

Equilibrium Selection in Stochastic Games

P. Jean-Jacques Herings* Ronald J.A.P. Peeters†

23rd October 2001

Abstract

In this paper a selection theory for stochastic games is developed. The theory itself is based on the ideas of Harsanyi and Selten to select equilibria for games in standard form. We introduce several possible definitions for the stochastic tracing procedure, an extension of the linear tracing procedure to the class of stochastic games. We analyze the properties of these alternative definitions. We show that exactly one of the proposed extensions is consistent with the formulation of Harsanyi-Selten for games in standard form and captures stationarity.

JEL Classification Codes: C72, C73

Keywords: Game theory; Stochastic games; Equilibrium selection; Linear tracing procedure; Correlated beliefs.

1 Introduction

Stochastic games were introduced by Shapley (1953). He considered both finite and infinite horizon two-person zero-sum stochastic games with finite state space and finite action spaces. Shapley proved that such games have a value and that both players possess optimal stationary strategies with respect to the discounted payoff criterion. Fink (1964), Takahashi (1964), and Sobel (1971) extended Shapley's model to n -person non-zero-sum stochastic games. For the model with finite state space and finite action spaces they showed the existence of a stationary equilibrium.

A stochastic game is played in stages. At each stage, the game is in one of finitely many states and every player observes the current state. In each stage the players have to make

*Department of Economics, Universiteit Maastricht, P.O. Box 616, 6200 MD Maastricht, The Netherlands, E-mail: P.Herings@algec.unimaas.nl

†Department of Economics, Universiteit Maastricht, P.O. Box 616, 6200 MD Maastricht, The Netherlands, E-mail: R.Peeters@algec.unimaas.nl

a choice (simultaneously and independently) out of the action sets which depend on the state. These choices result not only in a payoff (each state is coupled with a normal form game), but also in an action dependent probability measure on the set of states. Next, according to this probability measure, a chance experiment is carried out to determine the state of the next stage.

For many normal form games there is a vast multiplicity of equilibria (see McLennan (1999)). The situation is not better for the multiplicity of equilibria in stochastic games, even when one restricts attention to stationary equilibria. It can be shown that there is typically a finite number of equilibria (see Haller and Lagunoff (2000)), or even that there is an odd number of equilibria (see Herings and Peeters (2000)). However, the results of McLennan (1999) indicate that one should expect that the number of stationary equilibria is extremely large for an average stochastic game. It is therefore important to develop and analyze selection theories that select a particular stationary equilibrium. The aim of this paper is to extend the linear tracing procedure, which is proposed by Harsanyi and Selten to select equilibria in standard form games, in order to achieve equilibrium selection within the class of stationary equilibria in stochastic games.

In this paper four ways to extend the linear tracing procedure of Harsanyi and Selten (1988) to the setting of stochastic games are presented. The four alternatives originate from two choices of modeling beliefs in stochastic games. First, there is the assumption about beliefs within states. Second, there is the assumption about beliefs across time (stages). Both belief-types might be assumed to be correlated or not. After properly defining the four alternatives, they are analyzed. Only one of the extensions turns out to be consistent with the linear tracing procedure while capturing stationarity. We label this extension of the linear tracing procedure to the class of stochastic games as the stochastic tracing procedure.

The paper is organized as follows. In Section 2 we introduce the notation for stochastic games, we make the restriction to stationary strategies, and we define the concept of stationary equilibrium. Section 3 is an introduction to the linear tracing procedure. In Section 4, the four extensions of the linear tracing procedure to the class of stochastic games are proposed and defined. Section 5 analyzes the four alternatives. The last section, Section 6, summarizes comprehensively.

2 Notations

We study *finite discounted stochastic games*. A finite discounted stochastic game is given by an ordered sextuple

$$\Gamma = \langle N, \Omega, \{S_\omega^i\}_{(i,\omega) \in N \times \Omega}, \{u^i\}_{i \in N}, \pi, \delta \rangle.$$

Here, N is the finite set of players, Ω is the state space containing a finite number of states, and the set S_ω^i is the finite action set of player $i \in N$ in state $\omega \in \Omega$. Further, $u^i : H \rightarrow \mathbb{R}$ is the payoff function of player i and π is the transition map $\pi : H \rightarrow \Delta(\Omega)$, where $H = \{(\omega, s_\omega) \mid \omega \in \Omega, s_\omega \in S_\omega\}$ and $S_\omega = \prod_{i \in N} S_\omega^i$. If in state $\omega \in \Omega$ the players action choices are $s_\omega \in S_\omega$, then player i gets an instantaneous payoff of $u^i(\omega, s_\omega)$ and the probability that the system jumps to state $\bar{\omega}$ is $\pi(\bar{\omega} \mid \omega, s_\omega)$. Finally, $\delta \in [0, 1)$ is a discount factor that is used to discount future payoffs.

Such a stochastic game corresponds to a dynamic system which can be in different states and where at certain stages the players can influence the course of the play. We consider the infinite horizon model and the set of stages is assumed to be identical with the set $\mathbb{N} = \{0, 1, \dots\}$. Players know the game itself and that this knowledge is common knowledge among all the players. Moreover, the initial state ω^0 at stage $k = 0$ is common knowledge to the players.

The game proceeds as follows. All players i select at the initial state, simultaneously and independently of each other, an action $s_{\omega^0}^i \in S_{\omega^0}^i$. Now two things happen, both depending on the current state ω^0 and the action choices s_{ω^0} . (1) Player i earns $u^i(\omega^0, s_{\omega^0})$ for all $i \in N$. (2) The system jumps to the next state ω^1 according to the outcome of a chance experiment; the probability that the next state will be $\bar{\omega}$ equals $\pi(\bar{\omega} \mid \omega^0, s_{\omega^0})$. Subsequently, prior to the next stage $k = 1$, all players are informed about the previous actions chosen by the players, and of the new state ω^1 . At stage $k = 1$, the above procedure repeats itself, etc.

We assume that the game is of perfect recall, i.e. at each stage each player remembers all past actions chosen by all players and all past states that have occurred.

Note that for finite stochastic games, each stage game resembles a normal form game Γ_ω . However, contrary to the situation for normal form games, the game does not exist of a single play, but jumps according to the probability measure $\pi(\cdot \mid \omega, s_\omega)$ to the next state and continues dynamically. In choosing an action in a certain state, a player not only takes into account the immediate payoff, but also his opportunities in the future states.

Like in normal form games, the players are allowed to randomize their pure actions. A mixed strategy of player i in state ω is a probability distribution on S_ω^i . We identify the

set of all probability distributions on S_ω^i with Σ_ω^i . The strategy space of the normal form game in state ω is therefore equal to $\Sigma_\omega = \prod_{i \in N} \Sigma_\omega^i$. Given a mixed strategy combination $\sigma_\omega \in \Sigma_\omega$ and a strategy $\bar{\sigma}_\omega^i \in \Sigma_\omega^i$, we denote by $(\sigma_\omega^{-i}, \bar{\sigma}_\omega^i)$ the mixed strategy that results from replacing σ_ω^i by $\bar{\sigma}_\omega^i$. If a mixed strategy combination $\sigma_\omega \in \Sigma_\omega$ is played, then the instantaneous expected payoff of player i is denoted by $u^i(\omega, \sigma_\omega)$ and the expected transition to state $\bar{\omega}$ by $\pi(\bar{\omega} | \omega, \sigma_\omega)$.

The set of possible histories up to a stage k , $H^k = \prod_{\kappa=0}^{k-1} H$, consists of all sequences $h^k = (\omega^0, s_{\omega^0}, \omega^1, s_{\omega^1}, \dots, \omega^{k-1}, s_{\omega^{k-1}})$ that could have actually occurred up to time k . Here ω^κ represents the state and s_{ω^κ} the actions of the players at stage κ , $\kappa = 0, \dots, k-1$.

