CSEF
Centre for Studies in Economics and Finance

WORKING PAPER NO. 661

Asymptotic Behavior of Subgame Perfect Nash Equilibria in Stackelberg Games

Francesco Caruso, Maria Carmela Ceparano, and Jacqueline Morgan

December 2022

University of Naples Federico II

University of Salerno

Bocconi University, Milan

CSEF - Centre for Studies in Economics and Finance
DEPARTMENT OF ECONOMICS AND STATISTICS – UNIVERSITY OF NAPLES FEDERICO II
80126 NAPLES - ITALY
Tel. and fax +39 081 675372 – e-mail: csef@unina.it
ISSN: 2240-9696

WORKING PAPER NO. 661

Asymptotic Behavior of Subgame Perfect Nash Equilibria in Stackelberg Games

Francesco Caruso*, Maria Carmela Ceparano†, and Jacqueline Morgan‡

Abstract
The study on how equilibria behave when perturbations occur in the data of a game is a fundamental theme, since actions and payoffs of the players may be affected by uncertainty or trembles. In this paper we investigate the asymptotic behavior of the subgame perfect Nash equilibrium (SPNE) in one-leader one-follower Stackelberg games under perturbations both of the action sets and of the payoff functions. To pursue this aim, we consider a general sequence of perturbed Stackelberg games and a set of assumptions that fit the usual types of perturbations. We study if the limit of SPNEs of the perturbed games is an SPNE of the original game and if the limit of SPNE outcomes of perturbed games is an SPNE outcome of the original game. We fully positively answer when the follower's best reply correspondence is single valued. When the follower's best reply correspondence is not single valued, the answer is positive only for the SPNEs outcomes; whereas the answer for SPNEs may be negative, even if the perturbation does not affect the sets and affects only one payoff function. However, we show that under suitable non-restrictive assumptions it is possible to obtain an SPNE starting from the limit of SPNEs of perturbed games, possibly modifying the limit at just one point.

Keywords: Subgame perfect Nash equilibrium two-player Stackelberg game; action and/or payoff perturbation; convergence; asymptotic behavior; variational stability.

* Università di Napoli Federico II. E-mail: francesco.caruso@unina.it

† Università di Napoli Federico II. E-mail: mariacarmela.ceparano@unina.it

‡ Università di Napoli Federico II and CSEF. E-mail: morgan@unina.it

1 Introduction

The investigation on the asymptotic behavior of equilibria when perturbations affect the actions and/or the payoffs of the players represents a relevant issue in Game Theory, (see, for example, the seminal books [18] and [38]) and it goes back to classical topics in Optimization, as the so-called *variational stability* or *qualitative stability* (see, for example, the books [4, 1, 11, 13, 33]). Such an investigation requires to answer to questions like: given a perturbation of the original game, does the sequence of equilibria of the perturbed games converge to an equilibrium of the original game? what conditions on the perturbations ensure that the limit of the sequence of equilibria of the perturbed games is an equilibrium of the original game? and what about the convergence of players' payoffs at equilibrium?
While these questions have been extensively analyzed and positively addressed in simultaneous-move games, the literature on the behavior of selections of Nash equilibria in sequential games under perturbations is more confined due to the higher intrinsic complexity of the model. The results concerning the subgame perfect Nash equilibrium ([35], henceforth SPNE), a classic solution concept in sequential frameworks, can be collected in the following four strands of literature. Recall that an SPNE is a strategy profile and an SPNE *outcome* is the action profile played in an SPNE.

- For finite-action extensive-form games, properties of robustness under perturbations gave birth to renowned refinements of the SPNE: see e.g. [34, 21, 31, 37].
- In infinite-action games of perfect information with infinite horizon, the limit of ϵ-SPNEs of perturbed games has been shown to be an SPNE of the original game: see [15, 16, 5] where the perturbation consists of finite-horizon (truncated) games; see [7] where the payoff functions are perturbed via sequences of simple functions.
- For N-player N-stage infinite-action games of perfect information, the limit points of SPNE outcomes of finite-action (discretized) games are shown to be SPNE outcomes of the original infinite-action game and an analogous result does not hold for the SPNE strategies ([20]). Results on such an upper semicontinuity property of the SPNE outcomes have been proved also for infinite-action games with infinite horizon in [6].
- For two-stage infinite-action games of perfect information (one-leader one/two-follower Stackelberg games), in [30] a constructive method to approach an SPNE relying on the perturbation of the followers' payoff functions via the Tikhonov regularization is proposed. The authors showed that, starting from the limit of the unique SPNEs of the Tikhonov-perturbed Stackelberg games, one can define an SPNE of the original game. Inspired by this result, in [8] a selection method for SPNEs via a learning approach based on the proximal regularization of the payoff functions of both players has been defined. The selection is motivated by the players' aversion against the costs to face when they deviate from their current actions. Recently, the same methodology has been used in [29] where a perturbation involving the Shannon's entropy affects the payoff functions.

