - T \ln \left(\frac{x}{1-x} \right) + \epsilon(x) \right]$$

Accordingly, let $f_\epsilon(x; T, \gamma) := 2x - 1 - \gamma - T \ln \left(\frac{x}{1-x} \right) + \epsilon(x)$. Under the perturbation, the number of the interior steady states of the dynamics, as expressed by the roots of $f_\epsilon(x)$, may change significantly in comparison to the unperturbed system. In particular, assume that x^* is a steady state of the unperturbed system for some $T > 0$, then

$$f_\epsilon(x^*; T, \gamma) := \left[2x^* - 1 - \gamma - T \ln \left(\frac{x^*}{1-x^*} \right) \right] + \epsilon(x^*) = \epsilon(x^*)$$

Accordingly, there exists a neighborhood around x^* for which the system becomes unpredictable. For $\epsilon(x)$ arbitrary, it is impossible to argue about the exact state of the system within this region. However, perturbations that should retain the value of $\gamma + \epsilon(x) \in [0, 1]$ (larger perturbations change the nature of the system in favor of the originally undesirable equilibrium and are hence irrelevant) impose the natural upper bound $\epsilon_0 \in (0, \min\{\gamma, 1 - \gamma\})$ on $|\epsilon(x)|$. The key intuition is that in this case, we can still argue about the behavior of the system *outside* these regions. In fact, the stability analysis, cf. Theorem 4.4 and Figure 8, carries through, with proper neighborhoods instead of the steady states alone.

THEOREM 5.1. *Let $\gamma \in (0, 1)$ and $T \geq 0$. Also let $f(x; T, \gamma) = 2x - 1 - \gamma - T \ln \left(\frac{x}{1-x} \right)$ and let $f_\epsilon(x) := f(x; T, \gamma) + \epsilon(x)$ for $x \in (0, 1)$ where $|\epsilon(x)| \leq \epsilon_0$ for some constant $\epsilon_0 \in (0, \min\{\gamma, 1 - \gamma\})$. Finally, let $x^*(T; \gamma) = \min\{x : f(x; T, \gamma) = 0\}$ denote the minimum QRE of $f(x; T, \gamma)$. Then, it holds that*

- *If $T > 0$, and $x \in (0, x^*(T; \gamma + \epsilon))$, then $f(x; T, \gamma) > 0$.*

- If $T > T_c(\gamma - \epsilon_0)$, and $x \in (x^*(T; \gamma - \epsilon), 1)$, then $f(x; T, \gamma) < 0$.
- If $T < T_c(\gamma - \epsilon_0)$, and $x \in (x^*(T; \gamma - \epsilon), \frac{1}{2})$, then $f(x; T, \gamma) < 0$.

Moreover, deviation between the QRE of the perturbed and the original system admits the following bound

$$|x^*(T; \gamma - \epsilon) - x^*(T; \gamma + \epsilon)| \leq \min \left\{ \frac{\epsilon}{|1 - 2T|}, \frac{1}{2} \right\}$$

Remark 1. From the mechanism design perspective, it is also meaningful to study the effects of the perturbation on the implementation of the proposed mechanism and in particular, on the control that need to be exercised on T . To quantify this change, we solve equation $f_\epsilon(x) = 0$ for T , and recover the temperature as a function of x on the geometric locus of all QRE of the unperturbed system

$$T(x; \gamma) = (2x - 1 - \gamma) \left(\ln \left(\frac{x}{1-x} \right) \right)^{-1}$$

Accordingly, let $T_\epsilon(x)$ denote the temperature on the geometric locus of all QRE of the perturbed system. Then,

$$T_\epsilon(x; \gamma) = \frac{2x - 1 - \gamma + \epsilon(x)}{\ln \left(\frac{x}{1-x} \right)} = \frac{2x - 1 - c}{\ln \left(\frac{x}{1-x} \right)} + \frac{\epsilon(x)}{\ln \left(\frac{x}{1-x} \right)} = T(x; \gamma) + \frac{\epsilon(x)}{\ln \left(\frac{x}{1-x} \right)} \quad (16)$$

which implies that if some $x \in (0, 1)$ is a QRE of both systems, then

$$|T_\epsilon(x; \gamma) - T(x; \gamma)| \leq \epsilon_0 \cdot \left| \ln \left(\frac{x}{1-x} \right) \right|^{-1} \quad (17)$$

This reduces the study of the hysteresis mechanism and the stability properties of the dynamics in the perturbed system to the study of the worst case perturbation ϵ_0 . Equation (17) highlights an important property that comes from the inclusion of the entropy term in the dynamics. Namely, as x approaches the boundary, the effect of any bounded perturbation vanishes, as this terms gets dominated by the entropy. An illustration is given in Figure 10.

The main takeaway of this part is that the results of the unperturbed case, largely carry over also to the perturbed case and, intuitively, this rests on two concurrent effects. First, while the exact number of steady states in the perturbed system cannot be determined, all new steady states are all located in some bounded neighborhood around the old steady states. Outside these regions, the sign of \dot{x} remains the same as in the unperturbed case and hence, the dynamics converge to these regions in the same fashion that they converged to the exact steady states in the unperturbed case.

Second, and most important, since the perturbation term is divided by $|\ln \left(\frac{x}{1-x} \right)|$ (see equation (16) above), the change in the QRE of the perturbed system becomes negligible for values of x close to 0 and close 1. This implies, when T is reduced back to its initial level (after it has been raised above the critical level to cause the bifurcation in the dynamics), the perturbed dynamics will converge *exactly* to the desired equilibrium. Two illustrations are given in Figure 10.