A *behavior strategy* σ^{ik} of player i specifies for each stage k , each state ω^k at time k , and each history h^k a strategy $\sigma^{ik}(h^k, \omega^k) \in \Sigma_{\omega^k}^i$. So, a behavior strategy σ^i for player i is a sequence $\sigma^{i0}, \sigma^{i1}, \dots$, where $\sigma^{i0} \in \Sigma^i := \prod_{\omega \in \Omega} \Sigma_\omega^i$ and $\sigma^{ik} : H^k \rightarrow \Sigma^i$ for all $k \geq 1$.

A *stationary strategy* for player i is a behavior strategy for which $\sigma^{ik}(h^k, \omega^k)$ is of the form $\sigma^i(\omega^k)$, i.e. a history independent, but state dependent strategy. In the sequel, a stationary strategy for player i will be denoted by the symbol ρ^i . If player i decides to play the stationary strategy ρ^i , then every time that the system is in state ω , player i selects his pure action according to ρ_ω^i . So, a stationary strategy ρ^i for player i is an element of Σ^i .

Given initial state ω and a strategy σ , the stream of expected payoffs is evaluated by

$$U^i(\omega, \sigma) := \sum_{k=0}^{\infty} \delta^k \cdot U^{ik}(\omega, \sigma),$$

where $U^{ik}(\omega, \sigma)$ denotes the expected utility at stage k . Here, $U^i(\omega, \sigma)$ equals the total discounted expected payoff of player i when the discount factor equals δ , the starting state is ω and the strategy-tuple σ is played. Since the state and action spaces are assumed to be finite, $U^i(\omega, \sigma)$ exists.

A strategy-tuple σ is an *equilibrium* if and only if σ^i is a best response to σ^{-i} for all $i \in N$.

In the sequel of this paper we will restrict ourselves to stationary strategies. Motivations for restricting to stationary strategies are that stationary strategies prescribe the simplest form of behavior that is consistent with rationality, that stationarity captures the notion that ‘bygones are bygones’, and it embodies the principle that ‘minor causes should have minor effects’. Moreover, the focus on stationary strategies allows for clean, unobstructed analysis of the influence of the state variables, stationary strategies substantially reduce the number of parameters to be estimated in dynamic models, and stationary models can be simulated, see also Maskin and Tirole (2001).

Suppose the players decide to play a stationary strategy-tuple ρ and ω is the initial state. The total discounted expected payoff of player i is denoted by $U^i(\omega, \rho)$. The instantaneous payoff player i obtains in stage $k = 0$ equals $u^i(\omega, \rho_\omega)$. The probability that at the next stage the state will be $\bar{\omega}$ equals $\pi(\bar{\omega} | \omega, \rho_\omega)$.

When ρ is a stationary strategy-tuple and ω is the initial state, the expected payoffs are given by the following recursive formula (see e.g. Fink (1964))

$$U^i(\omega, \rho) = u^i(\omega, \rho) + \delta \cdot \sum_{\bar{\omega} \in \Omega} \pi(\bar{\omega} | \omega, \rho) U^i(\bar{\omega}, \rho).$$

A stationary strategy-combination $\rho \in \Sigma$ is a *stationary equilibrium* if it is a Nash equilibrium in stationary strategies. Every finite discounted stochastic game has at least one equilibrium point in stationary strategies (see Fink (1964), Takahashi (1964), and Sobel (1971)). In Haller and Lagonoff (2000) it is shown that the number of stationary equilibria is generically finite. In Herings and Peeters (2000) this result is sharpened to generic oddness. For $\delta = 0$, the number of stationary equilibria is the product of the number of Nash equilibria of the separate stage games. It is shown in McLennan (1999) that the number of Nash equilibria of each stage game might be enormous. So, for $\delta = 0$ the number of stationary equilibria is enormous. There is no reason to expect that the number of stationary equilibria will not be huge when δ is unequal to 0.

3 Linear tracing procedure

In this section, we fix a state ω and explain the linear tracing procedure for the corresponding normal-form game $\Gamma_\omega = \langle N, \{S_\omega^i\}_{i \in N}, \{u^i(\omega)\}_{i \in N} \rangle$.

The linear tracing procedure is a mathematical construct to model a process of convergent expectations, by which rational players come to adopt, and expect each other to adopt, one particular equilibrium point as the outcome for a given game. At the beginning of this outcome-selection process, the players will as yet lack any specific theory predicting the strategies to be used by the other players. Accordingly, each player will express his expectations about the strategy choice of any other player in the form of a subjective probability distribution over the other players' pure strategies. These subjective distributions will be called prior probability distributions, or simply priors. An important assumption of this model is that all players other than i will associate the same prior probability distribution p_ω^i with any given player i . Part of the Harsanyi-Selten theory is the determination of a prior.

The linear tracing procedure is based on a one-parameter family of auxiliary games Γ_ω^t with $t \in [0, 1]$. In any game Γ_ω^t , every player i has the same strategy set Σ_ω^i as he has in

the original game Γ_ω . But his payoff function $v^i(t; \omega)$ in Γ_ω^t is

$$v^i(t; \omega, \rho_\omega) = (1 - t)u^i(\omega, p_\omega^{-i}, \rho_\omega^i) + tu^i(\omega, \rho_\omega),$$

where $u^i(\omega)$ is his payoff in the original game Γ_ω . Clearly, we have $v^i(1; \omega, \rho_\omega) = u^i(\omega, \rho_\omega)$ so that $\Gamma_\omega^1 = \Gamma_\omega$. On the other hand, $v^i(0; \omega, \rho_\omega) = u^i(\omega, p_\omega^{-i}, \rho_\omega^i)$.

Thus Γ_ω^0 is a game of a rather special structure in which the payoff $v^i(0; \omega)$ of each player i will depend only on his own strategy ρ_ω^i and will be independent of the other players' strategy combination ρ_ω^{-i} . Consequently, the game Γ_ω^0 naturally decomposes into several mutually independent and separate maximization problems, one for each player. For almost all games Γ_ω , for almost all choices of the prior vector p_ω , the game Γ_ω^0 has exactly one equilibrium point ρ_ω^0 , which is in pure strategies.

For any auxiliary game Γ_ω^t , the set of all equilibrium points in Γ_ω^t will be denoted $\mathcal{E}(\Gamma_\omega^t)$. By Nash's (1951) existence theorem for equilibrium points, all of these sets will be nonempty. Let $\mathcal{L}_\omega = \mathcal{L}(\Gamma_\omega, p_\omega)$ be the graph of the correspondence $t \rightarrow \mathcal{E}(\Gamma_\omega^t)$ for $t \in [0, 1]$. \mathcal{L}_ω will typically be a collection of pieces of one-dimensional algebraic curves, though in degenerate cases it may also contain isolated points and/or subsets of more than one dimension, see Herings and Peeters (2001).

Suppose the graph \mathcal{L}_ω contains a path γ_ω connecting a point $x_\omega^0 = (0, \rho_\omega^0)$, corresponding to an equilibrium point ρ_ω^0 of the game Γ_ω^0 , with a point $x_\omega^1 = (1, \rho_\omega^*)$, corresponding to an equilibrium point ρ_ω^* of the original game $\Gamma_\omega^1 = \Gamma_\omega$. Then γ_ω will be called a feasible path, whereas x_ω^0 and x_ω^1 will be called the starting point and the end point of this path γ_ω , respectively. Moreover, the strategy part ρ_ω^* of this end point x_ω^1 will be called the outcome selected by path γ_ω . This strategy combination ρ_ω^* can be rationally selected as the outcome of the game because it will always be an equilibrium point of the original game Γ_ω .

