Defensive medicine, liability insurance and malpractice litigation in an evolutionary model

Defensive medicine, liability insurance and malpractice litigation in an evolutionary model

Angelo Antoci[a], Alessandro Fiori Maccioni[a], Marcello Galeotti[b], Paolo Russu[a,*]

[a] *Department of Economics, University of Sassari*
[b] *Department of Statistics, Informatics and Applications, University of Florence*

Abstract

We analyse the relations between defensive medicine, medical malpractice insurance and litigious behaviours, by an evolutionary game between physicians and patients. When medical treatment fails, patients may suit the physician and seek compensation. Conversely, physicians may prevent negligence charges by practising defensive medicine or by buying medical malpractice insurance. The latter choice transfers the risk of litigation from the physician to the insurer. By studying the population dynamics, we intend to describe how clinical and legal risks can shape the interactions between healthcare providers and patients, and how this can affect the diffusion and the price of medical malpractice insurance.

Keywords: defensive medicine; malpractice; insurance; clinical risk; evolutionary game.
JEL classification: C62, C73, I13

1. Introduction

Defensive medicine and liability insurance are both self-protective tools aimed at protecting physicians from malpractice claims (Ehrlich and Becker, 1972). Defensive medicine is a deviation from sound medical practice motivated by the threat of liability (Kessler and McClellan, 1996). It can expose patients to risks of harm from unnecessary or inappropriate procedures (Tancredi and Barondess, 1978). Health issues include, among others, the excessive use of Caesarean section to deliver babies (Dubay et al., 1999, 2001; Feess, 2012) and the excessive exposure to radiation in diagnosis (Hendee et al., 2010).

Defensive medicine in high-risk specialities is a worldwide issue. In the United States, 93% of respondent physicians reported practising it (Studdert et al., 2005), while comparable numbers emerged in Europe (Garcia-Retamero and Galesic, 2014; Osti and Steyrer, 2016; Palagiano, 2013; Ramella et al., 2015), China (He, 2014), and Japan (Hiyama et al., 2006). Throughout their career, U.S. surgeons face a claim almost certainly and pay an indemnity with 70% probability (Jena et al., 2011). Clinical safety may increase the probability of being sued when adverse events occur; for example, after a tenfold decrease in mortality rate (Eichhorn, 1989; Kohn et al., 2000), claims for anesthesia-related death barely declined (Cheney et al., 2006; Peng and Smedstad, 2000).

The magnitude of the problem is staggering. The medical liability system, including defensive medicine, has been estimated to cost the United States more than \$55 billion annually or between 2.4%–10% of total healthcare spending (Kessler and McClellan, 1996; Mello et al., 2010; PriceWaterhouseCoopers, 2006). In Italy, defensive medicine costs to the public healthcare system more than €10 billion per year or 10.5% of its overall expenses (Palagiano, 2013). In Austria, only for radiology, orthopaedic and trauma surgery, it costs to the public system around €420.8 million per year or 1.62% of its overall expenses (Osti and Steyrer, 2016).

*Corresponding author: DiSea University of Sassari, Via Muroni 25, 07100 Sassari (SS) Italy
Email addresses: angelo.antoci@virgilio.it (Angelo Antoci), alex.fiori.maccioni@gmail.com (Alessandro Fiori Maccioni), marcello.galeotti@dmd.unifi.it (Marcello Galeotti), russu@uniss.it (Paolo Russu)

Medical malpractice insurance is a type of professional liability insurance that protects healthcare providers against claims of medical negligence. Its market has occasionally experienced times of crisis (such as in the mid-1970s, mid-1980s, and early 2000s, see Baker, 2005), which may be partially attributable to increases in lawsuits. Indeed, malpractice claims increased at nearly 10% per year in the 1970s and 1980s (Danzon, 1991; U.S. General Accounting Office, 1986); whereas they have been moderately stable since then (Jena et al., 2011). The cost of medical malpractice insurance can fluctuate widely (Rodwin et al., 2008), more than other insurance markets over the course of the underwriting cycle (Baker, 2005). Such cost depends on the liability system (Danzon et al., 2004), but the impact of tort reforms is still theoretically ambiguous (Avraham and Schanzenbach, 2010).

Despite the interdependence between defensive medicine and liability insurance, no joint model of analysis exists to our knowledge in literature. We try to fill this gap in the present article. Our starting point is the evolutionary game proposed by Antoci et al. (2016), in which physicians are randomly paired with patients and provide them a risky medical treatment; patients can choose whether to suit or not their physician for medical malpractice, if an adverse event occurs, while physicians can choose whether to practise or not defensive medicine (which may be harmful to patients) to prevent malpractice charges. This framework amounts to a Lotka-Volterra game in which physicians can be seen as prey and litigious patients as their predators.

In the present article, we add to the game the physicians' strategy of insuring against liability claims. The evolutionary dynamics is represented by a four-dimensional system whose variables are the ratios of the adopted strategies, in both populations, and the price of the insurance policy. We show that our approach constitutes an innovative step in modelling the above interaction between physicians and patients, as it adds several new dynamic patterns with respect to the original game. Indeed, depending not only on the price level of the policy but also on the premium calculation principle, the eventual outcome of the game can be an attractive stationary state where some alternative strategy of physicians or patients disappears, or a mixed-strategy Nash equilibrium, or even an oscillatory behaviour in which strategies coexist in a recurrent fashion. Moreover, we find that the role of medical malpractice insurance can be ambiguous as it can discourage or even encourage undesirable behaviours (such as the practice of defensive medicine or the propensity to litigation), depending on the model parameters and, eventually, on the initial conditions of the game.

The article is organized as follows. The next section introduces the model and its assumptions. Sections 3–5 describe the evolutionary dynamics that arises from different premium calculation principles. We discuss our results and draw the conclusions in Section 6. Finally, we present in Appendix all the proofs of theorems in this article.

2. Model

We consider a game between a population of physicians and a population of patients. At each instant of time $t \in [0, +\infty)$, a large number of random pairwise encounters takes place between members of the two populations. In each encounter, a physician provides a risky medical treatment to a patient. An adverse event can affect the treatment with probability $p \in (0, 1)$; if that happens, the patient suffers a damage R and can choose, at a cost C_L, to sue the physician for medical malpractice. The outcome of the litigation is uncertain and depends on the physician's behaviour.

Physicians can choose to practise defensive medicine to minimize possible liability charges. Alternatively, physicians can act to the best of their competence, with no defensive behaviour; if so, they can choose either to retain their legal risk or to transfer it to an insurance company.

The physician who practises defensive medicine incurs a cost C_D, inflicts the patient a harm H, and loses a possible litigation with probability q_D. The physician who chooses not to defend (nor to insure), incurs a cost C_{ND} and loses a possible litigation with probability q_{ND}. The physician who chooses to insure, pays w for the insurance policy, incurs a medical cost C_{ND}, and loses a possible litigation with probability q_I.

We assume $C_D > C_{ND} \geq 0$ and $q_D < q_I \leq q_{ND}$, that is, defensive medicine has a higher immediate cost for physicians, but it protects them by decreasing their probability of being held liable in the event of litigation. Assuming a higher cost of defensive medicine avoids that the defensive strategy dominates the strategy of providing best practice without insurance, which would otherwise disappear. This assumption is consistent with the context in which an eventual additional compensation, received by defensive physicians for performing superfluous tests or procedures, is not enough to compensate them for the time and effort of being unnecessarily overcautious, or for the disadvantages of deviating from best practice (due, for example, to psychological, ethical or reputational reasons). This case is typical of public healthcare, where employed physicians normally receive a fixed salary and unnecessary care is generally discouraged.

If winning the litigation, the patient would get full compensation R either from the physician when not insured, or from the insurance company when the physician is insured. If losing, the patient would pay K either to the winning physician or to the insurance company, as compensation for litigation expenses. To simplify the notation, we define E_D, E_{ND} and E_I as the patient's expected settlement from litigating against respectively a defensive, a not-defensive and an insured physician:

$$E_D = Rq_D - K(1 - q_D)$$

$$E_{ND} = Rq_{ND} - K(1 - q_{ND})$$

$$E_I = Rq_I - K(1 - q_I)$$

From the preceding definitions, it results $E_{ND} \geq E_I > E_D$.

2.1. One-shot game

The one-shot game works as follows. The physician can play three pure strategies: practising defensive medicine (strategy D), neither defending nor insuring (strategy ND), and buying a malpractice insurance (strategy I). The patient can play two pure strategies, L or NL, representing, respectively, the choice of litigating or not in case that an adverse event occurs. Each player chooses the strategy without knowing *ex ante* the other player's choice.

The physician's payoffs of strategies D, ND and I are:

$$\begin{array}{c|cc} & L & NL \\ \hline D & -C_D - pE_D & -C_D \\ ND & -C_{ND} - pE_{ND} & -C_{ND} \\ I & -C_{ND} - w & -C_{ND} - w \end{array} \tag{1}$$

The patient's payoffs of strategies L and NL are:

$$\begin{array}{c|ccc} & D & ND & I \\ \hline L & -H - p(C_L - E_D) & -p(C_L - E_{ND}) & -p(C_L - E_I) \\ NL & -H & 0 & 0 \end{array} \tag{2}$$

We omitted in matrix (2) the term $-pR$, representing the patient's expected damage from the adverse event, because such term, appearing in every entry, does not affect the relative performance of strategies.

The terms in the payoff matrices satisfy the conditions:

$$p \in (0,1), \quad E_{ND} \geq E_I > E_D, \quad H > 0, \quad C_D > C_{ND} \geq 0, \quad C_L > 0, \quad w \geq 0.$$

2.2. Evolutionary dynamics

We define the evolutionary dynamics as follows. At each instant of continuous time t, a large number of physicians and patients are randomly paired and play the one-shot game described above.

Let $x(t)$, $y(t)$ and $1-x(t)-y(t)$ represent the shares of the population of physicians who adopt, respectively, the strategies D, ND and I, with $x, y \geq 0$ and $x+y \leq 1$. Let also $z(t)$ and $1-z(t)$ represent the shares of the population of patients who adopt, respectively, the strategies L and NL, with $1 \geq z \geq 0$.

From matrix (1), the physicians' expected payoffs from playing strategies D, ND and I result:

$$\begin{aligned}\Pi_D &= (-C_D - pE_D)\, z - C_D\,(1-z) \\ &= -C_D - pzE_D\end{aligned}$$

$$\begin{aligned}\Pi_{ND} &= (-C_{ND} - pE_{ND})\, z - C_{ND}\,(1-z) \\ &= -C_{ND} - pzE_{ND}\end{aligned}$$

$$\begin{aligned}\Pi_I &= (-C_{ND} - w)\, z + (-C_{ND} - w)\,(1-z) \\ &= -C_{ND} - w\end{aligned}$$

where z and $1-z$ represent the probability that a physician is matched with a patient who plays, respectively, strategy L or NL.