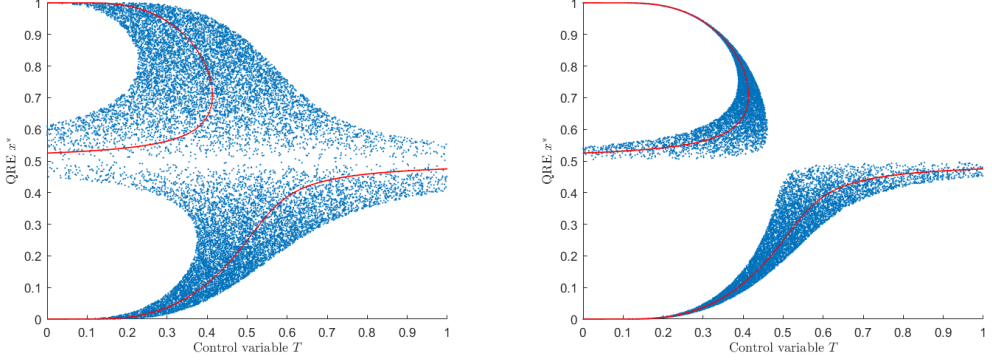


Fig. 10. QRE (blue dots) under state-dependent perturbations $\epsilon(x) = -\text{rand} \cdot \ln(x - x^{0.5})$ (left panel) and $\epsilon(x) = \text{rand} \cdot x^{1/6} (1 - x)$ (right panel), where rand denotes a random number in $[0, 0.1]$. In both cases, $\gamma = 0.05$. The perturbation in the left panel does not satisfy our working assumptions. Even if T is arbitrarily increased, the system may not equilibrate in the attracting region of $x = 0$. By contrast, in the right panel, the condition is satisfied. The largest deviations occur for population states close to $1/2$ whereas for values of x close to the boundary, the perturbed system reduces to the original system (red line). The memory encoding (or hysteresis mechanism) can be implemented with the only difference that, now, the dynamics can converge to any blue dot (instead of the red line) that lies in a neighborhood of the original QRE (in the unperturbed system). However, starting in the attracting region of $x = 0$, i.e., in $(0, 1/2)$, the dynamics will converge to the desired equilibrium as the perturbation gets increasingly dominated by the entropy term.

5.2 Technology Adoption and Generated Value

Parameter α in Equation (2) captures the value that can be created by each technology as a function of its adoption by the population of investors (miners). Values of $\alpha < 1$ indicate subadditive value, i.e., that the population payoff is maximized when the network splits between the two technologies. In this case, there exists an optimal split which is the unique stable state at which the population will stabilize. Such cases fall out of the present scope. Similarly, the case $\alpha = 1$ implies that the degree of adoption does not affect the generated value. In this case, adoption of the less costly technology constitutes a dominant strategy and the resolution of the game is trivial.

In the present context, we are interested in cases with superadditive value that are expressed by values of $\alpha > 1$ (an illustration is provided in Remark 3). Thus far, we have assumed that $\alpha = 2$ mainly for expositional purposes. In this part, we show that the results generalize essentially unaltered to any $\alpha \geq 1$ and hence, that the proposed control mechanism of Section 4.1 can be applied regardless of the underlying relationship. It is worth noting that $\alpha = 2$ corresponds in some sense to the most difficult case to treat from a practical perspective. Larger values of α capture networks for which splits are more detrimental. Accordingly, in such cases, it suffices to cost a small split in the network – i.e., an increase in $1 - x$ – to achieve the catastrophe. In practice, this translates to smaller values of the tipping point T and hence, to a less costly implementation of the proposed mechanism. These results are formalized next. Let $\alpha \geq 1$ arbitrary. Then, using equation

(3) with equation (10), we obtain the relationship

$$\dot{x} = x(1-x) \left[VK^{\alpha-1} x^{\alpha-1} - \gamma - VK^{\alpha-1} (1-x)^{\alpha-1} - T \ln \left(\frac{x}{1-x} \right) \right]$$

After normalizing $VK^{\alpha-1}$ to 1 – or equivalently after dividing the equation with $VK^{\alpha-1}$ and setting $\gamma \rightarrow \gamma/VK^{\alpha-1}$ and $T \rightarrow T/VK^{\alpha-1}$ – this yields the dynamics

$$\dot{x} = x(1-x) \left[x^{\alpha-1} - \gamma - (1-x)^{\alpha-1} - T \ln \left(\frac{x}{1-x} \right) \right] \quad (18)$$

To proceed, let $f_\alpha(x; T, \gamma) := x^{\alpha-1} - \gamma - (1-x)^{\alpha-1} - T \ln \left(\frac{x}{1-x} \right)$. To ease the proofs, we restrict attention to integer $\alpha \geq 1$, however the results apply for any $\alpha \geq 1$. The continuous case requires some more technical steps and since it does not add much intuition, it is omitted.

The main result in this direction is that for any $\alpha \geq 1$, there exists a unique QRE $x^* \in (0, 1/2)$, whenever $T \geq 1/2$. This implies that by increasing T to at most $1/2$, the system can be stabilized in a state x^* which lies in the attracting region of the desired equilibrium $x = 0$. Subsequently, T can be reset back to 0 and the population is theoretically guaranteed to converge to $x = 0$. This is formalized in Theorem 5.2

THEOREM 5.2. *Let $\gamma \in (0, 1)$ and $\alpha \geq 1$. Then, the Q-learning dynamics that describe the fraction x of the PoW miners in the population*

$$\dot{x} = x(1-x) \left[x^{\alpha-1} - \gamma - (1-x)^{\alpha-1} - T \ln \left(\frac{x}{1-x} \right) \right]$$

have a unique steady state (QRE), $x^ \in (0, 1/2)$ whenever $T \geq 1/2$. For $T = 0$, and any $\alpha > 1$, there are 3 equilibria⁷, and $[0, 1/2]$ lies in the attracting region of the socially beneficial equilibrium $x = 0$.*

The statement of Theorem 5.2 is illustrated in Figure 11. Although uniform, the bound $T = 1/2$ is extremely conservative, and essentially it is tight (or close to tight) only for the absolutely extreme cases with γ close to 0 and $\alpha = 1$. As α grows, i.e., as a network split becomes more and more detrimental, the system becomes more sensitive to T – a small split is enough to cause a catastrophe – and the unique x^* lies much closer to the origin. This can be seen from the thinning upper part in Figure 11. However, as can be seen from the lower stable part in the plots of Figure 11, the behavior of the dynamics may change as α increases for low values of T and x . For instance, when $\alpha = 8$ (red line), there exists an unstable part (s-shaped part of the bottom red line) at which the population may oscillate between three or marginally two QRE. However, the key is that all these equilibria lie in the attracting region of $x = 0$ and hence, if this instability will not imperil the outcome of the proposed control mechanism. More importantly, this instability part can be avoided by an abrupt change of T back to 0.

In sum, the mechanism of increasing T above the critical temperature (which is less or equal than $1/2$) and then reducing it back to zero, works essentially unaltered for any type of relationship between the technology and its adoption by the network as expressed by α . When splits are highly destructive for the generated value, i.e., for large values of α , the

⁷For $\alpha = 1$, there are two equilibria and the socially desirable one trivially dominates the socially undesirable.