We can now define the linear tracing procedure: it consists of selecting an outcome ρ_ω^* for any game Γ_ω by tracing a feasible path γ_ω from its starting point $x_\omega^0 = (0, \rho_\omega^0)$ to its end point $x_\omega^1 = (1, \rho_\omega^*)$. For any given pair $(\Gamma_\omega, p_\omega)$, we will call the linear tracing procedure feasible if the graph $\mathcal{L}_\omega = \mathcal{L}(\Gamma_\omega, p_\omega)$ contains at least one feasible path γ_ω , and we will call it well-defined if \mathcal{L}_ω contains exactly one feasible path γ_ω .

For any possible pair $(\Gamma_\omega, p_\omega)$, the linear tracing procedure is always feasible but is not always well-defined (see Harsanyi (1975)). For any specific vector p_ω of prior probability distributions, almost all games Γ_ω will give rise to a well-defined linear tracing procedure (see Herings and Peeters (2001)).

4 Stochastic tracing procedure

For the remainder of this paper, let a stochastic game Γ and a prior $p \in \Sigma$ be given. The prior p is taken out of the set of stationary strategy-combinations, because we are selecting on stationary equilibria. Moreover, a player always has a stationary best response if his opponents are playing stationary strategies. So, when the prior p is an element from Σ , each player i has a stationary best response to p^{-i} and we can start the linear tracing procedure in a point inducing a stationary equilibrium in Γ^0 .

For every $t \in [0, 1]$, the stochastic tracing procedure generates a stationary equilibrium of the stochastic game $\Gamma^t = \langle N, \Omega, \{\Sigma_\omega^i\}_{\omega \in \Omega, i \in N}, \{V^i(t)\}_{i \in N} \rangle$, where the total expected discounted payoff function $V^i(t)$ of player i is defined such that

$$V^i(0; \omega, \rho) = U^i(\omega, p^{-i}, \rho^i) \quad \text{and} \quad V^i(1; \omega, \rho) = U^i(\omega, \rho).$$

The stochastic game Γ^0 corresponds to a trivial stochastic game, where all players believe that all their opponents play with probability 1 according to the prior belief. The stochastic game Γ^1 coincides with the original stochastic game Γ . Alternative definitions of $V^i(t)$ for $t \in (0, 1)$ will give rise to alternative stochastic tracing procedures.

Given a proper definition of $V^i(t)$ for $t \in (0, 1)$, the stochastic tracing procedure $\mathcal{S}(\Gamma, p)$ is defined as the set of pairs (t, ρ) for which it holds that ρ is a stationary equilibrium of the stochastic game Γ^t , i.e.

$$\mathcal{S}(\Gamma, p) = \left\{ (t, \rho) \in [0, 1] \times \Sigma \mid \rho^i \text{ is a best stationary response to } \rho^{-i} \text{ in } \Gamma^t \right\}.$$

The stochastic tracing procedure is said to be feasible if there exists a path in $\mathcal{S}(\Gamma, p)$ connecting a best response against the prior to a stationary equilibrium of the stochastic game Γ , i.e. there exists a continuous function $\gamma : [0, 1] \rightarrow \mathcal{S}(\Gamma, p)$ such that $\gamma(0) \in \mathcal{S}(\Gamma, p) \cap (\{0\} \times \Sigma)$ and $\gamma(1) \in \mathcal{S}(\Gamma, p) \cap (\{1\} \times \Sigma)$. In general there may be many trajectories $\gamma([0, 1])$ that link a stationary equilibrium of Γ^0 to a stationary equilibrium of Γ^1 . If this trajectory is unique, then $\mathcal{S}(\Gamma, p)$ is said to be well-defined. If the stochastic tracing procedure is well-defined, then it selects a unique stationary equilibrium of the stochastic game Γ .

The concepts stochastic tracing procedure, feasibility and well-definedness are all based on the definition of $V^i(t)$ for $t \in (0, 1)$. There are at least four a priori reasonable ways in which $V^i(t)$ can be defined to extend the linear tracing procedure of Harsanyi and Selten to the setting of stochastic games. Choices have to be made whether a player holds correlated beliefs within a state or not, and whether a player holds correlated beliefs across time or not.

Correlation within states, C(S), means that when a player knows that some opponent plays according to the prior (which he expects with probability $1 - t$), he expects all opponents to play according to the prior. Alternatively, absence of correlation within states, I(S), implies that even when a player knows that some opponent is playing according to the prior he may not infer that other opponents are playing according to the prior; all opponents are expected to play $(1 - t)p^i + t\rho^i$, independent from one another.

Correlation across time, C(T), means that when a player knows that some opponent plays according to the prior today, he expects that opponent to play according to the prior in all future stages. Absence of correlation across time, I(T), implies that even when a player knows that some opponent is playing according to the prior today, this opponent might not play according to the prior in future stages. In all future events he faces independent lotteries which assign probability $1 - t$ to play against the prior strategies of his opponents.

In the rationalizability literature the issue of “correlated beliefs” versus “uncorrelated beliefs” is often discussed. The second requires players to believe that the opponent players choose their strategies independently, while the first does not. Rationalizability with uncorrelated beliefs has been studied extensively defined by Bernheim (1984) and Pearce (1984). The correlated rationalizability concept is defined by Brandenburger and Dekel (1987) and is related to the correlated equilibrium concept of Aumann (1974, 1987).

To analyze the four possible extensions, we first have to define the four alternatives properly by specifying the corresponding functions $V^i(t)$. Therefore, this section contains four subsections, with each subsection devoted to one of the alternatives.

4.1 Alternative 1: C(S), I(T)

Suppose beliefs are correlated within states, but not across time. Because here correlation within states is assumed, each player expects that in stochastic game Γ^t all opponents are playing according to the prior with probability $1 - t$ and playing strategically with probability t . The additional assumption of absence of correlation across time causes that each player faces this lottery at every stage. Therefore, the total discounted payoff to player i in stochastic game Γ^t when the initial state is ω and ρ is played, is

$$V_{C(S),I(T)}^i(t; \omega, \rho) = (1 - t)u^i(\omega, p_\omega^{-i}, \rho_\omega^i) + tu^i(\omega, \rho_\omega) \\ + \delta \cdot \sum_{\bar{\omega} \in \Omega} \left[(1 - t)\pi(\bar{\omega} | \omega, p_\omega^{-i}, \rho_\omega^i) + t\pi(\bar{\omega} | \omega, \rho_\omega) \right] V_{C(S),I(T)}^i(t; \bar{\omega}, \rho).$$

Stationarity implies that this payoff can be written by means of a recurrent relation. This stationarity results from the assumption that the beliefs players have depend only on the

state and not at the stage in which the state is reached.

4.2 Alternative 2: C(S), C(T)

When we assume correlation as well within states as across time, again each player expects that in a stochastic game Γ^t all opponents are playing according to the prior with probability $1 - t$ and playing strategically with probability t . But now, unlike in the previous case, we have correlation across time. This means that when a player knows that some opponent plays according to the prior today, he expects that opponent to play according to the prior in all future stages. Apparently, unlike in the previous case each player faces this lottery only once and this will be at $k = 0$. Therefore, the total discounted payoff to player i in stochastic game Γ^t when the initial state is ω and ρ is played, is

$$V_{C(S),C(T)}^i(t; \omega, \rho) = (1 - t)U^i(\omega, p^{-i}, \rho^i) + tU^i(\omega, \rho).$$

We can also explain the game Γ^t by means of payoffs and transitions. Each player expects with probability $1 - t$ to face the stochastic game without externalities, that is opponents' decisions do not influence the payoffs and transitions, which are determined as if the opponents are playing according to the prior, and with probability t , they expect to face the stochastic game with opponents playing an updated strategy ρ^{-i} .