From matrix (2), the patients' expected payoffs from playing strategies L and NL result:

$$\begin{aligned}\Pi_L &= [-H - p(C_L - E_D)]\, x + [-p(C_L - E_{ND})]\, y + [-p(C_L - E_I)]\,(1-x-y) \\ &= -Hx + (E_D - E_I)\, x + (E_{ND} - E_I)\, y + E_I - C_L\end{aligned}$$

$$\begin{aligned}\Pi_{NL} &= -Hx + 0\, y + 0\,(1-x-y) \\ &= -Hx\end{aligned}$$

where x, y and $1-x-y$ represent the probability that a patient is matched with a physician who plays, respectively, strategy D, ND and I.

We assume that individuals are not fully rational and they develop their decisions through a social process of imitation of successful strategies. Accordingly, physicians and patients update their choices by adopting the relatively more rewarding strategy that emerges from available observations of others' choices. At the end of this learning process, if the behaviour of all individuals reaches a stable outcome such that the evolutionary dynamics converges to a stationary state, then that state is a Nash equilibrium (see Hofbauer and Sigmund, 1988; Weibull, 1997), that is, a state in which individuals have no incentive to modify their own strategy if the others do not modify theirs.

Following Taylor and Jonker (1978), we assume that the adoption shares of strategies evolve according to the replicator equations; therefore, the growth or decline in the adoption rate of a strategy will be proportional to the difference between its payoff and the population average payoff:

$$\begin{aligned}\dot{x} &= x(\Pi_D - \Pi_{PH}) \\ &= x\left[(C_D - C_{ND} + pzE_D)(x-1) + pzE_{ND}y + w(1-x-y)\right]\end{aligned} \tag{3}$$

$$\begin{aligned}\dot{y} &= y(\Pi_{ND} - \Pi_{PH}) \\ &= y\left[(C_D - C_{ND} + pzE_D)x + pzE_{ND}(y-1) + w(1-x-y)\right]\end{aligned} \tag{4}$$

$$
\begin{aligned}
\dot{z} &= z(\Pi_L - \Pi_{PI}) \\
&= pz(1-z)\left[E_I - C_L + (E_D - E_I)\,x + (E_{ND} - E_I)\,y\right]
\end{aligned}
\tag{5}
$$

where Π_{PH} and Π_{PI} represent the average payoffs in the populations of physicians and patients, that is, $\Pi_{PH} = x\Pi_D + y\Pi_{ND} + (1-x-y)\Pi_I$ and $\Pi_{PI} = z\Pi_L + (1-z)\Pi_{NL}$.

We now consider the price of the insurance policy in the two alternative scenarios of limited and perfect foresight of agents' behaviours by the insurance company. In the former case, at any time t, the company sets the premium $w(t)$ without knowing *ex ante* the realizations $x(t)$, $y(t)$ and $z(t)$ of the dynamics (3)–(5), describing the adoption shares of strategies in the one-shot game that will be played at time t. The company knows these realizations only *ex post* and uses them to adjust the price according to the following adaptive rule, which determines the evolution over time of the policy premium:

$$
\dot{w} = \sigma\left[\pi(x, y, z) \ - w\right]
\tag{6}
$$

where $\sigma > 0$ is the speed of adjustment of the adaptive rule, w is the current price of the policy, and $\pi\,(x, y, z)$ is the premium that the company would have set if it had known in advance the realizations x, y and z (that is, *ex ante* instead of *ex post*). In other words, if the company had perfect knowledge and foresight of agents' behaviours, it would fix the premium w such that, at any time t:

$$
w = \pi(x, y, z)
\tag{7}
$$

We can also expect that the company (unless subsidized) aims at a policy premium which exceeds the actuarially fair value of the insurance provision, that is $\pi > pzE_I$. The product pzE_I represents the expected loss per policy for compensation awarded to patients.

In the following, we will consider the cases of both limited and perfect foresight by the insurance company, in which the premium will be calculated according to, respectively, condition (6) or (7).

We study the system dynamics in the prism:

$$
\overline{\Delta} = \{x, y \geq 0, x + y \leq 1, 0 \leq z \leq 1, 0 \leq w \leq \max \pi\}
$$

The open prism will be denoted by Δ. We will consider in Δ the change of variables:

$$
\begin{aligned}
u &= \ln \tfrac{x}{y} \\
v &= \ln\,(x + y)
\end{aligned}
\tag{8}
$$

which yields the system:

$$
\begin{aligned}
\dot{u} &= C_{ND} - C_D + pz\,(E_{ND} - E_D) \\
\dot{v} &= (1 - e^v)\left\{\tfrac{e^u}{e^u+1}\left[C_{ND} - C_D + pz\,(E_{ND} - E_D)\right] - pzE_{ND} + w\right\} \\
\dot{z} &= pz(1-z)\left[E_I - C_L + (E_D - E_I)\,\tfrac{e^{u+v}}{e^u+1} + (E_{ND} - E_I)\,\tfrac{e^v}{e^u+1}\right] \\
\dot{w} &= \sigma\left[\widetilde{\pi}\,(u, v, z) - w\right]
\end{aligned}
\tag{9}
$$

where $\widetilde{\pi}$ denotes the expression of π after the change of variables.

2.3. Additional conditions

We focus on the most interesting case of the defensive medicine game proposed by Antoci et al. (2016), in which no dominant strategy exists (without insurance) and the frequency of each behaviour fluctuates cyclically around its value at the unique stable stationary state. Accordingly, we operate under the following assumption.

Assumption 1. *Strategy D does not dominate strategy ND and vice versa, that is:*

$$C_D - C_{ND} < p\,(E_{ND} - E_D) \tag{10}$$

Furthermore, strategy L does not dominate strategy NL and vice versa, that is:

$$E_D < C_L < E_{ND} \tag{11}$$

This assumption will be maintained throughout the article. It avoids that the dynamics of the game becomes trivial (because of a unique globally attractive equilibrium) when no insurance is available. Antoci et al. (2016) have shown that, under such assumption, when no insurance is available, the unique stable stationary state $(\bar{x},\, \bar{z})$ has coordinates:

$$\bar{x} = \frac{E_{ND} - C_L}{E_{ND} - E_D} \tag{12}$$

$$\bar{z} = \frac{C_D - C_{ND}}{p\,(E_{ND} - E_D)} \tag{13}$$

and the shares of adoption of strategies D and L fluctuate cyclically around it; therefore, given that $\bar{x}, \bar{z} \in (0, 1)$, all available strategies (except insuring) coexist in that state. These coordinates also represent the mixed strategies of physicians and patients at the Nash equilibrium of the one-shot game, when no insurance is available. So, if they were perfectly rational, the physician would practice defensive medicine with probability $\bar{x}$ and the patient would choose to litigate with probability $\bar{z}$. It has been also shown that the expected payoffs of both populations are lower in the state $(\bar{x}, \bar{z})$, when existing, than in the state of perfect cooperation $(x, z) = (0, 0)$, in which nobody defends nor litigates. Note that the stationary state $(\bar{x},\, \bar{z})$ of the two-dimensional dynamics analysed by Antoci et al. (2016) corresponds to the stationary state $(\bar{x},\, \bar{y}{=}1{-}\bar{x},\, \bar{z},\, w{=}\pi(\bar{x}, \bar{y}, \bar{z}))$ of the system (3)–(6), in which no physician plays strategy I.

We now analyse the cases in which the insurance company has limited foresight and determines the premium according to the adaptive rule (6), by reacting to the behaviour of only patients (in Section 3) or of both patients and physicians (in Section 4). We complete our analysis by addressing (in Section 5) the case in which the insurance company has perfect foresight and adjusts instantaneously the premium according to equilibrium condition (7). To improve readability, all the proofs of theorems in this article are shown in Appendix.

3. The expected value principle

We assume that the insurance company adopts a premium calculation principle based on the expected loss of the policy, plus a proportional or fixed loading charge. Accordingly, in setting the price, the company pays attention to the observed behaviour of patients, without further considering the correlated behaviour of physicians.

We find that the qualitative behaviour of the system (3)–(6) depends on threshold levels related to the policy loading charge (be either proportional or fixed) and to the patients' cost of litigation. Namely, if the loading charge exceeds a higher threshold, the additional cost of insuring is always higher than its additional benefit. Therefore, insurance is too costly, physicians avoid buying it, and the system tends to cyclical dynamics where all strategies except insuring coexist, which are analogous to the two-dimensional dynamics described by Antoci et al. (2016).

If the loading charge falls into an intermediate range, insuring is sufficiently convenient to replace one of its two alternative strategies, which is gradually abandoned by physicians. The replaced strategy depends on patients' cost of litigation: if C_L exceeds (or falls below) the expected compensation E_I from litigating against an insured physician, the system tends to a stationary state where no physician practises defensive medicine (or, respectively, to an oscillatory behaviour in which all strategies except not defending coexist).

Finally, if the loading charge falls below a lower threshold, insurance is so low-priced and convenient that the system tends to a stationary state where all physicians insure and, therefore, defensive physicians tend to disappear.

The above results are exposed in detail in the following sections. The main policy implication is that introducing liability insurance can effectively deter defensive medicine as long as the premium depends upon the actuarially fair value of the policy, plus a sufficiently low loading charge. Defensive medicine may be also deterred if the loading charge is not so low, provided that the patients' cost of litigation is sufficiently high.

3.1. Proportional loading charge

The premium calculation principle based on the actuarially fair value of the policy plus a proportional loading charge is:

$$\pi(z) = pzE_I\,(1+\lambda)\,, \quad \lambda > 0 \tag{14}$$

where pzE_I is the expected loss per policy for compensation awarded to patients, and λ is the loading factor per unit of the premium.

Different dynamic regimes may arise depending on the value of λ with respect to the threshold levels:

$$\lambda_{Hi} = \frac{E_{ND}}{E_I} - 1 \tag{15}$$

$$\lambda_{Lo} = \frac{E_D}{E_I} + \frac{C_D - C_{ND}}{pE_I} - 1 \tag{16}$$

By condition (10), it results $\lambda_{Hi} > \max\{\lambda_{Lo}, 0\}$.

The higher threshold λ_{Hi} corresponds to the value of the loading charge λ that equalizes the expected payoffs Π_I and Π_{ND} in the context $z = 1$, that is, when a physician faces a litigious patient. Similarly, the lower threshold λ_{Lo} represents the value of λ that equalizes the expected payoffs Π_I and Π_D in the context $z = 1$. Such a value, however, can be taken into consideration only if it is positive, implying that E_I is sufficiently low. Indeed, by condition (14), the insurance premium cannot be lower than the expected loss because, as it is well know, this would lead in the long run to insurer's ruin (see, for instance, Bühlmann, 1970).