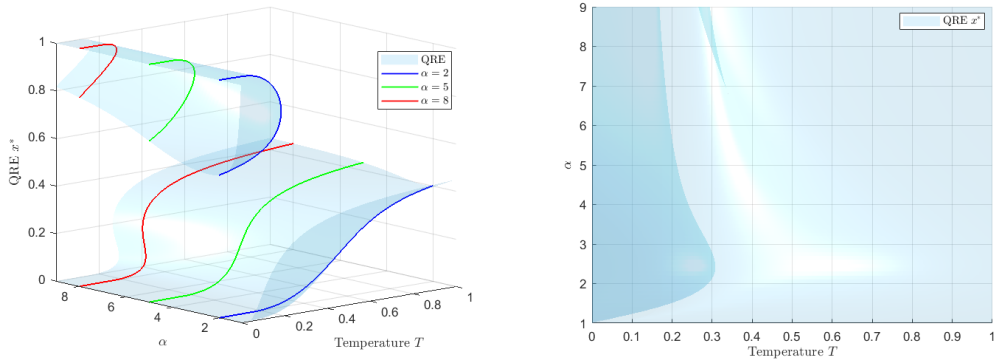


Fig. 11. The QRE correspondence for various values of α (left panel) and its projection on the (α, T) plane (right panel). In all cases, $\gamma = 0.2$, however, qualitatively equivalent plots can be generated for any value of $\gamma \in (0, 1)$. The case $\alpha = 2$ (blue slice) has been treated in Section 4. As α increases, the bottom stable part of the QRE correspondence may develop an unstable region: s-shaped part in the green (less obvious) and red (clearly visible) slices and darker thin part – darker parts correspond to more QRE – at the top of the right panel. Concerning the implementation of the controlled catastrophe and memory encoding (or hysteresis mechanism) to destabilize the undesired equilibrium $x = 1$, this instability is inconsequential: all new equilibria lie in the attracting region, $x < 1/2$ of the desired equilibrium $x_1 = 0$ for $T = 0$. Hence, starting in the upper branches, at $x_3 = 1$ and increasing T above the critical temperature, $T_c \leq 1/2$, the system will stabilize in some (unique) QRE x^* with $x^* < 1/2$. Subsequently, when resetting T back to zero, the population will converge to $x = 0$ by moving along the stable points of the bottom branches.

bifurcation occurs at increasingly smaller values of T , implying that the mechanism is less costly to implement.

REFERENCES

- Nick Arnosti and S. Matthew Weinberg. 2018. Bitcoin: A Natural Oligopoly. In *10th Innovations in Theoretical Computer Science Conference (ITCS 2019) (Leibniz International Proceedings in Informatics (LIPIcs))*, Avrim Blum (Ed.), Vol. 124. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, Dagstuhl, Germany, 5:1–5:1. <https://doi.org/10.4230/LIPIcs.ITCS.2019.5>
- Ido Bentov, Ariel Gabizon, and Alex Mizrahi. 2016. Cryptocurrencies Without Proof of Work. In *Financial Cryptography and Data Security*, Jeremy Clark, Sarah Meiklejohn, Peter Y.A. Ryan, Dan Wallach, Michael Brenner, and Kurt Rohloff (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 142–157. https://doi.org/10.1007/978-3-662-53357-4_10
- G. Bissias, B. N. Levine, and D. Thibodeau. 2019. Greedy but Cautious: Conditions for Miner Convergence to Resource Allocation Equilibrium. [arXiv:cs.GT/1907.09883](https://arxiv.org/abs/1907.09883)
- Jonah Brown-Cohen, Arvind Narayanan, Alexandros Psomas, and S. Matthew Weinberg. 2019. Formal Barriers to Longest-Chain Proof-of-Stake Protocols. In *Proceedings of the 2019 ACM Conference on Economics and Computation (Phoenix, AZ, USA) (EC '19)*. Association for Computing Machinery, New York, NY, USA, 459–473. <https://doi.org/10.1145/3328526.3329567>
- V. Buterin, D. Reijbergen, S. Leonardos, and G. Piliouras. 2019. Incentives in Ethereum’s Hybrid Casper Protocol. In *2019 IEEE International Conference on Blockchain and Cryptocurrency (ICBC)*. IEEE, USA, 236–244. <https://doi.org/10.1109/BLOC.2019.8751241>
- X. Chen, C. Papadimitriou, and T. Roughgarden. 2019. An Axiomatic Approach to Block Rewards. In *Proceedings of the 1st ACM Conference on Advances in Financial Technologies (Zurich, Switzerland) (AFT '19)*. ACM, New