A solution (t, ρ) for $t \in (0, 1)$ in $\mathcal{S}(\Gamma, p)$ is not necessarily a stationary equilibrium of the game Γ^t . A player i might have a non-stationary strategy that is better than ρ^i against ρ^{-i} in Γ^t . Namely, at $k = 0$ all players make their stationary decisions believing that with probability $1 - t$ all opponents play according to the prior and with probability t they play strategically. In fact, in $k = 0$ all players are facing a lottery. Once the players are in $k = 1$, they know the outcome of the lottery. It might be better, and most of the time it will be better, to revise their strategy having this knowledge, which would lead to a non-stationary strategy.

4.3 Alternative 3: I(S), I(T)

In this subsection we consider the situation in which absence of correlation both within states and across time is assumed. Absence of correlation within states means that in the stochastic game Γ^t a player expects that each opponent is playing independently according to the prior with probability $1 - t$ and playing strategically with probability t . The additional assumption of absence of correlation across time causes every player to face this lottery at every stage. Clearly, these assumptions imply the total discounted payoff to

player i in stochastic game Γ^t when the initial state is ω and the strategy ρ is played to be equal to

$$V_{\text{I(S),I(T)}}^i(t; \omega, \rho) = U^i(\omega, (1-t)p^{-i} + t\rho^{-i}, \rho^i).$$

Just like in Subsection 4.1 it is possible, because of absence of correlation across time, to rewrite the total discounted payoffs in recurrent form:

$$\begin{aligned} V_{\text{I(S),I(T)}}^i(t; \omega, \rho) &= u^i(\omega, (1-t)p_\omega^{-i} + t\rho_\omega^{-i}, \rho_\omega^i) \\ &\quad + \delta \cdot \sum_{\bar{\omega} \in \Omega} \pi(\bar{\omega} | \omega, (1-t)p_\omega^{-i} + t\rho_\omega^{-i}, \rho_\omega^i) V_{\text{I(S),I(T)}}^i(t; \bar{\omega}, \rho). \end{aligned}$$

Since the decision problem a player faces depends on the state only, his problem is stationary no matter at which stage of the game the state is realized. Assuming that beliefs are not correlated across time captures, therefore, the presence of stationarity. Stationarity means that a recurrent relation can be given. A direct result is that best stationary responses are best responses in the broader class of behavior strategies as well.

4.4 Alternative 4: I(S), C(T)

When absence of correlation within states, but correlation across time is assumed, again each player expects that in stochastic game Γ^t each opponent is playing according to the prior with probability $1-t$ and playing strategically with probability t . But this time, in contrast to Subsection 4.3, we have correlation across time. This means that when a player knows that some opponent plays according to the prior today, he expects that opponent to play according to the prior in all future stages. Therefore, the total discounted payoff to player i in stochastic game Γ^t when the initial state is ω and ρ is played, is

$$V_{\text{I(S),C(T)}}^i(t; \omega, \rho) = \sum_{S \subseteq N \setminus \{i\}} (1-t)^S t^{N-S-1} U^i(\omega, p^S, \rho^{N \setminus S}).$$

In the formula, $(1-t)^S t^{N-S-1}$ is the probability that the opponents from $S \subseteq N \setminus \{i\}$ are playing according to the prior, whereas the other opponents are playing strategically. When player i observes that the opponents from S are playing according to p and the others according to ρ , his total discounted payoff is $U^i(\omega, p^S, \rho^{N \setminus S})$. Player i 's payoff in the game Γ^t is therefore the sum over all subsets of $N \setminus \{i\}$ of the probability that only the players from S play according to p (given t) times the total expected discounted payoff in that case.

Like in Subsection 4.2, a solution (t, ρ) for $t \in (0, 1)$ in $\mathcal{S}(\Gamma, p)$ is not necessarily a stationary equilibrium of the game Γ^t . A player i might have a non-stationary strategy that performs better than ρ^i against ρ^{-i} in Γ^t . Apparently, a best stationary response is not necessarily a best response in the broader class of behavior strategies.

4.5 Properties

In order to be useful as a selection theory, it is important that for almost all stochastic games the stochastic tracing procedure determines a unique stationary equilibrium. When well-definedness holds, we are sure of selecting a unique stationary equilibrium. According to Herings and Peeters (2000), the following theorem holds:

Theorem 4.1 *Assume $C(S)$ and $I(T)$. For an open set of stochastic games Γ and priors $p \in \Sigma$ with full Lebesgue measure, the stochastic tracing procedure is well-defined.*

We also claim the following result:

Theorem 4.2 *Assume $I(S)$ and $I(T)$. For an open set of stochastic games Γ and priors $p \in \Sigma$ with full Lebesgue measure, the stochastic tracing procedure is well-defined.*

Proof Since the stochastic tracing procedure when $I(S)$ and $I(T)$ is assumed can be written as a recurrent relation, the proof of this theorem is completely analogous to the proof in Herings and Peeters (2000). \square

For the alternatives where $C(T)$ is assumed, it is not possible to rewrite the payoffs into a recurrent relation. Therefore, to prove similar theorems with $C(T)$ assumed, calls for different techniques of proof. Nevertheless, it should be expected that generic well-definedness holds as well when $C(T)$ is assumed. Since the number of variables is one less than the number of constraints imposed by the concept of stationary equilibrium, there is one degree of freedom left. Under suitable transversality conditions, this should be sufficient to show well-definedness.

5 Analysis of alternatives

We proceed by analyzing the four alternatives defined so far. The first part of this section is devoted to lay bare that the choice about beliefs within states matter. As the proof of the pudding is in the eating, this is done by an example. Next, in the second part of this section the same will be done for the assumption with respect to beliefs across time.

5.1 Beliefs within states

We create an example for which it matters whether correlation or absence of correlation within states is assumed, in the sense that different stationary equilibria will be selected starting from the same prior. Before doing so, it is important to know the minimal size of

such an example. In case there are two players, assuming absence or presence of correlation within states does not matter since the number of opponents is 1.

Proposition 5.1 *For $n = 2$, the stochastic tracing procedure based on $I(S)$ is identical to one based on $C(S)$.*

So, we need $n \geq 3$ to construct an example.

Consider the normal form game of Figure 1. This is a special case of a stochastic game, i.e. a stochastic game with only one state and discount factor equal to 0. The advantage of considering a normal form games, is that it disconnects assumptions on beliefs within states and assumptions on beliefs across time. Note that this game possesses three stationary

	s^2	$s^{2'}$		
s^1	2, 2, 2	0, 0, 0	0, 0, 0	0, 0, 0
$s^{1'}$	0, 0, 0	0, 0, 0	0, 0, 0	1, 1, 1
	s^3		$s^{3'}$	

$\delta = 0$

Figure 1: Game Γ : $C(S)$ vs. $I(S)$.

equilibria: the pure stationary equilibria $(s^1; s^2; s^3)$ and $(s^{1'}; s^{2'}; s^{3'})$, and a mixed stationary equilibrium all players play their first strategy with probability $\sqrt{2} - 1$ and their second with probability $2 - \sqrt{2}$. Because the discount factor is taken equal to 0, it does not matter whether $I(T)$ or $C(T)$ is assumed.

Suppose we take the following prior:

$$p = \left(\left(\frac{1}{6}, \frac{5}{6} \right); \left(\frac{1}{2}, \frac{1}{2} \right); \left(\frac{2}{3}, \frac{1}{3} \right) \right).$$

Note that for the best response against the prior it does not matter whether $C(S)$ or $I(S)$ is assumed. Both alternatives lead to the same best response to all possible priors.