All the results previously mentioned concern specific kinds of perturbation of the action sets or of the payoff functions and emphasize two features. On the one hand, the SPNE can exhibit a "bad" asymptotic behavior under perturbations of the data of the game. On the other hand, positive results can be achieved by looking at the SPNE outcomes, at the ϵ-SPNE and at the construction of an SPNE when the limit of perturbed SPNEs is not. Hence, it is natural to ask: what conditions a game and a perturbation must satisfy in order to guarantee that

- the limit of SPNEs of perturbed games is an SPNE of the original game?
- an SPNE can be obtained even when such a limit is not an SPNE?
- the limit of SPNE outcomes of perturbed games is an SPNE outcome of the original game?

The aim of this paper is to answer to these questions. To pursue such an aim, we investigate the asymptotic behavior of the SPNEs when general perturbations affect the action sets and the payoff functions. The analysis is provided for the class of continuous one-leader one-follower Stackelberg games, that is two-player two-stage perfect-information games where the players have a continuum of actions.

More precisely, we first consider a general sequence of perturbed Stackelberg games obtained by perturbing both the action sets and the payoff functions of the original game and we define the set of assumptions on such perturbed games that will be used in our investigation. The assumptions easily accommodate many types of perturbations of sets and of functions broadly used in literature on approximation theory.

Then, we provide the analysis on the asymptotic behavior of SPNEs under perturbation in three steps. First, the case of uniqueness of the follower's optimal reaction is addressed. In this situation both the SPNEs and the SPNE outcomes exhibit a "nice" asymptotic behavior under perturbations; in fact the limit of the sequence of SPNEs of the perturbed games is an SPNE of the original game and an analogous result happens for the SPNE outcomes. Secondly, the case of non-uniqueness of the follower's optimal reaction is considered. In this situation stability holds only for the SPNE outcomes (this represents a complementary result of [20]). Instead, stability for SPNEs is not guaranteed. Indeed, we show through an illustrative example that the limit of the sequence of SPNEs of the perturbed games could not be an SPNE of the original game, despite the perturbation affects only one player's payoff function, the action sets are unperturbed and both the original and each perturbed game have a unique SPNE. Nevertheless, stability for SPNEs actually happens for a particular class of games and perturbations. Thirdly, in light of the previous bad behavior result concerning the SPNEs, the issue of how "far" is the limit of the SPNEs of perturbed games from an SPNE of the original game is addressed. We prove that, starting from the limit of SPNEs of perturbed games, it is possible to obtain an SPNE of the original game by modifying the value of the follower's strategy at just one point. Hence, *stability-but-for-a-point* occurs for the SPNEs in Stackelberg games.

Therefore, the whole analysis allows to obtain a broad and complete view of the behavior of SPNEs in Stackelberg games when there are perturbations of all data of the game.

We mention that the asymptotic behavior of other solutions concepts in Stackelberg games has been also investigated in literature. See, for example, [25] for the *strong Stackelberg equilibrium* and the last works [24, 23] on the *weak Stackelberg equilibrium*. Note that stability results for both concepts have been achieved thanks to the introduction of approximate solutions. For further considerations and other results, see for example [9].

The paper is organized as follows. Section 2 displays the framework: the general sequence of perturbed games is defined and some convergence notions for sets and for functions are recalled. Section 3 is devoted to the investigation on the asymptotic behavior of SPNEs. We first display and discuss the assumptions that we consider in the analysis. Then, the following achievements are obtained: the stability for SPNEs and SPNE outcomes under the uniqueness of the follower's optimal reaction, the bad behavior of SPNEs and the stability for SPNE outcomes in the case of non-uniqueness of the follower's optimal reaction. In Section 4 we prove the key-result on the stability but for a point. Finally, conclusions and directions for future research are provided in Section 5.