- York, NY, USA, 124–131. <https://doi.org/10.1145/3318041.3355470>
- Digiconomist. 2020. Bitcoin Energy Consumption Index. Available [online]. [Accessed: 31-01-2020].
- I. Eyal and E. G. Sirer. 2018. Majority is Not Enough: Bitcoin Mining is Vulnerable. *Commun. ACM* 61, 7 (June 2018), 95–102. <https://doi.org/10.1145/3212998>
- A. Fiat, A. Karlin, E. Koutsoupias, and C. Papadimitriou. 2019a. Energy Equilibria in Proof-of-Work Mining. In *Proceedings of the 2019 ACM Conference on Economics and Computation* (Phoenix, AZ, USA) (EC '19). ACM, New York, NY, USA, 489–502. <https://doi.org/10.1145/3328526.3329630>
- A. Fiat, E. Koutsoupias, K. Ligett, Y. Mansour, and S. Olonetsky. 2019b. Beyond myopic best response (in Cournot competition). *Games and Economic Behavior* 113 (2019), 38–57. <https://doi.org/10.1016/j.geb.2013.12.006>
- J. Garay, A. Kiayias, and N. Leonardos. 2015. The Bitcoin Backbone Protocol: Analysis and Applications. In *Advances in Cryptology – EUROCRYPT 2015*, E. Oswald and M. Fischlin (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 281–310.
- G. Goren and A. Spiegelman. 2019. Mind the Mining. In *Proceedings of the 2019 ACM Conference on Economics and Computation* (Phoenix, AZ, USA) (EC '19). ACM, New York, NY, USA, 475–487. <https://doi.org/10.1145/3328526.3329566>
- Shihab S. Hazari and Qusay H. Mahmoud. 2019. Comparative evaluation of consensus mechanisms in cryptocurrencies. *Internet Technology Letters* 2, 3 (2019), e100. <https://doi.org/10.1002/itl2.100>
- J. Hofbauer and K. Sigmund. 1998. *Evolutionary Games and Population Dynamics*. Cambridge University Press, Cambridge, United Kingdom. <https://doi.org/10.1017/CBO9781139173179>
- Ardeshir Kianercy and Aram Galstyan. 2012. Dynamics of Boltzmann Q learning in two-player two-action games. *Phys. Rev. E* 85 (Apr 2012), 041145. Issue 4. <https://doi.org/10.1103/PhysRevE.85.041145>
- A. Kiayias, E. Koutsoupias, M. Kyropoulou, and Y. Tselekounis. 2016. Blockchain Mining Games. In *Proceedings of the 2016 ACM Conference on Economics and Computation* (Maastricht, The Netherlands) (EC '16). Association for Computing Machinery, New York, NY, USA, 365–382. <https://doi.org/10.1145/2940716.2940773>
- A. Kiayias, A. Russell, B. David, and R. Oliynykov. 2017. Ouroboros: A Provably Secure Proof-of-Stake Blockchain Protocol. In *Advances in Cryptology – CRYPTO 2017*, J. Katz and H. Shacham (Eds.). Springer International Publishing, Cham, 357–388.
- Yuri Kuznetsov. 2004. *Elements of Applied Bifurcation Theory, Third Edition*. Vol. 112. Springer-Verlag, New York. <https://doi.org/10.1007/978-1-4757-3978-7>
- R. D. McKelvey and T. R. Palfrey. 1995. Quantal Response Equilibria for Normal Form Games. *Games and Economic Behavior* 10, 1 (1995), 6–38. <https://doi.org/10.1006/game.1995.1023>
- S. Nakamoto. 2008. Bitcoin: A Peer-to-Peer Electronic Cash System. Available [online]. [Accessed: 31-01-2020].
- G. Palaiopoulos, I. Panageas, and G. Piliouras. 2017. Multiplicative Weights Update with Constant Step-Size in Congestion Games: Convergence, Limit Cycles and Chaos. In *Proceedings of the 31st International Conference on Neural Information Processing Systems* (Long Beach, California, USA) (NIPS'17). Curran Associates Inc., Red Hook, NY, USA, 5874–5884. <https://doi.org/10.5555/3295222.3295337>
- T. Puu. 1991. Chaos in duopoly pricing. *Chaos, Solitons & Fractals* 1, 6 (1991), 573–581. [https://doi.org/10.1016/0960-0779\(91\)90045-B](https://doi.org/10.1016/0960-0779(91)90045-B)
- William H. Sandholm. 2010. *Population Games and Evolutionary Dynamics*. The MIT Press, Cambridge, Massachusetts.
- S. H. Strogatz. 2000. *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering* (first ed.). Westview Press (Studies in nonlinearity collection), Cambridge, MA, USA.
- Ming Tan. 1997. *Multi-Agent Reinforcement Learning: Independent vs. Cooperative Agents*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 487–494.
- Karl Tuyls, Katja Verbeeck, and Tom Lenaerts. 2003. A Selection-mutation Model for Q-learning in Multi-agent Systems. In *Proceedings of the Second International Joint Conference on Autonomous Agents and Multiagent Systems* (Melbourne, Australia) (AAMAS '03). ACM, New York, NY, USA, 693–700. <https://doi.org/10.1145/860575.860687>
- University of Cambridge, Judge Business School. 2020. Cambridge Bitcoin Electricity Consumption Index. Available [online]. [Accessed: 31-01-2020].

- C. J.C.H. Watkins and P. Dayan. 1992. Technical Note: Q-Learning. *Machine Learning* 8, 3 (01 May 1992), 279–292. <https://doi.org/10.1023/A:1022676722315>
- David H. Wolpert, Michael Harré, Eckehard Olbrich, Nils Bertschinger, and Jürgen Jost. 2012. Hysteresis effects of changing the parameters of noncooperative games. *Phys. Rev. E* 85 (Mar 2012), 036102. Issue 3. <https://doi.org/10.1103/PhysRevE.85.036102>
- Ger Yang, Georgios Piliouras, and David Basanta. 2017. Bifurcation Mechanism Design - from Optimal Flat Taxes to Improved Cancer Treatments. In *Proceedings of the 2017 ACM Conference on Economics and Computation* (Cambridge, Massachusetts, USA) (EC '17). ACM, New York, NY, USA, 587–587. <https://doi.org/10.1145/3033274.3085144>

A APPENDIX

A.1 Additional Material and Omitted Proofs: Section 2

Remark 2. For games with general strategy sets, i.e., with possibly more than two strategies, the Q-learning agents update their strategies according to the following rule

$$\begin{aligned} \max \quad & \sum_{j=W,S} x_j Q_t(j) - T \sum_{j=W,S} x_j \ln x_j \\ \text{subject to:} \quad & \sum_{j=W,S} x_j = 1 \\ & x_j \geq 0, \quad \text{for } j = W, S. \end{aligned} \tag{S1'}$$

Since, in the current setting each agent has two strategies, we obtain the convenient representation in (S1). To intuitively explain the objective function in (S1'), observe that the first term, i.e., $\sum_{j=W,S} x_j Q_t(j)$ enforces maximization of the Q -values. Since it is linear in the x_j 's, it would simply choose (put full probability on) the strategy with the highest Q -value, if the second term (entropy) was missing. However, the introduction of the entropic term, i.e., of $-\sum_{j=W,S} x_j \ln x_j$, essentially requires from the agent to choose the distribution x with maximum entropy for every given weighted sum of the Q -values, and hence, to explore (assign positive probability on) suboptimal strategies.

The relative importance between maximization of Q -values and exploration of the strategy space, is controlled by the introduced parameter $T \geq 0$. Termed *temperature* in physics, T can be interpreted as a tuning parameter: as $T \rightarrow 0$, the agent always acts greedily and chooses the strategy corresponding to the maximum Q -value (pure exploitation), whereas as $T \rightarrow \infty$, the agent's chooses a strategy completely at random (pure exploration). In particular, for $T = 0$, the system reduces to the well known replicator (best response) dynamics which (under standard regularity assumptions that are met in the present model) recover the Nash equilibria of the underlying evolutionary game, cf. Section 2.1. For different values of $T > 0$, the resting points of the system change, sometimes in an abrupt way and this is precisely the intuition that we exploit here to design a mechanism that will stabilize the system in a desired state. In fact, as shown in [Yang et al., 2017], the temperature can be considered as a control parameter in the arsenal of a system designer. From the objective function, one discerns that parameter T essentially rescales all the Q -values in a multiplicative way. Hence, as shown in [Yang et al., 2017], T can be treated as a taxation parameter in economic systems or as a medically controlled substance in health related settings.