Here, the prior is chosen such that we have as starting point the point $(0, (s^1; s^{2'}; s^{3'}))$. Player 1 starts playing the equilibrium strategy of the first pure stationary equilibrium and players 2 and 3 start playing the equilibrium strategy of the second pure stationary equilibrium. To obtain a stationary equilibrium, at least one player has to switch to another pure strategy. It turns out that in the case of $I(S)$, player 2 or 3 is switching his strategy before player 1, whereas in the other case, $C(S)$, player 1 is the first player willing to switch his strategy.

5.1.1 When I(S) is assumed

When absence of correlation within states is assumed, Γ^t has the form as depicted in Figure 2. The first element in the first box is the payoff to player 1 if he plays s^1 and

	s^2	$s^{2'}$		s^3	$s^{3'}$
s^1	$\frac{1}{3}(1+t)(2+t)$ $\frac{1}{9}(1+5t)(2+t)$ $\frac{1}{6}(1+t)(1+5t)$	$\frac{1}{3}(1-t)(2+t)$ $\frac{5}{18}(1-t)(1-t)$ $\frac{1}{6}(1-t)(1+t)$	$\frac{2}{3}(1+t)(1-t)$ $\frac{2}{9}(1-t)(1+5t)$ $\frac{5}{12}(1-t)(1-t)$	$\frac{2}{3}(1-t)(1-t)$ $\frac{5}{18}(1-t)(1+2t)$ $\frac{1}{12}(1+t)(5+t)$	
$s^{1'}$	$\frac{1}{6}(1-t)(1-t)$ $\frac{1}{9}(1-t)(2+t)$ $\frac{1}{6}(1-t)(1+t)$	$\frac{1}{6}(1-t)(1+t)$ $\frac{1}{18}(1-t)(5+t)$ $\frac{1}{6}(1-t)(1-t)$	$\frac{1}{6}(1-t)(1+2t)$ $\frac{2}{9}(1-t)(1-t)$ $\frac{1}{12}(1-t)(5+t)$	$\frac{1}{6}(1+t)(1+2t)$ $\frac{1}{18}(1+2t)(5+t)$ $\frac{1}{12}(1+t)(5+t)$	

Figure 2: Game Γ^t when I(S) is assumed.

he expects player 2 to play according to the prior with probability $1 - t$ and s^2 with probability t and expects player 3 to play according to the prior with probability $1 - t$ and s^3 with probability t . So, player 1 expects player 2 to play $(\frac{1}{2}(1 - t) + t, \frac{1}{2}(1 - t))$ and player 3 to play $(\frac{2}{3}(1 - t) + t, \frac{1}{3}(1 - t))$ and the expected payoff when playing s^1 is therefore $(\frac{1}{2}(1 - t) + t)(\frac{2}{3}(1 - t) + t)2 = \frac{1}{3}(1 + t)(2 + t)$.

For all $t \in [0, 1]$ it is possible to compute all stationary equilibria of Γ^t . Since we are only interested in the stationary equilibrium selected, we only determine the path starting at $t = 0$ and terminating in $t = 1$. Obviously, by the way we constructed the prior, the path starts in the point $(s^1; s^{2'}; s^{3'})$. This point is a stationary equilibrium of Γ^t as long as $t \leq 1/10$. When $t = 1/10$, player 2 is indifferent between his two pure strategies. So, at $t = 1/10$ the path jumps from $(s^1; s^{2'}; s^{3'})$ to $(s^1; s^2; s^{3'})$. Next, the point $(s^1; s^2; s^{3'})$ is a stationary equilibrium of Γ^t until $t = 1/5(29 - 2\sqrt{189}) \approx 0.3009$, when player 3 is indifferent between his two pure strategies. Here, in $t = 1/5(29 - 2\sqrt{189})$, we jump from $(s^1; s^2; s^{3'})$ to $(s^1; s^2; s^3)$ which is a stationary equilibrium of Γ^t for all $t \in [1/5(29 - 2\sqrt{189}), 1]$. The point $(s^1; s^2; s^3)$ is therefore the stationary equilibrium selected by the stochastic tracing procedure when I(S) is assumed. The path is displayed graphically in Figure 3.

This figure shows a cube moving from $t = 0$ to $t = 1$. The cube itself displays the set of stationary strategy-combinations. The strategic possibilities of player 1, player 2 and player 3 are displayed in the horizontal, diagonal, respectively vertical direction. The point marked with a star is the point where all players play their first pure strategy.

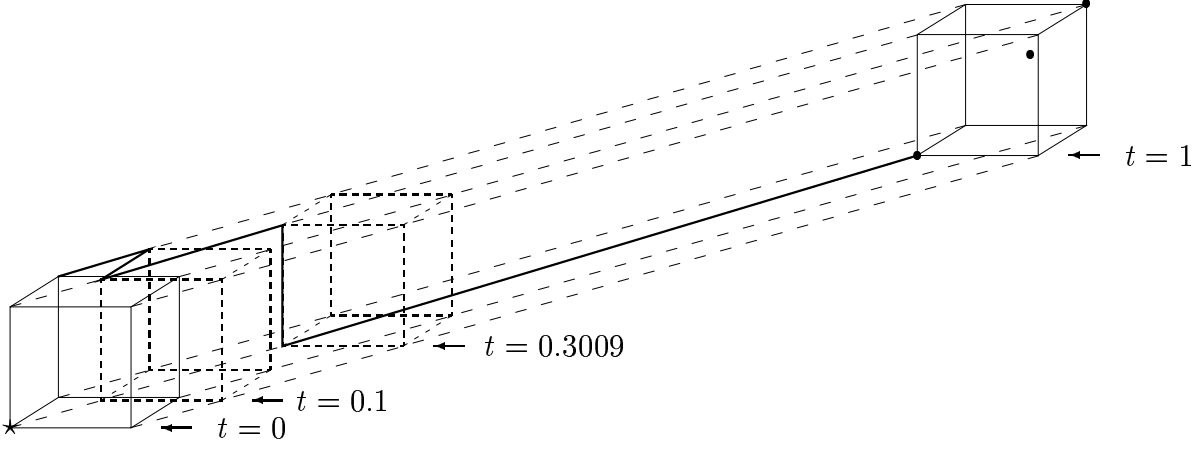


Figure 3: Feasible path when I(S) is assumed.

5.1.2 When C(S) is assumed

When correlation within states is assumed, Γ^t has the form as depicted in Figure 4. The

	s^2	$s^{2'}$		
s^1	$\frac{2}{3}(1-t)+2t$ $\frac{2}{9}(1-t)+2t$ $\frac{1}{6}(1-t)+2t$	$\frac{2}{3}(1-t)$ $\frac{5}{18}(1-t)$ $\frac{1}{6}(1-t)$	$\frac{2}{3}(1-t)$ $\frac{2}{9}(1-t)$ $\frac{5}{12}(1-t)$	$\frac{2}{3}(1-t)$ $\frac{5}{18}(1-t)$ $\frac{5}{12}(1-t)$
$s^{1'}$	$\frac{1}{6}(1-t)$ $\frac{2}{9}(1-t)$ $\frac{1}{6}(1-t)$	$\frac{1}{6}(1-t)$ $\frac{5}{18}(1-t)$ $\frac{1}{6}(1-t)$	$\frac{1}{6}(1-t)$ $\frac{2}{9}(1-t)$ $\frac{5}{12}(1-t)$	$\frac{1}{6}(1-t)+t$ $\frac{5}{18}(1-t)+t$ $\frac{5}{12}(1-t)+t$
	s^3		$s^{3'}$	

Figure 4: Game Γ^t when C(S) is assumed.

first element in the first box is the payoff to player 1 if he plays s^1 and he expects players 2 and 3 to play according to the prior with probability $1-t$ and their first pure strategy with probability t . So, player 1 expects with probability $1-t$ that players 2 and 3 are playing $(\frac{1}{2}, \frac{1}{2})$ respectively $(\frac{2}{3}, \frac{1}{3})$ and with probability t that players 2 and 3 are playing s^2 respectively s^3 . The expected payoff to player 1 when playing s^1 will therefore be equal to $((1-t)\frac{1}{2}\frac{2}{3}+t)2 = \frac{2}{3}(1-t) + 2t$.