The following theorem illustrates the basic features of the dynamics in the context in which the loading factor λ does not exceed the higher threshold λ_{Hi} (see Appendix for proofs of theorems). Let us remember that $x(t)$, $y(t)$ and $1 - x(t) - y(t)$ represent the shares of the population of physicians who adopt, respectively, the strategies D, ND and I, and that $z(t)$ and $1 - z(t)$ represent the shares of the population of patients who adopt, respectively, the strategies L and NL.

Theorem 1. *Assume $\lambda < \lambda_{Hi}$. Then we distinguish two cases.*

1. *If $C_L < E_I$, all the trajectories starting in Δ converge to the side $\{y = 0\}$, where no physician chooses the not-defensive strategy. Moreover:*
   (a) *if $\lambda > \max\{\lambda_{Lo}, 0\}$, then most*[1] *trajectories tend to an oscillatory behaviour on the side $\{y = 0\}$, staying away from the other sides of the boundary (see Figure 1);*
   (b) *if $\lambda_{Lo} > 0$ and $0 < \lambda < \lambda_{Lo}$, then all the trajectories starting in Δ converge to the boundary point $\{x = y = 0, z = 1, w = pE_I(1+\lambda)\}$;*
2. *If $C_L > E_I$, then any trajectory in Δ converges to one of the boundary points:*

$$\left\{x = z = w = 0, 0 \le y \le \frac{C_L - E_I}{E_{ND} - E_I}\right\} \tag{17}$$

[1] Here and in the following 'most' means all but a set of measure zero.

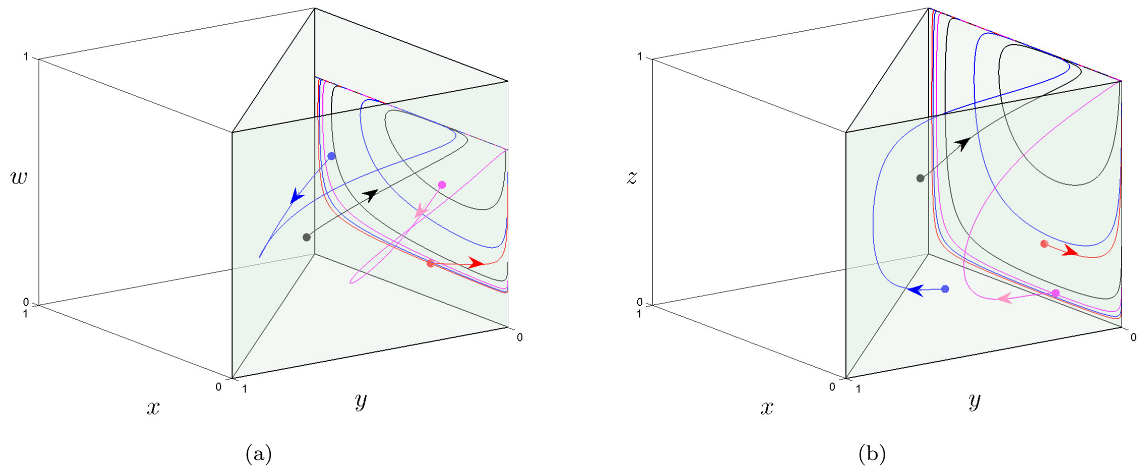


Figure 1: Phase portraits of the dynamical system (3)–(6). Panel (a): variables (x, y, w). Panel (b): variables (x, y, z). All the trajectories starting in Δ converge to the side $\{y{=}0\}$ of $\overline{\Delta}$, where no not-defensive physicians exists (Theorem 1). Parameter values: $E_{ND} = 8$, $E_D = 3$, $E_I = 6$, $C_L = 5$, $p = 0.1$, $C_D - C_{ND} = 0.4$, $\lambda = 0.12$, $\sigma = 0.2$.

According to Theorem 1, if the loading factor λ is lower than the higher threshold λ_{Hi}, and the legal cost C_L of litigating is lower than the expected settlement E_I from litigating against an insured physician, then the share y of not-defensive physicians always tends to zero (starting from every point in the open prism Δ). In such a context, the time evolution of the shares x and $1 - x - y$ of physicians adopting, respectively, strategies D and I depend on the loading factor λ. In particular (recalling that $\lambda < \lambda_{Hi}$ and $C_L < E_I$):

- If $\lambda > \max\{\lambda_{Lo}, 0\}$, then insuring is sufficiently inexpensive to deter the not-defensive strategy. However, defensive medicine is still more convenient than insurance when treating litigious patients, while the opposite holds when physicians face not litigious patients; this strategic scenario generates an oscillatory behaviour of the shares x, $1 - x - y$ and z according to which all strategies except the not-defensive one coexist (see Figure 1).

- If $\lambda_{Lo} > 0$ and $0 < \lambda < \lambda_{Lo}$, then insurance is so low-priced and convenient that the system tends to a state where all physicians insure (that is, the share $1 - x - y$ approaches 1) and all patients are litigious (that is, the share z approaches 1), while the insurance price w tends to its maximum value $pE_I(1 + \lambda)$ (which is observed when $z = 1$).

When the patients' cost of litigation C_L is higher than the expected settlement E_I from litigating against an insured physician (case 2 of Theorem 1), then the share of litigious patients z tends to zero, and so does the insurance premium w. Therefore, the physicians' strategies I and ND tend to have the same expected payoff, while the defensive strategy D tends to disappear. The share of insured physicians exceeds a certain threshold which effectively deters all patients from being litigious. Indeed, by interpreting the shares of adopted strategies as probabilities, fewer and fewer patients resort to litigation because the probability of meeting a non-defensive physician remains low (that is, not higher than $\frac{C_L - E_I}{E_{ND} - E_I}$).

The following theorem deals with the case in which the loading factor λ exceeds the higher threshold λ_{Hi}.

Theorem 2. *Assume, instead, $\lambda > \lambda_{Hi}$. Then, all the trajectories starting in Δ converge to a Hamiltonian dynamics taking place in the side $\{x + y = 1\}$, where no physician buys an insurance, and the system exhibits, on an invariant surface, a centre surrounded by periodic trajectories (see Figure 2).*

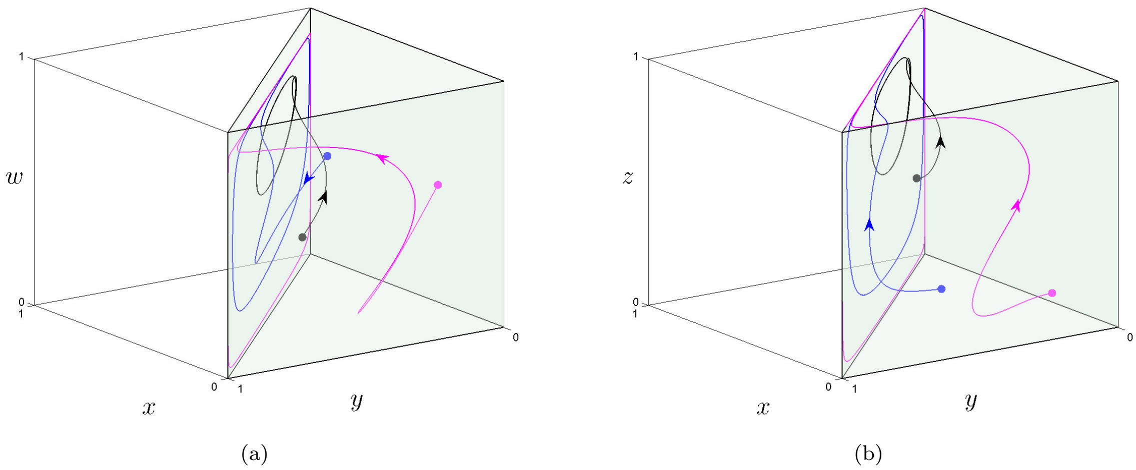


Figure 2: Phase portraits of the dynamical system (3)–(6). Panel (a): variables (x, y, w). Panel (b): variables (x, y, z). All the trajectories starting in Δ converge to the side $\{x + y = 1\}$, where no insured physician exists (Theorem 2). Parameter values: $E_{ND} = 8$, $E_D = 3$, $E_I = 6$, $C_L = 5$, $p = 0.1$, $C_D - C_{ND} = 0.4$, $\lambda = 1$, $\sigma = 0.2$.

It may be useful to compare the results in Theorems 1 and 2 with those by Antoci et al. (2016), already explained in Section 2.3, where the strategy I is not available and no dominant strategy exists among strategies D, ND, L and NL. In that paper, the authors have shown theoretically that a predator–prey relationship can emerge between patients and physicians. Namely, litigious patients who seek compensation are the 'predators' and physicians are their 'prey'. Physicians can adapt to the risk of being sued by performing defensive medicine. Then, a measure that may increase the share of not-defensive physicians, may also lead to an increase in the share of litigious patients, thus triggering this predatory-prey mechanism. Similarly, a measure favouring litigious behaviour (such as a reduction in the cost of litigation C_L) may also lead to an increase in the share of defensive physicians, which can be considered as 'adapted' prey who have improved their fitness through mutation.

Theorems 1 and 2 illustrate what happens if, in the context analysed by Antoci et al. (2016), the new strategy of insuring becomes available to physicians. Given the assumption $E_{ND} \geq E_I > E_D$, the patient's expected settlement from litigating against a defensive physician is lower than the expected settlement from litigating against a not-defensive and an insured physician; therefore, these latter physicians play the role of prey in the interactions with litigious patients.

According to Theorem 2, if the loading factor λ exceeds the higher threshold λ_{Hi}, then insurance is too costly and the system tends to predator–prey cyclical dynamics in which all strategies except insuring coexist (see Figure 2), similarly to the two-dimensional cyclic dynamics obtained by Antoci et al. (2016).

According to Theorem 1, for low-enough values of the loading factor and of the legal cost, such that $\lambda < \lambda_{Hi}$ and $C_L < E_I$, the share y of not-defensive physicians always tends to zero. In such a context, we have two scenarios. If $\lambda_{Lo} > 0$ and $0 < \lambda < \lambda_{Lo}$, then the insurance price w is close enough to the actuarially fair value of the policy, pzE_I, while E_I is in turn (possibly much) lower than E_{ND}, and all the trajectories starting in the open prism Δ approach the stationary state $\{x = y = 0, z = 1, w = pE_I(1 + \lambda)\}$, where all physicians insure and all patients are litigious. If, on the contrary, $\max\{\lambda_{Lo}, 0\} < \lambda < \lambda_{Hi}$, then oscillations of the predator–prey type can be observed. In such a case, oscillations occur because of the strategic context faced by physicians and patients, namely:

- strategy D performs better than strategy I when played against a litigious patient, while the opposite holds when a physician faces a not-litigious patient;

- strategy L performs better than strategy NL when played against a physician adopting strategy I, while the opposite holds when a patient faces a physician adopting strategy D.