2 Preliminaries

In the whole paper, we deal with a Stackelberg game denoted by Γ, a two-person two-stage perfect-information game where one player, called leader, chooses her action in the first stage and then one player, called follower, chooses his action after having observed the action of the leader. The action sets of the leader and of the follower are denoted by X and Y and they are assumed to be subsets of two Euclidean spaces $\mathbb{X}$ and $\mathbb{Y}$, respectively; their payoff functions are denoted by l and f, respectively, and they are real-valued functions defined on the set of action profiles $X \times Y$. To emphasize the data of the game, we refer to Γ as $\langle X, Y, l, f\rangle$. We denote with B the best reply correspondence of the follower, that is the set-valued map defined on X by

$$B(x) := \arg\max_{y\in Y} f(x,y) = \{y \in Y \mid f(x,y) \geq f(x,z) \text{ for each } z \in Y\}, \tag{1}$$

and with Y^X the set of strategies of the follower, that is $Y^X := \{\varphi \mid \varphi\colon X \to Y\}$. In this setting, an SPNE is defined as follows.

Definition 2.1 ([35]) A strategy profile $(\bar{x}, \bar{\varphi}) \in X \times Y^X$ is a subgame perfect Nash equilibrium of Γ if the following conditions hold:

(SPE1) $\bar{\varphi}(x) \in B(x)$ for each $x \in X$,

(SPE2) $\bar{x} \in \arg\max_{x\in X} l(x, \bar{\varphi}(x))$.

It is worth recalling that the existence of SPNEs is guaranteed under the compactness of the action sets and the continuity of the payoff functions ([17, Theorem 1]).

Having in mind the investigation on the asymptotic behavior of SPNEs in two-player Stackelberg games under perturbations, let us consider a general sequence of perturbed games $(\Gamma_n)_n$ obtained from Γ by perturbing the players' action sets and payoff functions. More precisely, let

$$\Gamma_n := \langle X_n, Y_n, l_n, f_n\rangle \quad \text{for each } n \in \mathbb{N},$$

that is, Γ_n is a two-player Stackelberg game whose action sets are X_n and Y_n, subsets of $\mathbb{X}$ and $\mathbb{Y}$ respectively, and whose payoff functions are l_n and f_n, real-valued functions defined on $X_n \times Y_n$.

About convergences. For the sake of completeness, we list some notions of convergence for points, sets and functions that we will use in the paper (see, e.g., [22, 25, 3]). Let $(U_n)_n$ and $(V_n)_n$ be sequences of subsets of the Euclidean spaces $\mathbb{U}$ and $\mathbb{V}$, respectively, and let $(\varphi_n)_n$ be a sequence of functions with $\varphi_n \colon U_n \to V_n$ for each $n \in \mathbb{N}$.

- The sequence $(u_n)_n \subseteq \mathbb{U}$ *converges* to $u \in \mathbb{U}$ (equivalently, u is the limit of the sequence $(u_n)_n$), that is $\lim_{n\to+\infty} u_n = u$, if for each $\epsilon > 0$ there exists $\nu_\epsilon \in \mathbb{N}$ such that $\|u_n - u\|_{\mathbb{U}} < \epsilon$ for each $n > \nu_\epsilon$;
- The *lower limit* of the sequence of sets $(U_n)_n$ is the set

$$\operatorname*{Liminf}_{n\to+\infty} U_n \coloneqq \{u \in \mathbb{U} \mid \exists\,(u_n)_n \text{ converging to } u \text{ s.t. } u_n \in U_n\,\forall n \in \mathbb{N}\},$$

  that is, $u \in \operatorname{Liminf}_{n\to+\infty} U_n$ if and only if u is the limit of a sequence $(u_n)_n$ with $u_n \in U_n$ for each $n \in \mathbb{N}$;
- The *upper limit* of the sequence of sets $(U_n)_n$ is the set

$$\begin{aligned}\operatorname*{Limsup}_{n\to+\infty} U_n \coloneqq \{u \in \mathbb{U} \mid \exists\,(u_k)_k &\text{ converging to } u \\ &\text{s.t. } u_k \in U_{n_k} \text{ for a subsequence } (n_k)_k \subseteq \mathbb{N}\},\end{aligned}$$