PROOF OF PROPOSITION 2.1. The resulting game

$$P = \begin{pmatrix} 1 - \gamma & -\gamma \\ 0 & 1 \end{pmatrix},$$

with $\gamma \in (0, 1)$ has three Nash equilibria: two pure, (W, W) with payoffs $(1 - \gamma, 1 - \gamma)$ and (S, S) with payoffs $(1, 1)$, and one (fully) mixed $\left(\left(\frac{1+\gamma}{2}, \frac{1-\gamma}{2}\right), \left(\frac{1+\gamma}{2}, \frac{1-\gamma}{2}\right)\right)$ with payoffs $\left(\frac{1-\gamma}{2}, \frac{1-\gamma}{2}\right)$. These correspond to population states $x_1 = 0$, $x_2 = \frac{1+\gamma}{2}$, and $x_3 = 1$. The corresponding payoffs can be also derived by substituting in the average payoff function, (4), which here becomes

$$\bar{u}(x) = 2x^2 - (2 + \gamma)x + 1$$

All three Nash equilibria are symmetric. The (S, S) (bottom right) Nash equilibrium is (strictly) payoff dominant, i.e., it *Pareto-dominates* the other two (recall that $\gamma \in (0, 1)$ is the cost from the current costly technology) and is also (strictly) *risk* dominant, since $0 + 1 > 1 - 2\gamma$. A symmetric mixed Nash equilibrium (x^*, x^*) is evolutionary stable if $u(x^*, x) > u(x, x)$ for all other mixed strategies $x \neq x^*$. Hence, the mixed Nash equilibrium is *not* evolutionary stable, since for any other $x \in (0, 1)$ with $x \neq \frac{1+\gamma}{2}$

$$\begin{aligned} u\left(\frac{1+\gamma}{2}, x\right) - u(x, x) &= \left(\frac{1+\gamma}{2} - x\right) [(1-\gamma)x + (-\gamma)(1-x)] + \left(\frac{1-\gamma}{2} - (1-x)\right) (1-x) \\ &= -\frac{1}{2} (2x - (1+\gamma))^2 < 0 \end{aligned}$$

By contrast, both pure strategy equilibria are evolutionary stable, since $u(W, W) = 1 - \gamma > 0 = u(S, W)$, and $u(S, S) = 1 > -\gamma = u(W, S)$. \square

A.2 Omitted Proofs: Section 4

PROOF OF LEMMA 4.3. Let $u := \sqrt{1 - 2T}$. Then, $T \in (0, 1/2]$ implies that $u \in [0, 1)$ and the transformation is one to one with inverse $T = \frac{1}{2}(1 - u^2)$. Hence, we need to show that $T_c(\gamma) = \frac{1}{2}(1 - u_c^2)$, where u_c is the unique root of the function

$$g_\gamma(u) := u - \gamma - \frac{1 - u^2}{2} \ln\left(\frac{1+u}{1-u}\right)$$

in $(0, 1)$. Note that $g_\gamma(u)$ is defined for any $u \in (-1, 1)$. The derivative of $g_\gamma(u)$ with respect to u is

$$\frac{d}{du} g_\gamma(u) = u \ln\left(\frac{1+u}{1-u}\right) > 0$$

for all $u \in (-1, 1)$. Hence, $g_\gamma(u)$ is strictly increasing with $\lim_{u \rightarrow -1^+} g_\gamma(u) = -1 + \gamma < 0$, $g_\gamma(0) = 0 - \gamma < 0$ and $\lim_{u \rightarrow 1^-} g_\gamma(u) = 1 - \gamma > 0$. Accordingly $g_\gamma(u)$ has precisely one root, $u_c \in (0, 1)$ which yields the unique solution of the equation in the statement of the Lemma, given by $T_c(\gamma) = \frac{1}{2}(1 - u_c^2) \in (0, 1/2)$. \square

PROOF OF LEMMA 4.5. For any $T > 0$ and $\gamma \in (0, 1)$, the function $f(x; T, \gamma)$ is continuous in $x \in (0, 1)$ with

$$\lim_{x \rightarrow 0^+} f(x) = +\infty, \quad f(1/2) = -\gamma, \quad \lim_{x \rightarrow 1^-} f(x; T, \gamma) = -\infty. \quad (19)$$

This implies that $f(x; T, \gamma)$ starts positive and ends up negative. The derivative of $f(x; T, \gamma)$ with respect to $x \in (0, 1)$ is $f'(x; T, \gamma) = 2 - \frac{T}{x(1-x)}$, with $f(x; T, \gamma) = 0$, if and only if $x^2 - x + \frac{T}{2} = 0$. This yields the critical points

$$x_{l,u} = \frac{1}{2} \left(1 \pm \sqrt{1 - 2T} \right)$$

Depending on the value of T relative to $T_c(\gamma)$, there are three cases for the number of the solutions of the equation $f(x; T, \gamma) = 0$ for $x \in (0, 1)$.

- $0 < T < T_c(\gamma)$. Since $T_c(\gamma) < 1/2$ for any $\gamma \in (0, 1)$ as shown in Lemma 4.3, we have that $1 - 2T > 0$ and it is immediate that $0 < x_l < 1/2 < x_u < 1$. Also,

$$\begin{aligned} f(x_l; T, \gamma) &= g_\gamma \left(-\sqrt{1 - 2T} \right) < g_\gamma(0) < 0 \\ f(x_u; T, \gamma) &= g_\gamma \left(\sqrt{1 - 2T} \right) > g_\gamma \left(\sqrt{1 - 2T_c(\gamma)} \right) = 0 \end{aligned}$$

where the inequalities follow from the strict monotonicity of $g_\gamma(u) \in (-1, 1)$ that was shown in the proof of Lemma 4.3. Hence, using equation (19), we have that $f(x; T, \gamma)$ has precisely one root $x_1 \in (0, x_l)$, one root $x_2 \in (1/2, x_u)$ and one root $x_3 \in (x_u, 1)$.

- $T = T_c(\gamma)$. As in the previous case, $T_c(\gamma) < 1/2$ implies that $0 < x_l < 1/2 < x_u < 1$. However, in this case, $f(x_u; T, \gamma) = g_\gamma \left(\sqrt{1 - 2T_c(\gamma)} \right) = 0$ and hence, using again equation (19), it follows that f has one root in $(0, x_l)$, and a second root in $(1/2, 1)$ which is precisely x_u .
- $T > T_c(\gamma)$. In this case,

$$f(x_u; T, \gamma) = g_\gamma \left(\sqrt{1 - 2T} \right) < g_\gamma \left(\sqrt{1 - 2T_c(\gamma)} \right) = 0$$

which implies that $f(x; T, \gamma)$ turns negative at some point $x_1 < x_l$ when $T < 1/2$ or $x_1 < 1/2$ when $T > 1/2$ (and x_l undefined) and remains negative thereafter since x_2 is a local maximum. Hence, this $x_1 \in (0, x_l)$ or $x_1 \in (0, 1/2)$ is also the unique root of $f(x; T, \gamma)$ in $(0, 1)$. □

The three cases of Lemma 4.5 are illustrated in Figure 12. In the depicted instantiation, $\gamma = 0.185$ which yields $T_c(\gamma) \approx 0.3$. The three curves correspond to $T = 0.25$, $T = T_c(\gamma) \approx 0.3$ and $T = 1$ which, in agreement with Lemma 4.5 give rise to 3, 2 and 1 solutions to the equation $f(x; T, \gamma) = 0, x \in (0, 1)$, respectively.