Again, it is possible to compute all stationary equilibria of Γ^t , for $t \in [0, 1]$. And, again, we restrict ourselves to determine the path starting at $t = 0$ and terminating in $t = 1$. Just as before, the path starts in the point $(s^1; s^{2'}; s^{3'})$. This point is a stationary equilibrium point as long as $t \leq 1/3$. When $t = 1/3$, player 1 is indifferent between both

his pure strategies and will be the first player to switch to another pure strategy unlike in the previous case. At $t = 1/3$, the path jumps from $(s^1; s^{2'}; s^{3'})$ to $(s^{1'}; s^{2'}; s^{3'})$. The point $(s^{1'}; s^{2'}; s^{3'})$ is a stationary equilibrium of Γ^t up to $t = 1$ and this point is therefore the stationary equilibrium selected by the stochastic tracing procedure when C(S) is assumed. The path is displayed graphically in Figure 5 and is displayed in the same style as Figure 3.

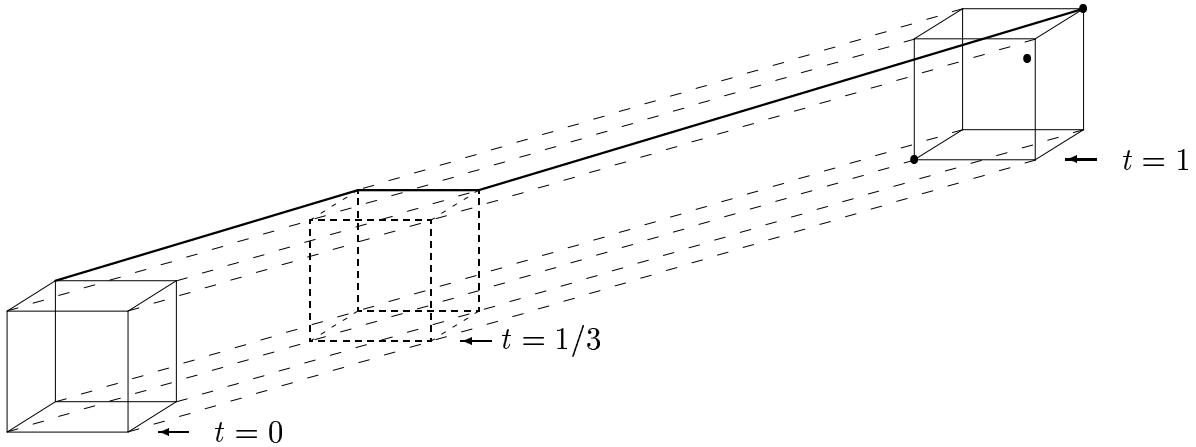


Figure 5: Feasible path when C(S) is assumed.

5.2 Beliefs across time

Again, it is important to know the minimal size of an example for which it matters whether correlation or absence of correlation across time is assumed. In the next proposition we claim that C(T) and I(T) select the same stationary equilibrium for stochastic games with one state (repeated game), when the process never leaves a state (finite number of independent repeated games) or when future payoffs are ignored (finite number of normal games).

Proposition 5.2 *When $\pi = \mathbb{1}_\Omega$ or $\delta = 0$, the stochastic tracing procedure based on $I(T)$ is identical to one based on $C(T)$.*

Proof Suppose we have a repeated game, i.e. $|\Omega| = 1$ and therefore trivially $\pi = \mathbb{1}_\Omega$. In this case, the total discounted payoffs for the four alternatives are:

$$V_{C(S), I(T)}^i(t; \rho) = (1 - t)u^i(p^{-i}, \rho^i) + tu^i(\rho) + \delta V_{C(S), I(T)}^i(t; \rho),$$

$$V_{C(S), C(T)}^i(t; \rho) = (1 - t)U^i(p^{-i}, \rho^i) + tU^i(\rho),$$

$$V_{I(S),I(T)}^i(t; \rho) = u^i((1-t)p^{-i} + t\rho^{-i}, \rho^i) + \delta V_{I(S),I(T)}^i(t; \rho),$$

and

$$V_{I(S),C(T)}^i(t; \rho) = \sum_{S \subseteq N \setminus \{i\}} (1-t)^S t^{N-S-1} U^i(p^S, \rho^{N \setminus S}).$$

First, note that

$$\begin{aligned} V_{C(S),C(T)}^i(t; \rho) &= (1-t)U^i(p^{-i}, \rho^i) + tU^i(\rho) \\ &= (1-t) [u^i(p^{-i}, \rho^i) + \delta U^i(p^{-i}, \rho^i)] + t [u^i(\rho) + \delta U^i(\rho)] \\ &= (1-t)u^i(p^{-i}, \rho^i) + tu^i(\rho) + \delta \cdot [(1-t)U^i(p^{-i}, \rho^i) + tU^i(\rho)] \\ &= (1-t)u^i(p^{-i}, \rho^i) + tu^i(\rho) + \delta V_{C(S),C(T)}^i(t; \rho). \end{aligned}$$

And secondly, note that

$$\begin{aligned} V_{I(S),C(T)}^i(t; \rho) &= \sum_{S \subseteq N \setminus \{i\}} (1-t)^S t^{N-S-1} U^i(p^S, \rho^{N \setminus S}) \\ &= \sum_{S \subseteq N \setminus \{i\}} (1-t)^S t^{N-S-1} [u^i(p^S, \rho^{N \setminus S}) + \delta U^i(p^S, \rho^{N \setminus S})] \\ &= \sum_{S \subseteq N \setminus \{i\}} (1-t)^S t^{N-S-1} u^i(p^S, \rho^{N \setminus S}) \\ &\quad + \delta \cdot \sum_{S \subseteq N \setminus \{i\}} (1-t)^S t^{N-S-1} U^i(p^S, \rho^{N \setminus S}) \\ &= u^i((1-t)p^{-i} + t\rho^{-i}, \rho^i) + \delta V_{I(S),C(T)}^i(t; \rho). \end{aligned}$$

Finally, observe that

$$V_{C(S),I(T)}^i(t; \rho) = V_{C(S),C(T)}^i(t; \rho),$$

and that

$$V_{I(S),I(T)}^i(t; \rho) = V_{I(S),C(T)}^i(t; \rho)$$

for all $(t; \rho) \in [0, 1] \times \Sigma$. So, for repeated games C(T) and I(T) are equivalent.

From this the equivalence between C(T) and I(T) follows trivially when $\pi = \mathbb{1}_\Omega$ as being a finite number of independent repeated games.

When $\delta = 0$, the stochastic game is equivalent to a one shot game in which time plays no role when stationary strategies are assumed. \square

Consider the stochastic game of Figure 6. The stochastic game displayed in that figure is a stochastic game with two states. In the first state both players have two pure strategies, whereas in the second state both players have one strategy. The upper-left corner of each box contains the payoffs. The transition probabilities are displayed in the lower-right corner, where the first element is the probability that in the next stage the game will be

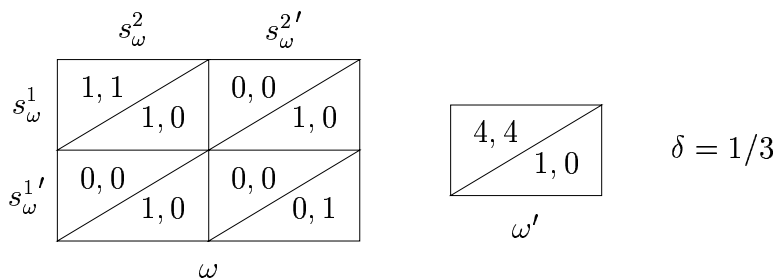


Figure 6: Game Γ : C(T) vs. I(T).

in the first state, the second element is the probability that the game will be in the second state at the next stage.