  that is, $u \in \operatorname{Limsup}_{n\to+\infty} U_n$ if and only if u is a cluster point of a sequence $(u_n)_n$ with $u_n \in U_n$ for each $n \in \mathbb{N}$.
- The sequence of sets $(U_n)_n$ *converges* to U, that is $\operatorname{Lim}_{n\to+\infty} U_n = U$, if $\operatorname{Limsup}_{n\to+\infty} U_n \subseteq U \subseteq \operatorname{Liminf}_{n\to+\infty} U_n$.
- The sequence $(\varphi_n)_n$ *pointwise converges* to the function $\varphi \colon U \to V$ (equivalently, φ is the pointwise limit of the sequence $(\varphi_n)_n$) if the sequence $(\varphi_n(u))_n$ converges to $\varphi(u)$ for each $u \in U$, that is $\lim_{n\to+\infty} \varphi_n(u) = \varphi(u)$ for each $u \in U$;
- The sequence $(\varphi_n)_n$ *continuously converges* to the function $\varphi \colon U \to V$ (equivalently, φ is the continuous limit of the sequence $(\varphi_n)_n$) if for each sequence $(u_n)_n$, such that $u_n \in U_n$ for each $n \in \mathbb{N}$, converging to $u \in U$ the sequence $(\varphi_n(u_n))_n$ converges to $\varphi(u)$, that is $\lim_{n\to+\infty} u_n = u$ implies $\lim_{n\to+\infty} \varphi_n(u_n) = \varphi(u)$.

Obviously, if $U_n = U$ for each $n \in \mathbb{N}$, every continuously convergent sequence is also pointwise convergent and the pointwise limit coincides with the continuous limit.
We say that the sequence $(u_n, \varphi_n)_n$ with $(u_n, \varphi_n) \in U_n \times {V_n}^{U_n}$ for each $n \in \mathbb{N}$ *pointwise converges* to $(u, \varphi) \in U \times V^U$ (equivalently, (u, φ) is the pointwise limit of the sequence $(u_n, \varphi_n)_n$), if $(u_n)_n$ converges to u and $(\varphi_n)_n$ pointwise converges to φ.

Now, let $(\bar{x}_n, \bar{\varphi}_n)_n$ be a pointwise convergent sequence of strategy profiles such that $(\bar{x}_n, \bar{\varphi}_n)$ is an SPNE of Γ_n for each $n \in \mathbb{N}$. The main question one aims to answer in the investigation on the asymptotic behavior of SPNEs under perturbation is:

(Q) What conditions on Γ and on the perturbed games $(\Gamma_n)_n$ guarantee that the pointwise limit of the sequence $(\bar{x}_n, \bar{\varphi}_n)_n$ is an SPNE of the original game Γ? And what about the limit of the sequence of the outcomes $(\bar{x}_n, \bar{\varphi}_n(\bar{x}_n))_n$?

In the next section we will give answers to this question.

3 Answers to question (Q)

We first display the assumptions that we will use in the sequel of the paper. Recall that $\Gamma = \langle X, Y, l, f\rangle$ and $\Gamma_n = \langle X_n, Y_n, l_n, f_n\rangle$.

($\mathcal{L}_1$) $X \subseteq X_n$ for each $n \in \mathbb{N}$ and $\operatorname{Liminf}_n X_n \subseteq X$;

($\mathcal{L}_2$) the sequence $(l_n)_n$ continuously converges to l;

($\mathcal{F}_1$) $Y \subseteq Y_n$ for each $n \in \mathbb{N}$ and $\operatorname{Limsup}_n Y_n \subseteq Y$;

($\mathcal{F}_2$) the sequence $(f_n)_n$ continuously converges to f;

($\mathcal{F}_3$) there exists a compact set $\Omega \subseteq \mathbb{Y}$ such that $\bigcup_{n\in\mathbb{N}} Y_n \subseteq \Omega$.