PROOF OF THEOREM 4.6. By Lemma 4.5, $f(x; T, \gamma) = 2x - 1 - \gamma - T \ln \left(\frac{x}{1-x} \right)$ has precisely one root $x_1(T)$ in $(0, x_l(T))$ for any value of $0 < T \leq T_c(\gamma) < 1/2$, where $x_l(T) = \frac{1}{2} \left(1 - \sqrt{1 - 2T} \right)$. Moreover, implicit differentiation of the function $f(x, T) := 2x(T) - 1 - \gamma - T \ln \left(\frac{x(T)}{1-x(T)} \right)$, with $f(x, T) = 0$, shows that $x(T)$ is strictly increasing in T for $x < 1/2$, since

$$\frac{\partial f(x, T)}{\partial T} \frac{dx}{dT} + \frac{\partial f(x, T)}{\partial T} \frac{dx}{dT} = -\ln \left(\frac{x}{1-x} \right) + \left(2 - \frac{x}{1-x} \right) \frac{dx}{dT} = 0$$

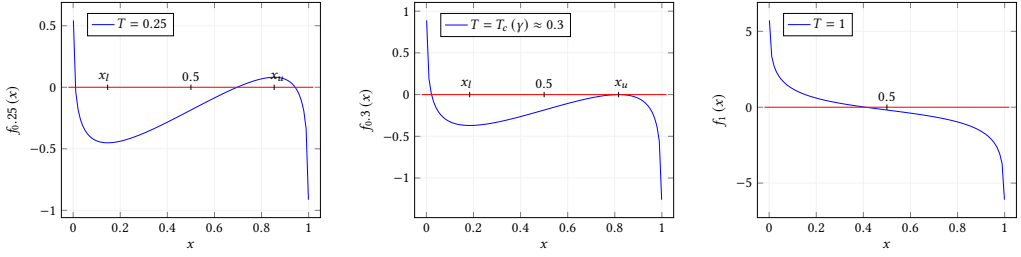


Fig. 12. The function $f(x; T, \gamma)$ for $\gamma = 0.185$ and $T = 0.25, T = T_c(\gamma) \approx 0.3$ and $T = 1$. For $T < T_c(\gamma)$, there are three steady states (roots of $f(x; T, \gamma)$), for $T = T_c(\gamma)$, there are precisely 2, and for $T > T_c(\gamma)$, only one. The smallest root is always less than $1/2$ (in particular less than $x_l = \frac{1}{2} \left(1 - \sqrt{1 - 2T}\right)$, if x_l exists), whereas the remaining ones (if any) are larger than $1/2$.

which yields

$$\frac{dx}{dT} = \frac{\ln\left(\frac{x}{1-x}\right)}{2 - \frac{T}{x(1-x)}} > 0$$

for $x < x_l(T)$ and $T < T_c(\gamma) < 1/2$. Hence, $0 < x_1(T) < x_l(T)$ implies that for any $\epsilon > 0$, there exists a $\delta > 0$ such that $x_1(T) < \epsilon$ for any $T < \delta$, since

$$\lim_{T \rightarrow 0^+} x_l = \lim_{T \rightarrow 0^+} \frac{1}{2} \left(1 - \sqrt{1 - 2T}\right) = 0$$

This implies, that $x_1(T) \rightarrow 0$ as $T \rightarrow 0^+$, which concludes the proof. \square

B APPENDIX

B.1 Omitted Results and Proofs: Section 5.1

LEMMA B.1. *The critical temperature $T_c(\gamma)$, which is the unique solution of the equation*

$$\sqrt{1 - 2T} - \gamma - T \cdot \ln\left(\frac{1 + \sqrt{1 - 2T}}{1 - \sqrt{1 - 2T}}\right) = 0$$

in $(0, 1/2)$, cf. Lemma 4.3, is decreasing in the cost parameter $\gamma \in (0, 1)$.

PROOF. For $\gamma \in (0, 1)$, the critical temperature $T_c(\gamma)$ is given by equation

$$\sqrt{1 - 2T} - \gamma - T \cdot \ln\left(\frac{1 + \sqrt{1 - 2T}}{1 - \sqrt{1 - 2T}}\right) = 0,$$

cf. Lemma 4.3. To proceed, let $F(\gamma, T) := \sqrt{1 - 2T} - \gamma - T \cdot \ln\left(\frac{1 + \sqrt{1 - 2T}}{1 - \sqrt{1 - 2T}}\right)$. Implicit differentiation of F with respect to γ , yields

$$0 = \frac{\partial F(\gamma, T)}{\partial \gamma} + \frac{\partial F(\gamma, T)}{\partial T} \cdot \frac{dT}{d\gamma} = -1 - \ln\left(\frac{1 + \sqrt{1 - 2T}}{1 - \sqrt{1 - 2T}}\right) \cdot \frac{dT}{d\gamma}$$

Hence

$$\frac{dT}{d\gamma} = - \left(\ln \left(\frac{1 + \sqrt{1 - 2T}}{1 - \sqrt{1 - 2T}} \right) \right)^{-1}$$

which is negative, since the argument of the \ln is larger than 1 for all $T > 0$. \square

LEMMA B.2. Let $f(x; \gamma, T)$ as in Definition 4.2 and let $f_\epsilon(x; T, \gamma) := f(x; T, \gamma) + \epsilon(x)$ where $\epsilon(x)$ denotes a noise term defined for $x \in (0, 1)$ such that $|\epsilon(x)| \leq \epsilon_0$ for some $\epsilon_0 \in (0, \min\{\gamma, 1 - \gamma\})$. Then, it holds that

$$f(x; T, \gamma + \epsilon_0) < f_\epsilon(x; T, \gamma) < f(x; T, \gamma - \epsilon_0)$$

PROOF. Since $\frac{\partial}{\partial \gamma} f(x; T, \gamma) = -1 < 0$, $f(x; T, \gamma)$ is decreasing in γ which implies that

$$f(x; T, \gamma + \epsilon_0) < f(x; T, \gamma) + \epsilon(x) = f_\epsilon(x; T, \gamma) < f(x; T, \gamma - \epsilon_0)$$

\square

Definition B.3. Given $T \geq 0$ and $\gamma \in (0, 1)$, we define $x^*(T; \gamma) = \min\{x : f(x; T, \gamma) = 0\}$. Note that by Theorem 4.4, $x^*(T; \gamma) < 1/2$ for any pair (T, γ) .