When both players play their second strategy in state ω , they both expect to earn 0 at the current stage, go to state ω' , earn 4 at the next stage and then return to state ω . Notice that the game at ω' is completely degenerate. There is no need for players to make a choice. They simply collect a payoff of 4. When ω is the initial state, both players expect to earn $\frac{4}{3}(1 + \frac{1}{9} + \frac{1}{81} + \dots) = \frac{3}{2}$ in total when both always play their second pure strategy. Suppose the box that belongs to strategy-combination $(s_\omega^{1'}; s_\omega^{2'})$ had 1,1 in the upper-left corner and 1,0 in the lower-right corner. Then the game is always in state ω when ω is the initial state, so coincides with a repeated game. When again both players play their second pure strategy at all times, they both expect again to earn $(1 + \frac{1}{3} + \frac{1}{9} + \dots = \frac{3}{2})$. Nevertheless, the two resulting stochastic games are not strategically equivalent when mixed strategies are played. It turns out that $((\frac{1}{2}, \frac{1}{2}); (\frac{1}{2}, \frac{1}{2}))$ is not a stationary equilibrium in the stochastic game of Figure 6 unlike in the adapted stochastic game (the repeated game). Although the stochastic game of Figure 6 is close to a repeated game, for which C(T) and I(T) are equivalent, it is not quite.

The stochastic game of Figure 6 possesses three stationary equilibria: the pure stationary equilibria $(s_\omega^1; s_\omega^2)$ and $(s_\omega^{1'}; s_\omega^{2'})$, and the symmetric mixed stationary equilibrium where both players play their first strategy with probability $4 - 2\sqrt{3} \approx 0.5359$ and the other strategy with the rest of the probability mass.

Just like in the previous example, we want the starting point of the stochastic tracing procedure not to be a stationary equilibrium strategy-combination of the original game. This can be arranged by taking for one player the prior play of the first strategy smaller than $4 - 2\sqrt{3}$ and the reverse for the other. Without loss of generality, let the first player mentioned be player 1. Given such a prior, player 1 will prefer strategy s_ω^1 and player 2 will prefer strategy $s_\omega^{2'}$. Consider the prior

$$p = ((\frac{1}{2}, \frac{1}{2}); (\frac{2}{3}, \frac{1}{3})),$$

which satisfies this property. Note that the starting point, in this case $(s_\omega^1; s_\omega^{2'})$, is independent of the choice made between the two assumptions I(T) and C(T). Because the stochastic game depicted is a game with two players, it does not matter whether C(S) or I(S) is assumed.

5.2.1 When I(T) is assumed

In Figure 7 the stage game for $t = 0$ is given. The figure is almost similar to Figure 6. The

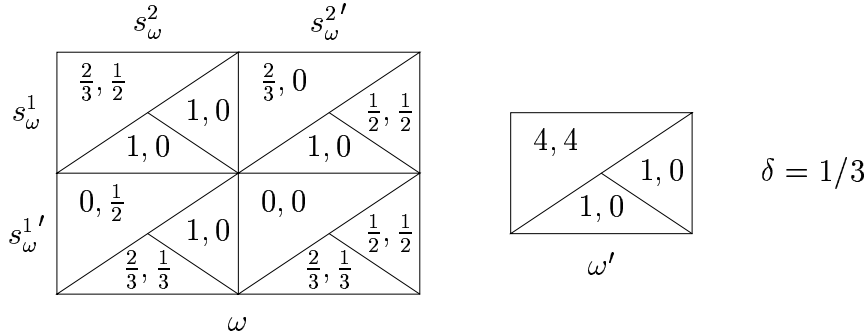


Figure 7: Game Γ^0 when I(T) is assumed.

difference is found in the lower-right corner of the boxes, which is divided into two sectors. The lower-left sector displays the transition probabilities as player 1 thinks they are, the upper-right sector the transition probabilities as player 2 thinks they are. For example, when player 1 plays $s_\omega^{1'}$ he expects to receive 0 and expects that in the next stage the state will be ω with probability $2/3$ and ω' with probability $1/3$.

The stage games Γ_ω^t can now be defined as $\Gamma_\omega^t = (1 - t)\Gamma_\omega^0 + t\Gamma_\omega^1$ for all $t \in (0, 1)$. For all $t \in [0, 1]$, it is possible to compute all stationary equilibria of Γ^t , but since we are only interested in the stationary equilibrium selected, we only determine the path starting at $t = 0$ and terminating in $t = 1$. As mentioned before, the path starts in the point $(s_\omega^1; s_\omega^{2'})$. This point is a stationary equilibrium point as long $t \leq \frac{1}{2}(23 - \sqrt{513}) \approx 0.1752$, the moment player 1 is indifferent between both his pure strategies. Player 2 is not willing to switch strategy as long as $t \leq \frac{1}{6}(11 - \sqrt{97}) \approx 0.1919$. So, player 1 is the first to change his strategy and at $t = \frac{1}{2}(23 - \sqrt{513})$ the path jumps from $(s_\omega^1; s_\omega^{2'})$ to $(s_\omega^{1'}; s_\omega^{2'})$. This point is a stationary equilibrium point for all $t \in [\frac{1}{2}(23 - \sqrt{513}), 1]$. The point $(s_\omega^{1'}; s_\omega^{2'})$ is therefore the stationary equilibrium selected by the stochastic tracing procedure when I(T) is assumed. The path is displayed graphically in Figure 8.

This figure shows the Cartesian product of the strategy space, a square, and the interval $[0, 1]$. The strategic possibilities of player 1 and player 2 are displayed in the horizontal,

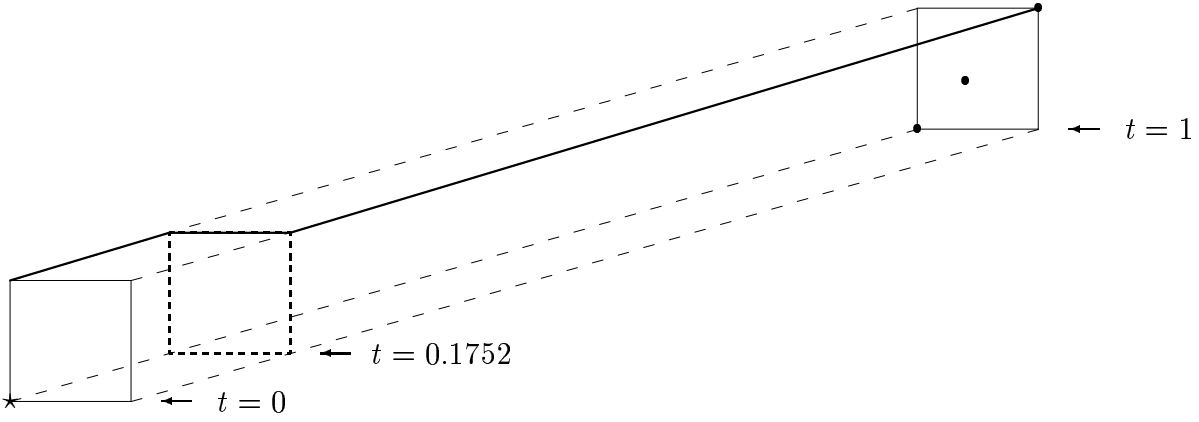


Figure 8: Feasible path when I(T) is assumed.

respectively vertical direction. The point marked with a star corresponds to both players playing their first pure strategy.