Consequently, when the share of litigious patients z is high enough, strategy D performs better than I, and thus the share x of physicians practising defensive medicine increases. The increase in x, in turn, pushes down the share z, until it becomes so low that strategy I starts to perform better than D; therefore, the share $1-x$ of insured physicians increases, pushing z high again, and so on.

Finally, when the patients' cost of litigation C_L is higher than the expected settlement E_I, it is not convenient to litigate with insured physicians, and the system tends to one of the stationary states (17) where no patient is litigious and (recalling that $\lambda < \lambda_{Hi}$) no physician practises defensive medicine (case 2 of Theorem 1). Then, even though the physicians' strategies I and ND tend to have the same expected payoff, the share of insured physicians always exceeds a certain threshold which prevents patients from litigating.

3.2. Fixed loading charge

We now consider the case in which the insurance company applies a fixed loading charge (instead of the proportional one assumed above). The premium, calculated as the expected loss of the policy plus a fixed premium loading, results:

$$\pi(z) = pzE_I + k, \quad k > 0 \tag{18}$$

The dynamics of the system, when the loading charge is fixed, can be described by the same arguments used in Theorems 1 and 2 for the case of proportional loading. Namely, there exist three ranges of values for the loading charge, which determine, according the patients' cost of litigation, the qualitative behaviour of the system.

In the higher range, insurance is so costly that physicians avoid buying it, and the system tends to cyclical dynamics where all strategies except insuring coexist. In the intermediate range, insuring is sufficiently convenient to replace one of its two alternative strategies, which is gradually abandoned by physicians. The replaced strategy depends on the patients' cost of litigation. Finally, in the lower range, insurance is so low-priced and convenient that the system tends to a stationary state where all physicians insure and, therefore, nobody practises defensive medicine.

To simplify the subsequent steps, we introduce the parameter:

$$l := k - \frac{E_{ND} - E_I}{E_{ND} - E_D}(C_D - C_{ND}) \tag{19}$$

Now, we can state the following proposition (which can be proved by the same arguments of Theorem 1 and 2, see Appendix).

Theorem 3. *If $l < 0$ and:*

1. *$C_L < E_I$, then all the trajectories starting in the open prism Δ tend to the side $\{y = 0\}$. Precisely, if $k < C_D - C_{ND} - pE_I$, the trajectories tend to the sink $\{x = y = 0, z = 1, w = pE_I + k\}$. If, instead, $k > C_D - C_{ND} - pE_I$, then most trajectories tend to an oscillatory behaviour on the side $\{y = 0\}$.*
2. *$C_L > E_I$, then all the trajectories starting in the open prism Δ tend to the side $\{x = 0\}$, where the system has a sink in:*

$$\left(x = 0,\, y = \frac{C_L - E_I}{E_{ND} - E_I},\, z = \frac{k}{p(E_{ND} - E_I)},\, w = \frac{kE_{ND}}{E_{ND} - E_I}\right) \tag{20}$$

Vice versa, if $l > 0$, then all the trajectories starting in Δ converge to a Hamiltonian dynamics taking place in the side $\{x + y = 1\}$, where no physician buys an insurance, and the system exhibits, on an invariant surface, a centre surrounded by periodic trajectories (that is, the dynamics is qualitatively the same as in Theorem 2, see Figure 2).

The parameter l can be seen as a measure of the advantage or disadvantage of buying an insurance, for the physician facing a litigious patient. Indeed, if $l < 0$, insuring is sufficiently low-priced and convenient to replace at least one of its alternative strategies; then, the defensive or not-defensive strategy may disappear depending on whether or not the patients' cost of litigation C_L exceeds the expected compensation E_I from litigating against an insured physician. Conversely, if $l > 0$, the insurance strategy becomes too costly and tends to disappear. Note that l is positive only when k is sufficiently high, given that the other term in (19) is always non-positive.

We intend to compare the results of case 2 in Theorems 1 and 3. In both cases, the physicians' defensive strategy D tends to disappear, as all the trajectories starting in Δ tend to the side $\{x = 0\}$ and reach a stationary state where both strategies ND and I coexist.

What differs is that, in Theorem 3, the share z of litigious patients never tends to zero, while it always does so in case 2 of Theorem 1. In the latter, when the share z tends to zero, the price $\pi(z) = pzE_I\,(1+\lambda)$ tends to zero as well because of the proportional loading λ and, therefore, the strategies ND and I tend to have the same payoff. This allows the convergence of trajectories to one of the stationary states (17), where $z = 0$ and both strategies ND and I coexist.

Conversely, in Theorem 3, because of the fixed premium loading $k > 0$, the price $\pi(z) = pzE_I + k$ would never tend to zero even if the share z of litigious patients tended to zero as well. Indeed, when the share z is low enough, the not-defensive strategy ND becomes more remunerative than strategy I. Then, a positive share z is necessary to guarantee that, at the stationary state (20), the payoffs of both strategies of physicians become the same.

4. Market-dependent premium

We still assume that the premium calculation principle consists of the expected loss plus a positive loading charge. We now consider such loading to be somehow related to physicians' behaviour[2]. Accordingly, in setting the price, the company pays attention to the observed behaviours of both physicians and patients. The simplest formulation of such mechanism is:

$$\pi\,(x, y, z) = p\,z\,E_I\;+ k - ax - by \tag{21}$$

with k, $k - a$, $k - b > 0$; these inequalities guarantee a positive premium loading. The parameter k represents the fixed part of the premium loading, while a and b represent its variable part related to, respectively, the share of defensive and not-defensive physicians.

We make no *a priori* assumption on the signs of a and b. We could expect them to be negative in presence of economies of scale (the more policies sold, the lower their price); conversely, they might be positive because of diseconomies of scale or of marketing strategy, so that the company reduces the premium when fewer policies are sold and increase it otherwise. Notably, if $a = b$ it results $k - ax - by = k - a(x + y)$, that is, the loading charge only depends on the share of not-insured physicians.

First of all, we observe that, if the insurance premium is given by formulae (18) or (21), implying a positive loading for any (x, y) such that $0 \le x + y \le 1$, then, irrespectively of the patients' cost of litigation, litigious behaviour never disappears (while it is easily checked that, if all patients tend to litigate, then all physicians tend to self-protect, through defensive medicine or liability insurance). More precisely, the following proposition holds:

[2]Clearly, insurance companies take count of potential policyholders' behaviours when setting policy premiums. However, the literature devoted to this subject is, as far as we know, rather scarce. An interesting very recent contribution to the topic, although in a field different from ours (life policies) and with a different mathematical approach, is given by Baione et al. (2018).

Proposition 4. *Assume that the premium is calculated according to (18) or (21). Then, no trajectory starting in Δ can approach the side $\{z = 0\}$. Moreover:*

1. *if $C_L < E_I$, no trajectory starting in Δ can approach the side $\{x = 1\}$, while trajectories approaching $\{x = 0\}$ approach the side $\{y = 0\}$ as well;*
2. *if, instead, $C_L > E_I$, then the previous point holds exchanging x with y.*

The proof, which we skip, is given by arguments similar to those used in demonstrating Theorems 1 and 2. What is important to observe is that, when $z(t)$ is very close to zero and thus $u(t)$ keeps decreasing, so that $x(t)$ becomes, in its turn, very small, then, since the insurance premium is bounded away from zero, $y(t)$ increases, such that, eventually, z increases again. Such a behaviour can be better understood by interpreting the shares in term of probabilities. Consider, in particular, the case $C_L > E_I$. Then litigiousness can decrease, since the expected insurance reimbursement is less than patients' litigation cost. However, as the insurance premium does not tend to zero, precisely that fact will encourage more and more physicians to adopt the *ND* strategy. Hence, when the probability, for a patient, of meeting a *ND* physician becomes sufficiently high, the litigious behaviour will increase again.

From now on we assume that the premium is given by formula (21). Let us investigate, first of all, the conditions for the existence of an interior equilibrium point. Recalling the definition (19), if the linear system:

$$\begin{cases} E_I - C_L + (E_D - E_I)\,x + (E_{ND} - E_I)\,y = 0 \\ l = ax + by \end{cases} \tag{22}$$

has one solution $(x^*, y^*) \in (0,1)^2$ satisfying $x^* + y^* < 1$, then there exists exactly one stationary point $P^* \in \Delta$, which is also the mixed-strategy Nash equilibrium.

The first equation in the above system means, by equation (5), that patients have the same expected payoff from litigious and not-litigious strategies. As regards the interpretation of the second equation, consider, for example, $l > 0$. Then, if the premium were given by formula (18), there would be a *disadvantage* for the physician in buying an insurance with respect to the other strategies; hence, the second equation of (22) says that, by taking into account physicians' behaviours in formulating the premium, such *disadvantage* can be completely offset (while the opposite conclusion holds in case of $l < 0$).

If instead, the system (22) has no solution in Δ, we are, substantially, re-conducted to the cases of the previous section. Namely, the premium parameters determine a threshold level, above which the premium is too costly and nobody tends to insure. Below the threshold, on the contrary, defensive (or not-defensive) physicians tend to disappear, depending on whether (or not) the patients' cost of litigation exceeds the expected compensation from insured physicians. These results can be summarized as follows

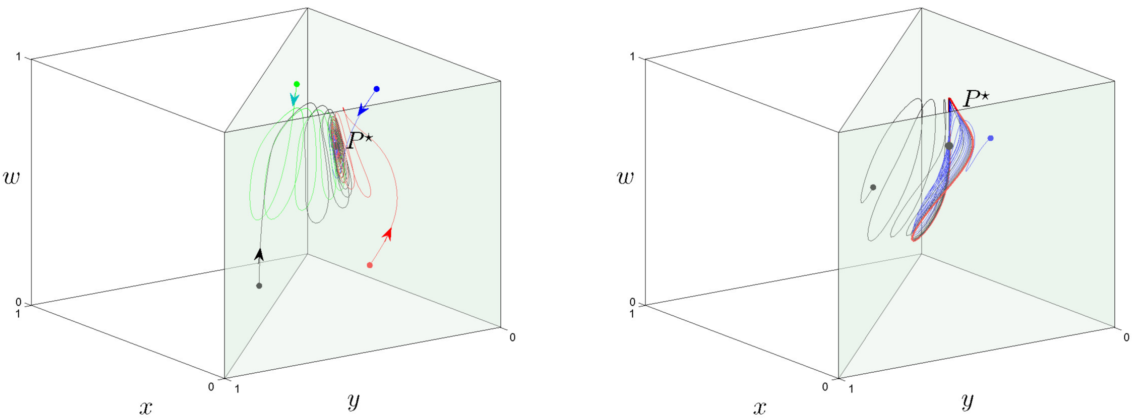
Proposition 5. *Assume that no interior stationary point exists. Then:*

1. *if $l\,(E_{ND} - E_D) - (a - b)\,(E_I - C_L) < 0$ and:*
   (a) *$C_L < E_I$, then all the trajectories in Δ tend to the side $\{y = 0\}$;*
   (b) *$C_L > E_I$, then all the trajectories in Δ tend to the side $\{x = 0\}$;*
2. *if $l\,(E_{ND} - E_D) - (a - b)\,(E_I - C_L) > 0$, then all the trajectories in Δ tend to the side $\{x + y = 1\}$.*

By Cramer's rule, it is possible to solve the linear system (22) and obtain the conditions for having an interior stationary point, which represents the mixed-strategy Nash equilibrium of the game. The following proposition, which can be obtained by algebraic manipulation of (22), states such conditions.