For the sake of readability, we denote by *($\mathcal{A}$)* the set of all the assumptions *($\mathcal{L}_1$)*, *($\mathcal{L}_2$)*, *($\mathcal{F}_1$)*, *($\mathcal{F}_2$)*, *($\mathcal{F}_3$)*.
Assumption *($\mathcal{A}$)* accommodates the particular cases where the perturbation of Γ involves either only the action sets or only the payoff functions. In fact, on the one hand, if $X_n = X$ and $Y_n = Y$ for each $n \in \mathbb{N}$ and if X is closed and Y is compact, then assumptions *($\mathcal{L}_1$)*, *($\mathcal{F}_1$)*, *($\mathcal{F}_3$)* are satisfied. On the other hand, if l and f are continuous over $\bigcup_{n\in\mathbb{N}} X_n \times Y_n$ and if $l_n(x, y) = l(x, y)$ and $f_n(x, y) = f(x, y)$ for each $(x, y) \in X_n \times Y_n$ and each $n \in \mathbb{N}$, then assumptions *($\mathcal{L}_2$)*, *($\mathcal{F}_2$)* hold.
Assumptions *($\mathcal{L}_1$)*, *($\mathcal{F}_1$)* require that the action sets of the perturbed games approach the sets X and Y from outside. This is a common condition in approximation theory (see, e.g, [26, 32]). Note that assumptions *($\mathcal{L}_1$)*, *($\mathcal{F}_1$)*, *($\mathcal{F}_3$)* imply that X is closed and Y is compact (see, e.g., [3]).
A traditional instance of perturbation of sets satisfying assumptions *($\mathcal{L}_1$)*, *($\mathcal{F}_1$)* comes by defining X and Y as the sets of solutions to a finite number of inequalities. In fact, let

$$X = \bigcap_{i=1}^{m_X} \{x \in \mathbf{X} \mid g_i(x) \leq 0\} \quad \text{and} \quad Y = \bigcap_{j=1}^{m_Y} \{y \in \mathbf{Y} \mid h_j(y) \leq 0\},$$

where $\mathbf{X}$ and $\mathbf{Y}$ are subsets of $\mathbb{X}$ and $\mathbb{Y}$, respectively, and g_i and h_j are continuous real-valued functions defined on $\mathbf{X}$ and $\mathbf{Y}$, respectively. Consider two sequences $(\epsilon_n^X)_n$ and $(\epsilon_n^Y)_n$ of non-negative real numbers decreasing to 0 and define the perturbed sets as follows:

$$X_n = \bigcap_{i=1}^{m_X} \{x \in \mathbf{X} \mid g_i(x) \leq \epsilon_n^X\} \quad \text{and} \quad Y_n = \bigcap_{j=1}^{m_Y} \{y \in \mathbf{Y} \mid h_j(y) \leq \epsilon_n^Y\},$$

for each $n \in \mathbb{N}$. Then, assumptions *($\mathcal{L}_1$)*, *($\mathcal{F}_1$)* are satisfied (actually, $\text{Lim}_n X_n = X$ and $\text{Lim}_n Y_n = Y$). Assumption *($\mathcal{F}_3$)* holds provided that either $\mathbf{Y}$ or Y_1 is a compact set.
Assumptions *($\mathcal{L}_2$)*, *($\mathcal{F}_2$)* hold for many kinds of perturbations of the payoff functions. A broadly used instance consists in adding to l and f a noisy term that vanishes when n grows. Let k_l and k_f be real-valued functions defined on $X \times Y$ and define

$$l_n(x, y) = l(x, y) + \alpha_n k_l(x, y) \quad \text{and} \quad f_n(x, y) = f(x, y) + \beta_n k_f(x, y),$$

for each $(x, y) \in X \times Y$ and each $n \in \mathbb{N}$, where $\alpha_n, \beta_n \in \mathbb{R}$ (for the sake of readability, we suppose that $X_n = X$ and $Y_n = Y$ for each $n \in \mathbb{N}$). The noisy terms $\alpha_n k_l(x, y)$ and $\beta_n k_f(x, y)$ can be interpreted as discomforts, costs, or general entropic penalties (see, e.g., [10] and [2] where entropic and costs-to-move perturbations, respectively, are used to approach Nash equilibria in simultaneous-move games). If X and Y are compact, l and f are continuous and the sequences $(\alpha_n)_n$ and $(\beta_n)_n$ converge to 0, then the assumptions *($\mathcal{L}_2$)*, *($\mathcal{F}_2$)* are satisfied.
Finally, looking at assumptions *($\mathcal{L}_1$)* and *($\mathcal{F}_1$)*, we point out that *($\mathcal{L}_1$)* is strictly weaker than the condition $X \subseteq X_n$ for each $n \in \mathbb{N}$ and $\text{Limsup}_n X_n \subseteq X$. In fact, consider the following slight modification of [3, Example 1.1]:

$$X = \{(0, 0)\} \quad \text{and} \quad X_n = \begin{cases} (\{1/n\} \times [0, 1]) \cup \{(0, 0)\}, & \text{if } n \text{ is even} \\ (\{1/n\} \times [-1, 0]) \cup \{(0, 0)\}, & \text{if } n \text{ is odd.} \end{cases}$$

We have $X \subseteq X_n$ for each $n \in \mathbb{N}$, $\text{Liminf}_n X_n = \{(0, 0)\}$ and $\text{Limsup}_n X_n = \{0\} \times [-1, 1]$. Hence *($\mathcal{L}_1$)* is satisfied, but $\text{Limsup}_n X_n \not\subseteq X$. Note that the equivalence can be shown if the sequence $(X_n)_n$ is assumed to be monotone.

Remark 3.1 (well-definedness of the limits of perturbed strategy and action profiles) Assumptions *($\mathcal{L}_1$)*, *($\mathcal{F}_1$)* guarantee that, when a sequence of strategy profiles of the perturbed games pointwise converges, then the limit is a well-defined strategy profile of the original game. In fact, let $(x_n, \varphi_n) \in X_n \times Y_n^{X_n}$ for each $n \in \mathbb{N}$ and suppose that $(x_n, \varphi_n)_n$ pointwise converges to (x^*, φ^*). Since $\text{Liminf}_n X_n \subseteq X$ then $x^* \in X$; moreover, since $X \subseteq X_n$ then $\varphi^*(x) = \lim_{n \to +\infty} \varphi_n(x)$ is well-defined for each $x \in X$; finally, since $\text{Limsup}_n Y_n \subseteq Y$ then $\varphi^*(x) \in Y$ for each $x \in X$. Hence, (x^*, φ^*) belongs to $X \times Y^X$ and the definition of pointwise convergence of $(x_n, \varphi_n)_n$ towards $(x^*, \varphi^*) \in X \times Y^X$ is well-posed, with $(x_n, \varphi_n) \in X_n \times Y_n^{X_n}$ for each $n \in \mathbb{N}$. Analogous considerations hold also for the sequence of action profiles of perturbed games. Indeed, if $(x_n, y_n) \in X_n \times Y_n$ for each $n \in \mathbb{N}$ and $(x_n, y_n)_n$ converges to (x^*, y^*), then by assumptions *($\mathcal{L}_1$)*, *($\mathcal{F}_1$)* we have that (x^*, y^*) belongs to $X \times Y$.

Now let us show a fundamental technical result that will be used in the proofs of the theorems.

Lemma 3.1. *Assume that ($\mathcal{L}_1$)-($\mathcal{L}_2$) and ($\mathcal{F}_1$)-($\mathcal{F}_2$) hold.*
Let $(\bar{x}_n, \bar{\varphi}_n)_n$ be an SPNE of Γ_n for each $n \in \mathbb{N}$ and fix $x \in X$ and $y \in Y$.

(i) If $\bar{y}(x)$ is a cluster point of the sequence $(\bar{\varphi}_n(x))_n$, then $f(x, \bar{y}(x)) \geq f(x, y)$.
(ii) If $(\bar{x}_n)_n$ converges to $\bar{x}$ and $\bar{y}$ is a cluster point of the sequence $(\bar{\varphi}_n(\bar{x}_n))_n$, then $f(\bar{x}, \bar{y}) \geq f(\bar{x}, y)$.

(iii) If $(\bar{x}_n)_n$ converges to $\bar{x}$ and there exists a subsequence of integers $(n_k)_k \subseteq \mathbb{N}$ such that $(\bar{\varphi}_{n_k}(\bar{x}_{n_k}))_k$ converges to $\bar{y}$ and that $(\bar{\varphi}_{n_k}(x))_k$ converges to $\bar{y}(x)$, then $l(\bar{x}, \bar{y}) \geq l(x, \bar{y}(x))$.