PROOF OF THEOREM 5.1. The stability part of the Proposition is derived in the same way as in Theorem 4.4 and follows directly from Lemma B.2 and the fact that $|\epsilon(x)| \leq \epsilon_0$. Our primary analysis indicates that $x^*(T, \gamma) < \frac{1}{2}$, $\forall T \in (0, \infty)$ and $\forall \gamma \in (0, 1)$; which immediately gives a most basic bound to the uncertain region $|x^*(T; \gamma - \epsilon) - x^*(T; \gamma + \epsilon)| < \frac{1}{2}$. However this bound can be tightened for specific values of T as we show in the rest of this section.

LEMMA B.4. Given $\gamma \in (0, 1)$ and $T \in [0, \infty)$, then $\forall x \in (0, 1)$

$$f'(x; T, \gamma) \in (-\infty, 2 - 4T)$$

PROOF. Let $x \in (0, 1)$. Then

$$\begin{aligned} f''(x; T, \gamma) \leq 0 &\implies \frac{T(1 - 2x)}{x^2(1 - x)^2} \leq 0 \\ &\implies 1 - 2x \leq 0 \\ &\implies x \geq \frac{1}{2} \end{aligned}$$

Hence, $\forall x \in (0, 1)$ it follows that $f'(x; T, \gamma) \leq f'(\frac{1}{2}) = 2 - 4T$

Furthermore, $\lim_{t \rightarrow 0^+} f'(t; T, \gamma) = \lim_{t \rightarrow 1^-} f'(t; T, \gamma) = -\infty$, which, by continuity implies that $f'(x; T, \gamma) \in (-\infty, 2 - 4T)$, $\forall x \in (0, 1)$

\square

Let $T > \frac{1}{2}$. Using Taylor's Approximation on $x^*(T; \gamma - \epsilon)$, it follows that $\forall x \in (0, 1)$ there exists some $\lambda(x) \in [\min\{x^*(T; \gamma - \epsilon), x\}, \max\{x^*(T; \gamma - \epsilon), x\}]$ such as

$$\begin{aligned} f(x; T, \gamma - \epsilon) &= f(x^*(T; \gamma - \epsilon); T, \gamma - \epsilon) + f'(\lambda(x); T, \gamma - \epsilon)(x - x^*(T; \gamma - \epsilon)) \\ &\implies f(x; T, \gamma) + \epsilon = f'(\lambda(x); T, \gamma - \epsilon)(x - x^*(T; \gamma - \epsilon)) \\ &\implies f(x; T, \gamma + \epsilon) = f'(\lambda(x); T, \gamma - \epsilon)(x - x^*(T; \gamma - \epsilon)) - 2\epsilon \end{aligned}$$

Hence, for $x = x^*(T; \gamma + \epsilon)$ and by Lemma B.4 we have

$$\begin{aligned} f(x^*(T; \gamma + \epsilon); T, \gamma + \epsilon) &= f'(\lambda(x^*(T; \gamma + \epsilon)); T, \gamma - \epsilon)(x^*(T; \gamma + \epsilon) - x^*(T; \gamma - \epsilon)) - 2\epsilon \\ \implies x^*(T; \gamma + \epsilon) - x^*(T; \gamma - \epsilon) &= \frac{2\epsilon}{f'(\lambda(x^*(T; \gamma - \epsilon)); T, \gamma - \epsilon)} \\ \implies |x^*(T; \gamma + \epsilon) - x^*(T; \gamma - \epsilon)| &= \frac{2\epsilon}{|f'(\lambda(x^*(T; \gamma - \epsilon)); T, \gamma - \epsilon)|} \\ \implies |x^*(T; \gamma + \epsilon) - x^*(T; \gamma - \epsilon)| &\leq \frac{\epsilon}{|1 - 2T|} \end{aligned}$$

As a final remark, note that $|x^*(T; \gamma + \epsilon) - x^*(T; \gamma - \epsilon)| < \epsilon$, $\forall T > 1$. In fact $|x^*(T; \gamma + \epsilon) - x^*(T; \gamma - \epsilon)| \rightarrow 0$ as $T \rightarrow \infty$. \square

B.2 Omitted Results and Proofs: Section 5.2

Remark 3. Parameter α expresses the value that is created by the technology in response to its adoption. In particular, there are three interesting cases, depending on whether α is smaller, larger than or equal to 1.

- $\alpha < 1$: Subadditive value. In this case, the total value generated from the two technologies is subadditive implying that a (perfect) split of the network is more beneficial for the society.
- $\alpha = 1$: Linear value. In this case, the aggregate value that is generated in increasing in the rate of adoption of the innovative (less costly) technology.
- $\alpha > 1$: Superadditive value. In this case, the aggregate value is (locally) maximized when either technology is fully adopted. This is the most interesting case in the current context.

Typical instantiations of these cases are depicted in the two panels of Figure 13.

LEMMA B.5. For any $\alpha \geq 2$, and any $x \in [0, 1]$, it holds that

$$x^\alpha (1 - x) + x (1 - x)^\alpha \leq \frac{1}{2\alpha}.$$

PROOF. Since $x \in [0, 1]$, we can rewrite the inequality as

$$x^\alpha (1 - x) + x (1 - x)^\alpha \leq \frac{x}{2\alpha} + \frac{1 - x}{2\alpha}$$

which is equivalent to

$$\left[x^\alpha (1 - x) - \frac{x}{2\alpha} \right] + \left[x (1 - x)^\alpha - \frac{1 - x}{2\alpha} \right] \leq 0$$

Hence, by symmetry, it suffices to show that $x^\alpha (1 - x) - \frac{x}{2\alpha} \leq 0$, for any $x \in [0, 1]$ and $\alpha > 2$. The last equation is equivalent to $x^{\alpha-1} (1 - x) \leq \frac{1}{2\alpha}$. By differentiating the left hand side with respect to x , we find that $\frac{d}{dx} (x^{\alpha-1} (1 - x)) = \frac{x^{\alpha-2}}{\alpha} (\frac{\alpha-1}{\alpha} - x)$ which is zero for $x = \frac{\alpha-1}{\alpha}$. Since $x^{\alpha-1} (1 - x)$ is equal to 0 for both $x = 0$ and $x = 1$, it attains a maximum at