Proposition 6. *The system has exactly one interior stationary point P^* if and only if the following inequalities hold:*

$$\begin{gathered} 0 < \frac{l(E_{ND} - E_I) + b(E_I - C_L)}{a(E_{ND} - E_I) + b\,(E_I - E_D)}, \\ \frac{l(E_I - E_D) - a(E_I - C_L)}{a(E_{ND} - E_I) + b\,(E_I - E_D)} < \frac{l(E_{ND} - E_D) - (a - b)(E_I - C_L)}{a(E_{ND} - E_I) + b\,(E_I - E_D)} < 1 \end{gathered} \tag{23}$$



(a) $\sigma = 0.3 > \sigma_H$, trajectories converge to P^*

(b) $\sigma = 0.128 < \sigma_H$, trajectories approach the cycle

Figure 3: Phase portraits of the dynamical system (3)–(6) with variables (x, y, w). Panel (a): σ above the Hopf bifurcation value. Panel (b): σ below the Hopf bifurcation value. Case with $\det J(P^*) > 0$ and $l > 0$: no trajectories starting in Δ can approach the boundary of $\overline{\Delta}$ (Theorem 7). $P^* = (\overline{x}, \overline{y}, \overline{z}, \overline{w}) = (0.5, 0.25, 0.8, 0.64)$. Bifurcation value: $\sigma_H = 0.21664$. Parameter values: $E_{ND} = 8$, $E_D = 3$, $E_I = 6$, $C_L = 5$, $p = 0.1$, $C_D - C_{ND} = 0.4$, $a = 0.2$, $b = 0$, $k = 0.26$.

By straightforward calculations and adopting the coordinates (u, v, z, w), we obtain the Jacobian matrix:

$$J(P^*) = \begin{pmatrix} 0 & 0 & \alpha & 0 \\ 0 & 0 & \beta & \gamma \\ \delta & \zeta & 0 & 0 \\ \sigma\varphi & \sigma\vartheta & \sigma\eta & -\sigma \end{pmatrix}$$

where:

$\alpha = p(E_{ND} - E_D)$; $\quad \beta = p\left(e^{v^*} - 1\right)\left[E_{ND} - \frac{x^*}{x^*+y^*}(E_{ND} - E_D)\right]$; $\quad \gamma = 1 - e^{v^*}$;

$\delta = -pz^*(1 - z^*)(E_{ND} - E_D)\frac{e^{u^*+v^*}}{\left(e^{u^*}+1\right)^2}\left[\frac{e^{u^*+v^*}}{\left(e^{u^*}+1\right)^2} = \frac{x^* y^*}{x^*+y^*}\right]$; $\quad \zeta = -pz^*(1 - z^*)(E_I - C_L)$;

$\varphi = -(a - b)\frac{e^{u^*+v^*}}{\left(e^{u^*}+1\right)^2}$; $\quad \vartheta = -l$; $\quad \eta = pE_I$.

Recall that, at the interior stationary point $P^* = (x^*, y^* z^*, w^*)$, it results $x^*, y^* z^*, w^* \in (0, 1)$. Then, by simple calculation, it follows:

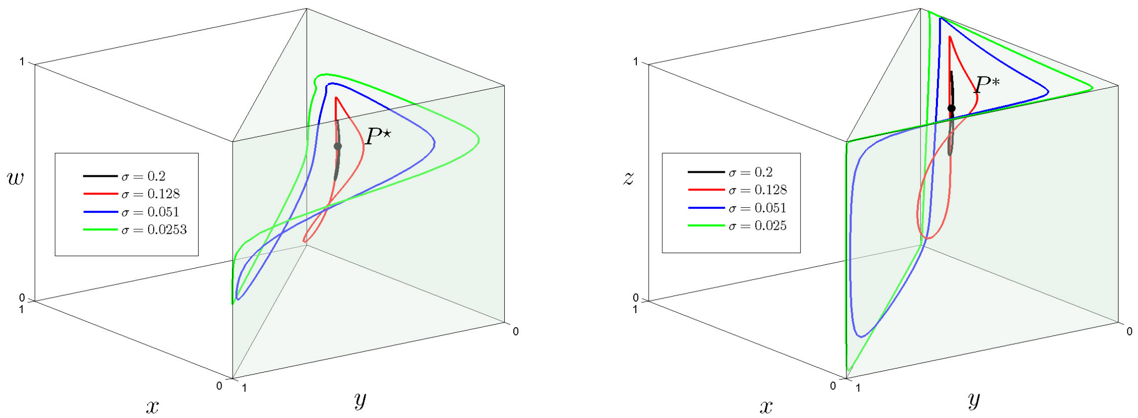
$$sgn\,(\det J(P^*)) = sgn\,(l(E_{ND} - E_D) - (a - b)(E_I - C_L)) \tag{24}$$

Remark 1. *It can be checked that* $\det J(P^*) > 0$ *implies that at least one, between* a *and* b*, is positive, while* $\det J(P^*) < 0$ *implies that at least one, between* a *and* b*, is negative.*

In the following theorem (proved in Appendix), we analyse the coexistence of all available strategies of physicians and patients. If the assumptions are satisfied, then ω-limit sets do not intersect the boundary of Δ, meaning that, along any trajectory in the open prism Δ, all strategies tend to coexist.

Theorem 7. *Suppose the stationary point* P^* *exists, with* $\det J(P^*) > 0$ *and* $l > 0$*. Then, no trajectory starting in* Δ *can approach the boundary of* $\overline{\Delta}$ *(see Figure 3).*

The assumption $l > 0$ in the above theorem is satisfied only for a sufficiently high value of the parameter k (the fixed premium loading, see also Section 3.2). Similarly, the assumption $\det J(P^*) > 0$ (which implies P^* to be neither a sink nor a source nor a saddle with two-dimensional stable manifold) is satisfied only for



(a) (b)

Figure 4: Phase portraits of the dynamical system (3)–(6). Panel (a): variables (x, y, w). Panel (b): variables (x, y, z). Family of cycles arising from a Hopf bifurcation, with $\sigma < \sigma_H$. Parameter values: same of Figure 3.

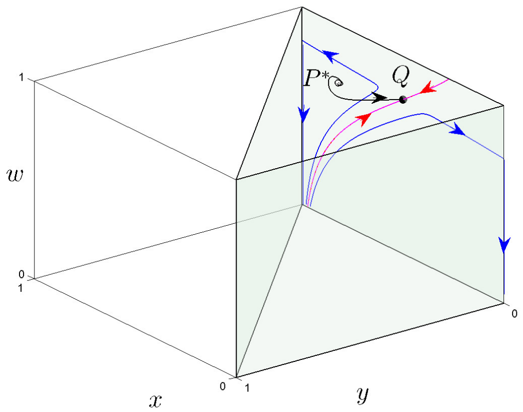
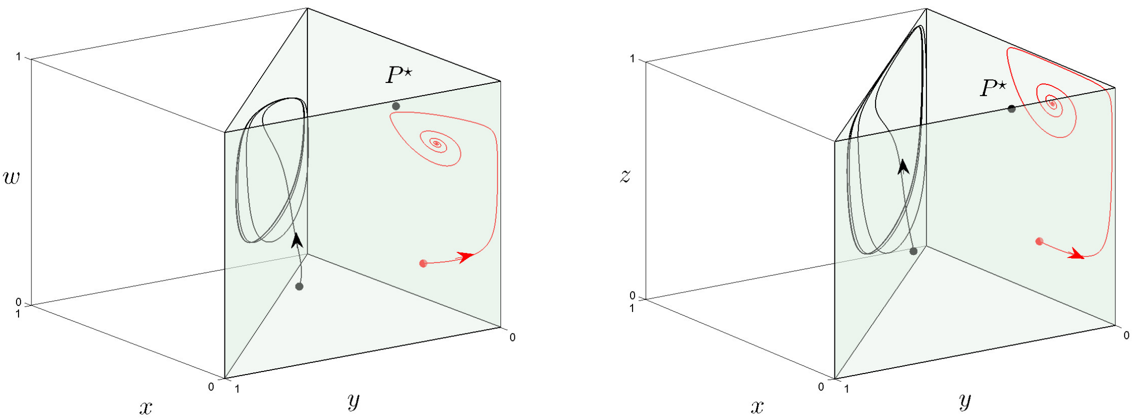


Figure 5: Phase portrait of the dynamical system (3)–(6), with variables (x, y, w). Case with $\det J(P^*) > 0$ and $l < 0$ (Remark 2). $P^* = \left(\frac{11}{15}, \frac{1}{15}, \frac{29}{30}, \frac{58}{75}\right)$ (2-dimensional stable manifold), $Q = (0.5, 0, 1, 0.78)$ (3-dimensional stable manifold). Parameter values: $E_{ND} = 8$, $E_D = 2$, $E_I = 5$, $C_L = 3$, $p = 0.1$, $C_D - C_{ND} = 0.58$, $a = -0.1$, $b = 0.2$, $k = 0.23$, $\sigma = 0.2$.

positive values of either a or b (the proportional components of the premium loading negatively related to the shares of, respectively, defensive and not-defensive physicians). Indeed, a positive value of a or b may reduce the insurance premium when fewer policies are sold, thus counterbalancing the effects of a high k and giving rise, for example, to cyclical behaviour of the system.

What is most important, the above theorem affirms that, when the model parameters satisfy its conditions, then the action of the insurance company, aimed at maximizing profits, can have the *collateral* effect of preventing each strategy, of patients and physicians, from being dominated and eventually disappearing. Let us observe, finally, that the theorem's conditions are satisfied for quite reasonable values of the parameters: for example, this is the case if $0 < l < a = b$ and the difference $a - l$ is sufficiently small.