Proof. Note that the functions l_n and f_n are well-defined on $X \times Y$ by assumptions *($\mathcal{L}_1$)* and *($\mathcal{F}_1$)*.
Proof of (i). Let $(\bar{\varphi}_{n_k}(x))_k$ be a subsequence of $(\bar{\varphi}_n(x))_n$ converging to $\bar{y}(x)$. By *($\mathcal{F}_1$)*, then $\bar{y}(x) \in Y$. In light of assumption *($\mathcal{F}_2$)* and the definition of SPNE for Γ_{n_k}, we have that

$$f(x, \bar{y}(x)) = \lim_{k\to+\infty} f_{n_k}(x, \bar{\varphi}_{n_k}(x)) \geq \lim_{k\to+\infty} f_{n_k}(x, y) = f(x, y).$$

Proof of (ii). Let $(\bar{\varphi}_{n_k}(\bar{x}_{n_k}))_k$ be a subsequence of $(\bar{\varphi}_n(\bar{x}_n))_n$ converging to $\bar{y}$. By *($\mathcal{L}_1$)* and *($\mathcal{F}_1$)*, then $\bar{x} \in X$ and $\bar{y} \in Y$, respectively. Since $(\bar{x}_{n_k}, \bar{\varphi}_{n_k}(\bar{x}_{n_k}))_k$ converges to $(\bar{x}, \bar{y})$, assumption *($\mathcal{F}_2$)* and the definition of SPNE for Γ_{n_k} guarantee that

$$f(\bar{x}, \bar{y}) = \lim_{k\to+\infty} f_{n_k}(\bar{x}_{n_k}, \bar{\varphi}_{n_k}(\bar{x}_{n_k})) \geq \lim_{k\to+\infty} f_{n_k}(\bar{x}_{n_k}, y) = f(\bar{x}, y).$$

Proof of (iii). In light of *($\mathcal{L}_1$)* and *($\mathcal{F}_1$)*, then $\bar{x} \in X$, $\bar{y} \in Y$ and $\bar{y}(x) \in Y$. Since $(\bar{x}_{n_k}, \bar{\varphi}_{n_k}(\bar{x}_{n_k}))_k$ converges to $(\bar{x}, \bar{y})$, by assumption *($\mathcal{L}_2$)*, the definition of SPNE for Γ_{n_k} and the convergence of $(\bar{\varphi}_{n_k}(x))_k$ to $\bar{y}(x)$, we get

$$l(\bar{x}, \bar{y}) = \lim_{k\to+\infty} l_{n_k}(\bar{x}_{n_k}, \bar{\varphi}_{n_k}(\bar{x}_{n_k})) \geq \lim_{k\to+\infty} l_{n_k}(x, \bar{\varphi}_{n_k}(x)) = l(x, \bar{y}(x)).$$

□

We now investigate question **(Q)** by analyzing first the case of uniqueness of the follower's optimal reaction.

The follower's optimal reaction is unique: positive answer to (Q) (stability for both SPNE and SPNE outcomes)

In this case the best reply correspondence of the follower B is single valued. In the next proposition, the pointwise limit of the sequence $(\bar{x}_n, \bar{\varphi}_n)_n$ is shown to be an SPNE of Γ and the limit of the sequence $(\bar{x}_n, \bar{\varphi}_n(\bar{x}_n))_n$ is shown to be an SPNE outcome of Γ. The results heavily rely on the convergence properties of the sequence $(\bar{\varphi}_n(\bar{x}_n))_n$. Such properties have been firstly achieved in [27] when the follower's optimal reaction is unique and the perturbation involves only the payoff functions. There, the asymptotic behavior of Stackelberg equilibria under perturbations is investigated and the continuous convergence of $(\bar{\varphi}_n)_n$ is proved.

Theorem 3.1 (SPNE stability and SPNE outcomes stability)**.** *Assume that ($\mathcal{A}$) holds and that the follower's best reply correspondence B (defined in (1)) is single-valued.*
Let $(\bar{x}_n, \bar{\varphi}_n)$ be an SPNE of Γ_n for each $n \in \mathbb{N}$.