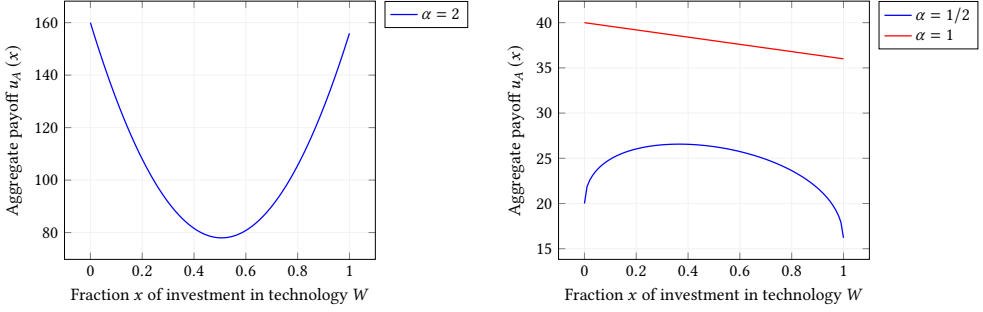


Fig. 13. Aggregate payoff $u_A(x)$ created by the population as a whole at population state $x \in [0, 1]$ (investment in technology W) for various values of parameter α . The aggregate payoff (vertical axis) is calculated by the formula: $u_A(x) = VK^\alpha \left[x^\alpha - \frac{\gamma}{VK^{\alpha-1}}x + (1-x)^\alpha \right]$, with $x \in [0, 1]$ and selected values of the parameters $V = 10, K = 4$ and $\gamma = 1$. For $\alpha = 2$ (left panel) and in general for $\alpha > 1$, the total aggregate value is (locally) maximized at the boundaries, i.e., when either technology is fully adopted (superadditive value). The global maximum is attained when the less costly technology S is fully adopted, i.e., when $x = 0$. By contrast, for $\alpha = 1/2$ (right panel, blue line), and in general for $\alpha < 1$, the aggregate wealth is maximized when the population is split between the two technologies (subadditive value). For $\alpha = 1$ (right panel, red line), the aggregate wealth is increasing in the adoption of the less costly technology S (linear value).

$x = \frac{\alpha-1}{\alpha}$ with value $\left(\frac{\alpha-1}{\alpha}\right)^{\alpha-1} \frac{1}{\alpha}$. Accordingly, it suffices to show that

$$\left(\frac{\alpha-1}{\alpha}\right)^{\alpha-1} \frac{1}{\alpha} \leq \frac{1}{2\alpha}$$

or equivalently that $\left(\frac{\alpha-1}{\alpha}\right)^{\alpha-1} \leq 1/2$ for any $\alpha \geq 2$. However, the term on the left side is decreasing in α , since by taking the logarithm and applying the inequality $\ln(x) \leq x - 1$, we obtain that

$$\frac{d}{d\alpha} \ln \left(\frac{\alpha-1}{\alpha}\right)^{\alpha-1} = \ln \left(\frac{\alpha-1}{\alpha}\right) + \frac{1}{\alpha} \leq \frac{\alpha-1}{\alpha} - 1 + \frac{1}{\alpha} = 0$$

Hence, the maximum of the left side is attained for $\alpha = 2$, yielding a value of $\left(\frac{2-1}{2}\right)^{2-1} = \frac{1}{2}$, which concludes the proof. \square

PROOF OF THEOREM 5.2. The case $\alpha = 1$ is trivial and $\alpha = 2$ has been treated in the main part of paper. For $\alpha \geq 3$, $\gamma \in (0, 1)$ and $T \geq \frac{1}{2}$, the function $f_\alpha(x) := x^{\alpha-1} - (1-x)^{\alpha-1} - \gamma - T \ln \frac{x}{1-x}$ is continuous and satisfies

$$\begin{aligned} \lim_{x \rightarrow 0^+} f(x) &= \lim_{x \rightarrow 0^+} \left(x^{\alpha-1} - (1-x)^{\alpha-1} - \gamma - T \ln \frac{x}{1-x} \right) = \infty \\ f\left(\frac{1}{2}\right) &= -\gamma < 0 \end{aligned}$$

Hence, since f is continuous in $(0, 1)$, by Bolzano's theorem that there exists $x^* \in (0, \frac{1}{2})$ such that $f(x^*) = 0$. To prove uniqueness, it will be sufficient to prove that for $T \geq 1/2$, $f_\alpha(x)$ is decreasing in x . This implies that for $T \geq 1/2$, there is a unique steady state x^* and

hence, the critical temperature T_c will necessarily satisfy $T_c \leq 1/2$. Taking the derivative of $f_\alpha(x)$ with respect to x , we obtain

$$\frac{d}{dx}f_\alpha(x) = (\alpha - 1)x^{\alpha-2} + (\alpha - 1)(1 - x)^{\alpha-2} - T\frac{1}{x(1 - x)}.$$

The last expression is decreasing in T and hence it suffices to prove that $f_\alpha(x) \leq 0$, for $T = 1/2$. In this case, $f_\alpha(x) \leq 0$ is equivalent to

$$x^{\alpha-1}(1 - x) + x(1 - x)^{\alpha-1} \leq \frac{1}{2(\alpha - 1)}$$

and the claim follows from Lemma B.5. In particular, equality holds only if $T = \frac{1}{2}$, $\alpha = 3$ ($\alpha - 1 = 2$), and $x = \frac{1}{2}$. Hence, f is decreasing in $(0, 1)$ which proves the claim. Finally, for $T = 0$ and any $\alpha > 1$ (the case $\alpha = 1$ is trivial), the first derivative of f with respect to x becomes

$$\frac{d}{dx}f_\alpha(x) = x(1 - x) [x^{\alpha-1} - \gamma - (1 - x)^{\alpha-1} = (\alpha - 1)x^{\alpha-2} > 0,]$$

for all $x \in (0, 1)$ and all $\alpha > 1$ which implies that $f_\alpha(x)$ is monotonically increasing. Since f starts negative and $f(1/2) = -\gamma < 0$, the third state lies in $(1/2, 1)$. Hence, the dynamics

$$\dot{x} = x(1 - x) [x^{\alpha-1} - \gamma - (1 - x)^{\alpha-1}]$$

has two obvious steady states $x = 0$ and $x = 1$ and one third steady state in $(1/2, 1)$. Accordingly, the usual stability analysis applies which proves that $[0, 1/2]$ lies in the attracting region of $x = 0$ as claimed. \square

Remark 4. In the continuous interval, $\alpha \in [2, 3]$, $f_\alpha(x)$ is not monotone decreasing. However the statement of Theorem 5.2 continues to hold. Since this case requires more technical details without providing any additional insight, its proof is omitted.