In Figure 4 we illustrate a family of cycles arising from a Hopf bifurcation. Each cycle corresponds to a different value of the bifurcation parameter σ, representing the speed of adjustment of the insurance premium to its equilibrium value, defined in equation (6). The bifurcation value, for the given parameter set, is $\sigma_H = 0.21664$. Then, the interior stationary point P^* is a sink for $\sigma > \sigma_H$ while it is a saddle with a two-dimensional stable manifold for $\sigma < \sigma_H$ (hence a Hopf bifurcation can be conjectured).



(a) (b)

Figure 6: Multi-dimensional system whose variables are the shares of strategies adopted by physicians and, respectively, the price of insurance and the share of litigious patients. Case with $\det J\left(P^{*}\right)<0$ and $l>0$: trajectories starting in Δ with $x+y$ sufficiently large tend to the side $\{x+y=1\}$, where no physician insures (Theorem 8). $P^{*}=(\overline{x},\overline{y},\overline{z},\overline{w})=(0.4,0.1,0.8,0.64)$. Parameter values: $E_{ND}=8$, $E_D=3$, $E_I=6$, $C_L=5$, $p=0.1$, $C_D-C_{ND}=0.4$, $a=0.2$, $b=-0.2$, $k=0.22$, $\sigma=0.2$.

Remark 2. *The situation is different if we assume, instead, $l<0$. In such a case, there may exist an invariant hypersurface whose trajectories tend to the side $\{y=0\}$ (see Figure 5).*

The following numerical example illustrates such a possibility. Let:

$$E_{ND}=8,\ E_D=2,\ E_I=5,\ C_L=3,\ p=0.1,\ C_D-C_{ND}=0.58,\ a=-0.1,\ b=0.2,\ k=0.23.$$

Then the interior stationary point is $P^{*}=\left(\frac{11}{15},\frac{1}{15},\frac{29}{30},\frac{58}{75}\right)$, which can be seen to be a saddle with two-dimensional stable manifold (hence $\det J(P^{*})>0$). Moreover, it is easily computed that the boundary stationary point $Q=(0.5,0,1,0.78)$ is a saddle with a three-dimensional stable manifold having two-dimensional intersection with $\{y=0\}$. The simulations referred to such a case are shown in Figure 5.

We now consider what happens to the system if $\det J\left(P^{*}\right)<0$. If so, P^{*} is a saddle endowed with a three- or one-dimensional stable manifold. In such a case, we prove that there exist trajectories starting in Δ tending to the boundary of $\overline{\Delta}$, where one of the physicians' strategies disappears; that is, to the sides $\{x+y=1\}$ or $\{y=0\}$ or $\{x=0\}$. The patients' cost of litigation may discourage the defensive (or not-defensive) behaviour by physicians, if exceeding (or not) the expected compensation from insured physicians. We summarize these results in the following theorem (see Appendix for the proof).

Theorem 8. *Assume the stationary point P^{*} exists and $\det J\left(P^{*}\right)<0$. Then:*

1. *if $l>0$, trajectories starting in Δ with $x+y$ sufficiently large tend to the side $\{x+y=1\}$ (see Figure 6);*
2. *if $C_L<E_I$ and $a>0$, trajectories starting in Δ with a sufficiently small y tend to the side $\{y=0\}$; while in the case $C_L<E_I$ and $a<0$, implying $l<0$, such trajectories may also converge to the side $\{x+y=1\}$;*
3. *if $C_L>E_I$ and $b>0$, trajectories starting in Δ with a sufficiently small x tend to the side $\{x=0\}$; while in the case $C_L>E_I$ and $b<0$, implying $l<0$, such trajectories may also converge to the side $\{x+y=1\}$.*

The results of the above theorem lend themselves to interesting comments. First of all, the assumption is that the system, with the premium given by formula (21), has a stationary point which is neither a sink nor a source nor a saddle with two-dimensional stable manifold ($\det J\left(P^{*}\right)<0$). Then there exist trajectories converging to the prism boundary, that is, along which one of the physicians' strategies tends to disappear. The theorem classifies the various cases. So, if $l > 0$, when the initial number of insured physicians is already low, the insurance strategy tends to disappear, in agreement with previous analogous observations. Figure 6 illustrates such a case.

On the other hand, independently of the sign of l, a low initial number of *ND* physicians leads to the disappearance of the *ND* strategy if $C_L < E_I$ and $a > 0$, meaning that the insurance premium decreases if defensive physicians increase (there is a competition between insurance and defensive strategy). An analogous result, exchanging y with x, holds if $C_L > E_I$ and $b > 0$. However, still under the theorem's assumption, if $C_L < E_I$ and $a < 0$ (or if $C_L > E_I$ and $b < 0$), an initial situation characterized by a *low* y (or, respectively, a *low* x) may lead, instead, to the disappearance of the insurance strategy. In fact, in such a case the insurance premium increases when the initially prevailing strategy (respectively, defensive and not-defensive) increases, which, in turn, may reinforce precisely that strategy and make the share of insured physicians become smaller and smaller.

5. Instantaneous adjustment

Let us imagine that the insurance company has perfect foresight of agents' behaviours and is able to adjust instantaneously the policy premium to its equilibrium value so that it results $w(t) = \pi\left(x\left(t\right), y\left(t\right), z\left(t\right)\right)$, at any time t. We find that some of the results obtained in the previous sections do not change.

In such a case, the dynamic system (9) becomes three-dimensional, namely:

$$\begin{aligned}
\dot{u} &= C_{ND} - C_D + pz\left(E_{ND} - E_D\right) \\
\dot{v} &= (1 - e^v)\left\{\tfrac{e^u}{e^u+1}\left[C_{ND} - C_D + pz\left(E_{ND} - E_D\right)\right] - pzE_{ND} + \widetilde{\pi}\left(u, v, z\right)\right\} \\
\dot{z} &= pz(1 - z)\left[E_I - C_L + \left(E_D - E_I\right)\tfrac{e^{u+v}}{e^u+1} + \left(E_{ND} - E_I\right)\tfrac{e^v}{e^u+1}\right]
\end{aligned} \tag{25}$$

where $\widetilde{\pi}$ denotes the expression of π after the change of variables (8).

Then if the expected value principle (14) is assumed (as in Section 3) the statements of Theorems 1 and 2 hold *verbatim*. Indeed, in the proof of Theorem 1 in Appendix, the function $\mathcal{L}\left(u, v, w\right) = \ln\frac{e^v}{1-e^v} + \frac{w}{\sigma} + \ln\frac{e^{qu}}{1+e^u}$ can be simply replaced by $\widetilde{\mathcal{L}}\left(u, v\right) = \ln\frac{e^v}{1-e^v} + \ln\frac{e^{qu}}{1+e^u}$; moreover, it can be checked that, under the assumptions of point 1a of Theorem 1, the system (25) is equivalent to a Hamiltonian one with centre in $(\overline{x}, \overline{z}) = \left(\frac{E_I - C_L}{E_I - E_D}, \frac{C_D - C_{ND}}{p[E_I(1+\lambda) - E_D]}\right)$.

If, instead, the premium calculation principle is market-dependent according to (21) (as in Section 4), then an interior stationary point P^* exists for the original system (9) if and only if an interior stationary point S^* exists for system (25). Moreover, it is easily checked that:

$$sign(\det J\left(S^*\right)) = -sign(\det J\left(P^*\right)) = sgn((a - b)\left(E_I - C_L\right) - l\left(E_{ND} - E_D\right))$$

Hence the statements of Theorems 7 and 8 can be rephrased by changing the sign of the Jacobian determinant, while the demonstrations follow the steps of the original ones by just cancelling the expression $\frac{w}{\sigma}$ in the functions that we considered for the proofs of Theorems 7 and 8 in Appendix. However, as mentioned in the captions of Figures 3 and 4, a low value of σ can give place to Hopf bifurcations and, thus, to the birth of limit cycles. Vice versa, instantaneous adjustment is equivalent to let $\sigma = +\infty$. Hence, in the latter case we can expect a lower tendency to cyclicity or recurrence of the model.

6. Conclusions

We analysed self-protective behaviours of physicians and their relation with medical malpractice litigation by patients. We built upon the evolutionary game proposed by Antoci et al. (2016), where the interaction between physicians (choosing whether to practise defensive medicine or not) and patients (choosing whether to be litigious or not) leads to cyclic behaviours typical of predator-prey models. In the present paper we introduced, as an additional strategy, the possibility for the physicians to insure against medical malpractice liability. We assumed different methods of calculation of the policy premium, and studied how they can influence the long-term behaviour of physicians and patients. The evolutionary dynamics was described by a four-dimensional system whose variables are the ratios of the adopted strategies, in both populations, and the price of the insurance policy. We found that the availability of liability insurance can effectively deter defensive medicine. This result is possible when the price of insurance is calculated according to its actuarially fair value, plus a loading charge (either proportional or fixed).

We also considered another possible premium calculation principle, which takes into account the behaviour of potential policy-holders. Accordingly, the premium loading charge consists of a fixed part plus a proportional part related to the shares of defensive and not-defensive physicians. In such a case, the system can exhibit an interior stationary state where all the physicians' and patients' strategies coexist. The most interesting result is that, when this state exists, its Jacobian determinant is positive, and the fixed part of the loading charge is high enough, then, if the initial conditions of the system are such that all of the strategies co-exist, none of these strategies will tend to disappear over time. The meaning is that, when the price of insurance is market-dependent, we may have the permanence of all the possible behaviours by physicians and patients. On the contrary, when the Jacobian determinant of the interior stationary state is negative, there exist trajectories along which one strategy of physicians tends to disappear, and the behaviour of the system becomes path-dependent.

In conclusion, we cannot affirm that in general the introduction of a medical malpractice insurance can completely discourage the practice of defensive medicine. However, the availability of such a further choice for physicians introduces many new scenarios, which, in turn, can suggest a broad range of strategies that a policy-maker may undertake to pursue the goal of a fair and efficient public health care. The design of such strategies can well be the goal of further research on this topic.

Acknowledgment

The authors would like to express their sincere appreciation to the anonymous reviewers for their significant suggestions from which the paper has been substantially improved.

References

Angelo Antoci, Alessandro Fiori Maccioni, and Paolo Russu. The ecology of defensive medicine and malpractice litigation. *PloS One*, 11(3):e0150523, 2016.

Ronen Avraham and Max Schanzenbach. The impact of tort reform on private health insurance coverage. *American Law and Economics Review*, 12(2):319–355, 2010.

Fabio Baione, Davide Biancalana, Paolo De Angelis, and Ivan Granito. Dynamic policyholder behaviour and surrender option evaluation for life insurance. In *Mathematical Methods for Actuarial Sciences and Finance (in press)*. Springer, 2018.

Tom Baker. Medical malpractice and the insurance underwriting cycle. *DePaul Law Review*, 54, 2005.

Hans Bühlmann. *Mathematical methods in risk theory.* Springer-Verlag, Berlin, 1970.