5.2.2 When C(T) is assumed

When C(T), correlation across time, is assumed, Γ^t cannot be represented easily in a figure. Again, the path starts in the point $(s_\omega^1; s_\omega^{2'})$. Given that player 2 plays $s_\omega^{2'}$ during the whole play with probability t and according to the prior during the whole play with probability $1 - t$, the best stationary response of player 1 is playing s_ω^1 as long as t is less than $4/19$. When t is larger than $4/19$ his other pure stationary strategy, that is $s_\omega^{1'}$, will be his best stationary response. Vice versa, when player 2 expects player 1 to play according to the prior the whole stochastic game with probability $1 - t$ and s_ω^1 the whole play with probability t , his best stationary response is playing strategy $s_\omega^{2'}$ when $t \leq 1/15$ and s_ω^2 otherwise. So, player 2 will be the first player to switch strategy, since $1/15 < 4/19$. Therefore, at $t = 1/15$, the path will jump from $(s_\omega^1; s_\omega^{2'})$ to $(s_\omega^1; s_\omega^2)$. From $t = 1/15$ up to $t = 1$ the point $(s_\omega^1; s_\omega^2)$ is the stationary equilibrium and is also the stationary equilibrium selected by the stochastic tracing procedure when C(T) is assumed. The path is plotted in Figure 9 and is displayed in the same style as Figure 8.

5.3 Extending the linear tracing procedure

The first subsection showed that the choice between assuming absence of correlation or assuming presence of correlation within states can cause different stationary equilibria to be selected. Although this was shown by means of a normal form game, the same result would have been found if we made it a repeated game by taking a positive discount factor,

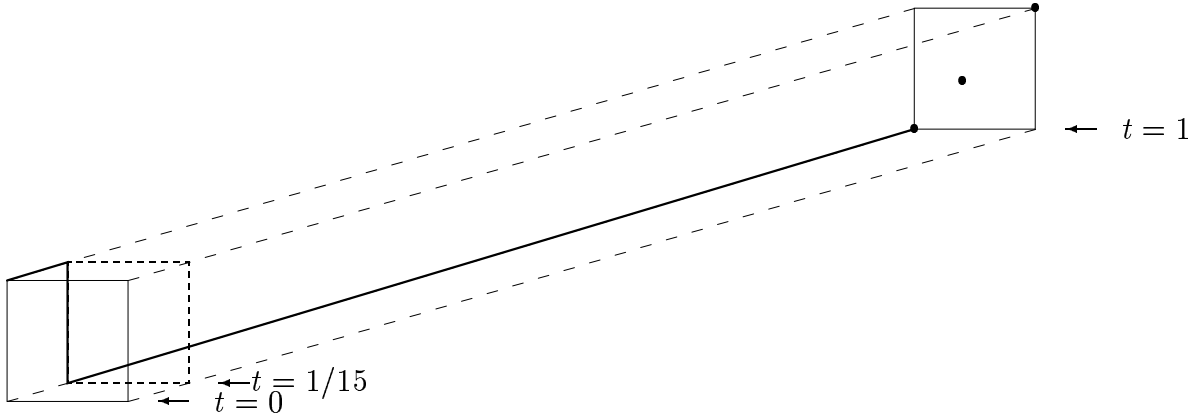


Figure 9: Feasible path when $C(T)$ is assumed.

i.e. $\delta \in (0, 1)$.

In the selection theory of Harsanyi and Selten (1988), the linear tracing procedure of Harsanyi (1975) is used to select on a Nash equilibrium of a normal form game. The structure of the linear tracing procedure there assumed beliefs to be correlated. The most natural extension of the linear tracing procedure is therefore the one that assumes correlation within states. For instance, this choice implies that for repeated games, the stationary equilibria which is selected is the repeated Nash equilibrium that is selected by the linear tracing procedure for the stage game using the same prior.

The second subsection showed that the choice between assuming absence of correlation or assuming presence of correlation across time can cause different stationary equilibria to be selected. When $C(T)$ is assumed, a point (t, ρ) on the feasible path is not necessarily a stationary equilibrium of Γ^t . For some player i there might be a better behavior strategy than ρ^i against ρ^{-i} in Γ^t which is non-stationary, i.e. a best stationary response is not necessarily a best response in the broader class of behavior strategies. However, assuming that beliefs are not correlated across time, that is assuming $I(T)$, captures the assumption of stationarity. The beliefs of a player will depend only on the state reached and not on the stage at which it is reached. We would therefore argue that the assumption of $C(S)$ and $I(T)$ leads to a stochastic tracing procedure that is the most natural extension of the linear tracing procedure of Harsanyi and Selten.

6 Summary

In this paper the linear tracing procedure of Harsanyi (1975) is extended in order to select on an equilibrium for stochastic games, more precisely, to select on a stationary equilibrium

for stochastic games. There are four reasonable extensions. After having defined all four properly, it is shown by examples that these extensions are independent. The most natural extension, assumes that players hold correlated beliefs within states, but do not hold correlated beliefs across time, since this is the only possible extension consistent with the formulation of Harsanyi for normal form games that captures stationarity. This extension is defined as the stochastic tracing procedure. It follows immediately from the results in Herings and Peeters (2000) that it is well-defined for almost all stochastic games.

References

- Aumann, R.J. (1974), "Subjectivity and correlation in randomized strategies," *Journal of Mathematical Economics*, 1, 67-96.
- Aumann, R.J. (1987), "Correlated equilibrium as an expression of Bayesian rationality," *Econometrica*, 55, 1-18.
- Bernheim, D. (1984), "Rationalizable strategic behavior," *Econometrica*, 52, 1007-1028.
- Brandenburger, A., and E. Dekel (1987), "Rationalizability and correlated equilibria," *Econometrica*, 55, 1391-1402.
- Fink, A.M. (1964), "Equilibrium in a stochastic n -person game," *Journal of Science of Hiroshima University Series A-I*, 28, 89-93.
- Haller, H., and R. Lagunoff (2000), "Genericity and Markovian behavior in stochastic games," *Econometrica*, 68, 1231-1248.
- Harsanyi, J.C. (1975), "The tracing procedure: A Bayesian approach to define a solution for n -person noncooperative games," *International Journal of Game Theory*, 4, 61-94.
- Harsanyi, J.C., and R. Selten (1988), *A General Theory of Equilibrium Selection in Games*, MIT Press, Cambridge, Massachusetts.
- Herings, P.J.J., and R.J.A.P. Peeters (2000), "Stationary equilibria in stochastic games: Structure, selection, and computation," Meteor Research Memorandum 00/031, Universiteit Maastricht, Maastricht.
- Herings, P.J.J., and R.J.A.P. Peeters (2001), "A differentiable homotopy to compute Nash equilibria of n -person games," *Economic Theory*, 18, 159-185.
- Maskin, E., and J. Tirole (2001), "Markov perfect equilibrium, I: Observable actions,"

Journal of Economic Theory, 100, 191-219.

McLennan, A. (1999), "The expected number of Nash equilibria of a normal form game," mimeo.

Nash, J.F. (1951), "Non-cooperative games," *Annals of Mathematics*, 54, 286-295.

Pearce, D. (1984), "Rationalizable strategic behavior and the problem of perfection," *Econometrica*, 52, 1029-1050.

Shapley, L.S. (1953), "Stochastic games," *Proceedings of the National Academy of Sciences of the United States of America*, 39, 1095-1100.

Sobel, M.J. (1971), "Non-cooperative stochastic games," *The Annals of Mathematical Statistics*, 42, 1930-1935.

Takahashi, M. (1964), "Equilibrium points of stochastic, noncooperative n -person games," *Journal of Science of Hiroshima University Series A-I*, 28, 95-99.