Frederick W. Cheney, Karen L. Posner, Lorri A. Lee, Robert A. Caplan, and Karen B. Domino. Trends in anesthesia-related death and brain damagea closed claims analysis. *The Journal of the American Society of Anesthesiologists*, 105(6):1081–1086, 2006.

Patricia M. Danzon. Liability for medical malpractice. *The Journal of Economic Perspectives*, 5(3):51–69, 1991.

Patricia M. Danzon, Andrew J. Epstein, and Scott J. Johnson. The 'crisis' in medical malpractice insurance. *Brookings-Wharton Papers on Financial Services*, 2004(1):55–96, 2004.

Lisa Dubay, Robert Kaestner, and Timothy Waidmann. The impact of malpractice fears on cesarean section rates. *Journal of Health Economics*, 18(4):491–522, 1999.

Lisa Dubay, Robert Kaestner, and Timothy Waidmann. Medical malpractice liability and its effect on prenatal care utilization and infant health. *Journal of Health Economics*, 20(4):591–611, 2001.

Isaac Ehrlich and Gary S. Becker. Market insurance, self-insurance, and self-protection. *Journal of political Economy*, 80(4):623–648, 1972.

John H. Eichhorn. Prevention of intraoperative anesthesia accidents and related severe injury through safety monitoring. *Anesthesiology*, 70(4):572–577, 1989.

Eberhard Feess. Malpractice liability, technology choice and negative defensive medicine. *The European Journal of Health Economics*, 13(2):157–167, 2012.

Rocio Garcia-Retamero and Mirta Galesic. On defensive decision making: how doctors make decisions for their patients. *Health Expectations*, 17(5):664–669, 2014.

Alex Jingwei He. The doctor–patient relationship, defensive medicine and overprescription in chinese public hospitals: Evidence from a cross-sectional survey in shenzhen city. *Social Science & Medicine*, 123:64–71, 2014.

William R. Hendee, Gary J. Becker, James P. Borgstede, Jennifer Bosma, William J. Casarella, Beth A. Erickson, C. Douglas Maynard, James H. Thrall, and Paul E. Wallner. Addressing overutilization in medical imaging 1. *Radiology*, 257(1):240–245, 2010.

Toru Hiyama, Masaharu Yoshihara, Shinji Tanaka, Yuji Urabe, Yoshihiko Ikegami, Tatsuma Fukuhara, and Kazuaki Chayama. Defensive medicine practices among gastroenterologists in japan. *World J Gastroenterol*, 12(47):7671–7675, 2006.

Josef Hofbauer and Karl Sigmund. *The Theory of Evolution and Dynamical Systems: Mathematical Aspects of Selection*. Cambridge University Press Cambridge, 1988.

Anupam B. Jena, Seth Seabury, Darius Lakdawalla, and Amitabh Chandra. Malpractice risk according to physician specialty. *New England Journal of Medicine*, 365(7):629–636, 2011.

Daniel Kessler and Mark McClellan. Do doctors practice defensive medicine? *The Quarterly Journal of Economics*, 111(2):353–390, 1996.

Linda T. Kohn, Janet M. Corrigan, Molla S. Donaldson, et al. *To err is human: building a safer health system*, volume 6. National Academies Press, 2000.

Michelle M. Mello, Amitabh Chandra, Atul A. Gawande, and David M. Studdert. National costs of the medical liability system. *Health Affairs*, 29(9):1569–1577, 2010.

Michael Osti and Johannes Steyrer. A perspective on the health care expenditures for defensive medicine. *The European Journal of Health Economics*, pages 1–6, 2016. doi:10.1007/s10198-016-0848-4.

Antonio Palagiano. Relazione conclusiva della commissione parlamentare di inchiesta sugli errori in campo sanitario e sulle cause dei disavanzi sanitari regionali. Technical report, Camera dei Deputati, Roma, 2013.

Philip W.H. Peng and Kari G. Smedstad. Litigation in canada against anesthesiologists practicing regional anesthesia. a review of closed claims. *Canadian journal of anaesthesia*, 47(2):105–112, 2000.

PriceWaterhouseCoopers. *The Factors Fueling Rising Healthcare Costs 2006*. PriceWaterhouseCoopers, 2006.

Sara Ramella, Giovanni Mandoliti, Lucio Trodella, and Rolando Maria D'Angelillo. The first survey on defensive medicine in radiation oncology. *La radiologia medica*, 120(5):421–429, 2015.

Marc A. Rodwin, Hak J. Chang, Melissa M. Ozaeta, and Richard J. Omar. Malpractice premiums in massachusetts, a high-risk state: 1975 to 2005. *Health Affairs*, 27(3):835–844, 2008.

David M. Studdert, Michelle M. Mello, William M. Sage, Catherine M. DesRoches, Jordon Peugh, Kinga Zapert, and Troyen A. Brennan. Defensive medicine among high-risk specialist physicians in a volatile malpractice environment. *JAMA: the Journal of the American Medical Association*, 293(21):2609–2617, 2005.

Laurence R. Tancredi and Jeremiah A. Barondess. The problem of defensive medicine. *Science*, 200(4344): 879–882, 1978.

Peter D. Taylor and Leo B. Jonker. Evolutionary stable strategies and game dynamics. *Mathematical biosciences*, 40(1):145–156, 1978.

U.S. General Accounting Office. *Medical Malpractice: Six State Case Studies Show Claims and Insurance Costs Still Rise Despite Reforms*, volume GAO/HRD-87-21. Washington, DC, 1986.

Jörgen W. Weibull. *Evolutionary Game Theory*. MIT press, 1997.

Appendix

Proof of Theorem 1. Consider system (9) and recall Assumption 1. We construct a *Lyapunov-like* function, namely:

$$\mathcal{L}(u,v,w) = \ln \frac{e^v}{1-e^v} + \frac{w}{\sigma} + \ln \frac{e^{qu}}{1+e^u}$$

where $q = \frac{E_{ND} - E_I(1+\lambda)}{E_{ND} - E_D}$, $0 < q < 1$. Then, it is easily verified that:

$$\frac{d}{dt}\mathcal{L}(u,v,w) = q\left(C_{ND} - C_D\right) < 0$$

which, by straightforward computations, implies:

$$\ln x^q y^{1-q} \to -\infty$$

that is, $xy \to 0$.

Now, assume $E_I > C_L$. If a trajectory converged to $\{x = 0\}$, then, recalling the expression of $\dot{z}$, it easily follows that $z(t) \to 1$, which, in turn, implies (see 9) $u(t) \to +\infty$, hence $y(t) \to 0$.

Next, suppose $C_D - C_{ND} < p\left[E_I(1+\lambda) - E_D\right]$.

In the invariant side $\{y = 0\}$ the dynamic system becomes:

$$\begin{aligned}
\dot{x} &= x(1-x)\left[C_{ND} - C_D - pzE_D + w\right] \\
\dot{z} &= pz(1-z)\left[E_I - C_L + (E_D - E_I)x\right] \\
\dot{w} &= \sigma\left[pzE_I(1+\lambda) - w\right]
\end{aligned} \tag{26}$$

Assume a trajectory $\gamma(t) = (x(t), y(t), z(t), w(t))$ *approaches*[3] the side $\{x = 0\}$. As a consequence, fixed any small $\varepsilon > 0$, there exists a sequence of intervals $I_n = [t_n, t_{n+1}]$ such that $x(t) < \varepsilon$ as $t \in I_n$. It follows that, for ε sufficiently small, $z(t)$ is increasing in I_n. Moreover, we can choose these intervals in such a way that in each one $x(t)$ is decreasing and more and more close to 0. Hence, being, correspondingly, $\dot{x}$ smaller and smaller, the length $t_{n+1} - t_n$ tends to $+\infty$. But this implies $z(t)$ becoming, in each succeeding I_n, closer and closer to 1 and therefore $x(t)$ increasing, which leads to a contradiction. The proof that a trajectory cannot tend to the other sides either is similar.

Moreover, system (26) exhibits exactly one stationary point. By calculating the characteristic polynomial of the Jacobian matrix, this point turns out to be a saddle with one-dimensional stable manifold. Hence any trajectory is *oscillating* (possibly, a limit cycle surrounding that saddle).

Assume, instead, $C_D - C_{ND} > p\left[E_I(1+\lambda) - E_D\right]$ and consider again system (26). Then, for a suitable α it results $\frac{d}{dt}\left(\frac{x}{1-x} + \frac{w}{\sigma}\right) < \alpha < 0$, which easily implies claim 1b.

Suppose, now, $E_I < C_L$. It follows, by an argument analogous to the one above, that, along any trajectory starting in Δ, $x(t) \to 0$. Suppose $u(t) \to -\infty$. Then it is easily computed that $\frac{d}{dt}\left(-\ln(1+e^u) + \frac{w}{\sigma} + \ln \frac{e^v}{1-e^v}\right) = -p(E_{ND} - E_I^*)z$. Then, if $z(t)$ does not tend to zero, it follows $e^v \to 0$, i.e. $y(t) \to 0$ as well, which, in turn, implies $z(t) \to 0$, leading to a contradiction. Therefore the theorem's conclusions follow by straightforward steps. ■

Proof of Theorem 2. Considering system (9) and recalling Assumption 1, from the same arguments of the previous theorem, it follows $\frac{d}{dt}\mathcal{L}(u,v,w) > 0$, which is easily seen to imply that, along any trajectory lying

[3] We say that a trajectory $\gamma(t)$ starting in Δ *approaches* a side Σ of $\overline{\Delta}$ if there exists a sequence of times $t_n \to +\infty$, such that $\lim_{t_n \to +\infty} \operatorname{dist}(\gamma(t_n), \Sigma) = 0$.

inside the prism, it results $x(t)+y(t)\to 1$ as $t\to+\infty$. Considering on such invariant side the variables u, z, w, the relative dynamic system decouples into the system:

$$\begin{aligned}\dot{u} &= C_{ND}-C_D+pz\left(E_{ND}-E_D\right)\\ \dot{z} &= pz(1-z)\left[E_I-C_L+\left(E_D-E_I\right)\tfrac{e^u}{e^u+1}+\left(E_{ND}-E_I\right)\tfrac{1}{e^u+1}\right]\end{aligned}$$

which is easily checked to be equivalent to a Hamiltonian system, having a centre in $\left(\ln\frac{E_{ND}-C_L}{C_L-E_D}, \frac{C_D-C_{ND}}{p(E_{ND}-E_D)}\right)$ and the differential equation $\dot{w}=\sigma\left[pzE_I\left(1+\lambda\right)-w\right]$, so that, by equations (6) and (14), the theorem easily follows. ■

Proof of Theorem 7. Suppose, first, that a trajectory $\gamma(t)=(x(t), y(t), z(t), w(t))$ *approaches* the side $\{y=0\}$. According to our definition, this means that there exists a sequence of times $t_n\to+\infty$ such that, along γ, $y(t_n)\to 0$. If, correspondingly, $\limsup u(t_n)<+\infty$, then $x(t_n)\to 0$ as well, i.e. $v(t_n)\to-\infty$. Since $|\dot{v}(t)|$ is bounded, it is easily seen that we can find a sequence of intervals $I_n=(t_n, t_{n+1})$ with $t_{n+1}-t_n\to\infty$, along which $v(t_{n+1})-v(t_n)<-M$, for any arbitrarily large M. On the other hand in I_n, posed $q=\frac{E_{ND}-E_I}{E_{ND}-E_D}$, it results:

$$\frac{d}{dt}\left(\ln\frac{e^v}{1-e^v}+\frac{w}{\sigma}+\ln\frac{e^{qu}}{1+e^u}\right)=l+\varepsilon>0$$

as $|\varepsilon|$ can be chosen as small as we want if n is large enough. Hence we are led to a contradiction, and the same occurs if we suppose $\lim u(t_n)=+\infty$ in a sequence of I_n satisfying the above properties. Indeed, if, in that case, v would tend to 0 in larger and larger I_n, then x would tend to 1 implying $z\to 0$, so that u would tend to $-\infty$, leading to a contradiction.

Assume, then, that a trajectory $\gamma(t)=(x(t), y(t), z(t), w(t))$ *approaches* the side $\{x=0\}$. In order not to get back to the previous case, we can suppose there exists a sequence of intervals $I_n=(t_n, t_{n+1})$ with $t_{n+1}-t_n\to+\infty$ and $u(t_{n+1})-u(t_n)<-M$ for any arbitrarily large M. Then, for a large n, it holds in I_n:

$$\frac{d}{dt}\left(\ln\frac{e^v}{1-e^v}+\frac{w}{\sigma}+\ln\frac{e^{qu}}{1+e^u}+\frac{b}{p\left(E_{ND}-E_I\right)}\ln\frac{z}{1-z}\right)=l+\frac{b(E_I-C_L)}{E_{ND}-E_I}+\varepsilon$$

which is positive when ε is small enough (i.e., n is large enough). Hence we are led again to a contradiction. In fact, if $z\to 1$ then $u\to+\infty$, while, if $z\to 0$ and $b<0$, then $\frac{d}{dt}(\ln\frac{e^v}{1-e^v}+\frac{w}{\sigma})=k-by+\varepsilon>0$ for a large n, implying v and therefore y tending to 1, so that z would tend to 1 as well.

Suppose, now, that a trajectory approaches $\{x+y=1\}$. Then, there exists a sequence of times $t_n\to+\infty$ such that, along γ, $\ln\frac{e^v}{1-e^v}(t_n)\to+\infty$, and the previous arguments can be repeated, replacing u by $\ln\frac{e^v}{1-e^v}$. It follows that, for n sufficiently high and I_n sufficiently large, we can compute:

$$\tfrac{d}{dt}\left(\ln\tfrac{e^v}{1-e^v}+\tfrac{w}{\sigma}+\ln\tfrac{e^{qu}}{1+e^u}-\tfrac{a-b}{p(E_{ND}-E_D)}\ln\tfrac{z}{1-z}\right)=l-\tfrac{(a-b)(E_I-C_L)}{E_{ND}-E_D}-b-\tfrac{(a-b)(E_{ND}-E_I)}{E_{ND}-E_D}+\varepsilon$$

so that, for a sufficiently high n, the sign of the above derivative, in I_n, is that of:

$$l\left(E_{ND}-E_D\right)-(a-b)\left(E_I-C_L\right)-\left[b(E_I-E_D)+a(E_{ND}-E_I)\right]$$

On the other hand, by solving (22), we get, by our assumptions:

$$\xi^*=x^*+y^*=\frac{l\left(E_{ND}-E_D\right)-(a-b)\left(E_I-C_L\right)}{b\left(E_I-E_D\right)+a(E_{ND}-E_I)}<1$$

where the numerator is positive. Hence the sign of the above derivative is negative in I_n, when n is high enough. In order not to be led to a contradiction, we have to assume that, in I_n, either $u\to-\infty$ or $z\to 0$ $(a-b>0)$ or $z\to 1$ $(a-b<0)$. But $u\to-\infty$ implies $x\to 0$ and $z\to 1$, so that $u\to+\infty$, while $z\to 0$ implies $u\to-\infty$

and $x \to 0$, so that $z \to 1$; finally, $z \to 1$ implies $y \to 0$ and $x \to 1$, so that $z \to 0$. Hence we are led to a contradiction. ■

Proof of Theorem 8. Let us prove, first, point 1. Consider a trajectory $\gamma(t) = (x(t), y(t), z(t), w(t))$ in Δ such that, at the initial time $t = 0$, it results $(x + y)(0) \geq 1 - \varepsilon$, with $\varepsilon > 0$ sufficiently small. Setting $q = \frac{E_{ND}-E_I}{E_{ND}-E_D}$, $0 < q < 1$, let us compute, along γ, the time derivative:

$$\frac{d}{dt}\left(\ln \frac{e^v}{1-e^v} + \frac{w}{\sigma} + \ln \frac{e^{qu}}{1+e^u} - \frac{a-b}{p(E_{ND}-E_D)} \ln \frac{z}{1-z}\right) = l - \frac{(a-b)(E_I - C_L)}{E_{ND}-E_D} + \left[-b - \frac{(a-b)(E_{ND}-E_I)}{E_{ND}-E_D}\right](x+y) \qquad (27)$$

It follows from the theorem's assumptions and the properties of system (22) that $l - \frac{(a-b)(E_I-C_L)}{E_{ND}-E_D} < 0$ and $\left|l - \frac{(a-b)(E_I-C_L)}{E_{ND}-E_D}\right| < -b - \frac{(a-b)(E_{ND}-E_I)}{E_{ND}-E_D}$.

Hence, if ε is small enough, the above derivative is positive at $t = 0$ and, moreover, $\ln \frac{e^v}{1-e^v} > M$, where $M > 0$ is *very large*, if ε is *very small*. If the derivative becomes 0 at a certain $\bar{t} > 0$, then:

$$(x + y)(\bar{t}) = \frac{(a-b)(E_I - C_L) - l(E_{ND} - E_D)}{-b(E_{ND} - E_D) - (a-b)(E_{ND} - E_I)} = r, \quad 0 < r < 1$$

so that $\ln \frac{e^v}{1-e^v}(\bar{t}) << M$. As $\left|\frac{d}{dt} \ln \frac{e^v}{1-e^v}\right|$ is bounded, this implies that the smaller is ε the larger is $\bar{t}$. Moreover, the derived function in (27), say $G(u, v, z, w)$, is larger in $\bar{t}$ than in 0, since $\frac{d}{dt}G(t) > 0$ as $t \in [0, \bar{t})$. Suppose $a - b > 0$. It follows that in G a quantity different from $\ln \frac{e^v}{1-e^v}$ has had a major increase, which can only be $-\ln \frac{z}{1-z}$. Therefore, $z(\bar{t})$ must be sufficiently close to 0. It follows, by the same arguments as above, that if, say, $z(t) \leq \frac{C_D - C_{ND}}{2p(E_{ND}-E_D)}$ when $t \in [\tilde{t}, \bar{t}]$, then $\bar{t} - \tilde{t}$ is also *very large* if ε is *very small*, As, consequently, $\dot{u}(t) \leq -\frac{1}{2}(C_D - C_{ND})$ in $[\tilde{t}, \bar{t}]$, recalling $u = \ln \frac{x}{y}$, that is easily seen to imply that $x(t)$ has been *quite small* too for a *long time* before $\bar{t}$. But eventually this leads to a contradiction, since, when $x(t)$ is sufficiently small, $z(t)$ increases, so that $z(\bar{t})$ cannot be *infinitesimal* with ε. The case $a - b < 0$ can be discussed by analogous arguments, observing that in such a case (27) implies $z(t)$ close to 1, whereas $l > 0$ implies $E_I < C_L$.

Next, consider case 2.

Suppose $y \leq \varepsilon$ at the time $t = 0$ of a trajectory $\gamma(t)$. Then, if ε is small enough:

$$\frac{d}{dt}\left(\ln \frac{e^v}{1-e^v} + \frac{w}{\sigma} + \ln \frac{e^{qu}}{1+e^u} - \frac{a}{p(E_I - E_D)} \ln \frac{z}{1-z}\right) = l - \frac{a(E_I - C_L)}{E_I - E_D} + \left[-b - \frac{a(E_{ND} - E_I)}{E_I - E_D}\right] y \qquad (28)$$

is negative. Moreover, $y(0) \leq \varepsilon$ for a *very small* ε implies that either $u(0) \geq M$ or $v(0) \leq -M$ for a *very large* $M > 0$.

Assume that, at a time $\bar{t} > 0$, $y(\bar{t}) = \frac{a(E_I - C_L) - l(E_I - E_D)}{-b(E_I - E_D) - a(E_{ND} - E_I)}$. Then $u(\bar{t})$, $-v(\bar{t})$ are no more *so large*, while, named again $G(u, v.z, w)$ the derived function in (28), G has been decreasing in $[0, \bar{t}]$. Hence, if $a > 0$, it follows by the above arguments that $z(\bar{t})$ must be very close to 1. However that implies $\dot{u}(t) > r > 0$ for a long time, say, $\bar{t} - \tilde{t}$. Since $u = \ln \frac{x}{y}$, this leads to affirm that $y(\bar{t})$ must be very small and therefore to a contradiction.

If, instead, $a < 0$, the above arguments lead to $z(\bar{t})$ being very close to 0. Hence we can assume that for a sufficiently long interval we have had $z(t) << 1$. Then it may happen $u(\bar{t}) << 0$ and therefore $x(\bar{t})$ very close to 0, which, though, since $C_L < E_I$, implies that $z(t)$ has been increasing in a sufficiently long interval $[\tilde{t}, \bar{t}]$, leading to a contradiction. Otherwise, it is possible that for a long time, say $[t', t'']$, $y(t)$ has been very small compared to $x(t)$, with $z(t)$ being also sufficiently small. In such a case the following derivative

$$\frac{d}{dt}\left(\ln \frac{e^v}{1-e^v} + \frac{w}{\sigma} + \ln \frac{e^{ru}}{1+e^u}\right), \quad r = \frac{k}{C_D - C_{ND}}$$

is checked to have the sign of $-ax(t)$ in $[t', t'']$, i.e. to be positive and bounded away from zero. But this is easily seen to imply that $e^v = x + y$ tends to 1.

The case 3 can be treated analogously, exchanging the role of a with that of b and recalling (23). $\blacksquare$