(i) If $(\bar{x}_n, \bar{\varphi}_n)_n$ pointwise converges to $(\bar{x}, \bar{\varphi})$, then $(\bar{x}, \bar{\varphi})$ is an SPNE of Γ.
(ii) If $(\bar{x}_n, \bar{\varphi}_n(\bar{x}_n))_n$ converges to $(\bar{x}, \bar{y})$, then $(\bar{x}, \bar{y})$ is an SPNE outcome of Γ.

Therefore, stability holds both for the SPNEs and for the SPNE outcomes.

Proof. Preliminarily, note that assumptions *($\mathcal{L}_1$)*, *($\mathcal{F}_1$)* ensure that $l_n(x,y)$, $f_n(x,y)$, $\bar{\varphi}_n(x)$ are well-defined for each $x \in X$, $y \in Y$, $n \in \mathbb{N}$ and that $(\bar{x}, \bar{\varphi})$ belongs to $X \times Y^X$ and $(\bar{x}, \bar{y})$ belongs to $X \times Y$ (see Remark 3.1).
Proof of (i). Let us prove that $(\bar{x}, \bar{\varphi})$ satisfies (SPE1) of Definition 2.1. Fix $x \in X$. Since $(\bar{\varphi}_n(x))_n$ converges to $\bar{\varphi}(x)$, by Lemma 3.1(i) we have that $f(x, \bar{\varphi}(x)) \geq f(\bar{x}, y)$ for each $y \in Y$. Thus $\bar{\varphi}(x) \in B(x)$. So (SPE1) holds and, by the single-valuedness of the follower's best reply correspondence, $B(x) = \{\bar{\varphi}(x)\}$ for each $x \in X$.
In order to show (SPE2) of Definition 2.1, we first prove that $\lim_{n\to+\infty} \bar{\varphi}_n(\bar{x}_n) = \bar{\varphi}(\bar{x})$. In light of assumption *($\mathcal{F}_3$)*, the sequence $(\bar{\varphi}_n(\bar{x}_n))_n \subseteq \Omega$ has a subsequence $(\bar{\varphi}_{n_k}(\bar{x}_{n_k}))_k$ converging to $\bar{y}$ and, by assumption *($\mathcal{F}_1$)*, $\bar{y} \in Y$. Hence, since $(\bar{x}_n)_n$ converges to $\bar{x}$, Lemma 3.1(ii) guarantees that $f(\bar{x}, \bar{y}) \geq f(\bar{x}, y)$ for each $y \in Y$. So $\bar{y} \in B(\bar{x})$ and, since $B(x) = \{\bar{\varphi}(x)\}$ for each $x \in X$, then $\bar{y} = \bar{\varphi}(\bar{x})$. The result above holds for each convergent subsequence of $(\bar{\varphi}_n(\bar{x}_n))_n$, hence the whole sequence $(\bar{\varphi}_n(\bar{x}_n))_n$ converges to $\bar{\varphi}(\bar{x})$.
Now, since $(\bar{x}_n, \bar{\varphi}_n(\bar{x}_n))_n$ converges to $(\bar{x}, \bar{\varphi}(\bar{x}))$ and $(\bar{\varphi}_n)_n$ pointwise converges to $\bar{\varphi}$, by Lemma 3.1(iii) it follows that $l(\bar{x}, \bar{\varphi}(\bar{x})) \geq l(x, \bar{\varphi}(x))$ for each $x \in X$. Hence also condition (SPE2) holds for $(\bar{x}, \bar{\varphi})$.
Proof of (ii). Since the follower's best reply correspondence is single valued, let $\varphi^*\colon X \to Y$ such that $B(x) = \{\varphi^*(x)\}$ for each $x \in X$. Let us prove that $(\bar{x}, \bar{y})$ is the outcome of the strategy profile $(\bar{x}, \varphi^*)$ and that $(\bar{x}, \varphi^*)$ is an SPNE of Γ.
Since $(\bar{x}_n, \bar{\varphi}_n(\bar{x}_n))_n$ converges to $(\bar{x}, \bar{y})$, then $f(\bar{x}, \bar{y}) \geq f(\bar{x}, y)$ for each $y \in Y$ by Lemma 3.1(ii). So $\bar{y} \in B(\bar{x}) = \{\varphi^*(\bar{x})\}$; then $\bar{y} = \varphi^*(\bar{x})$ and $(\bar{x}, \bar{y})$ is the outcome of the strategy profile $(\bar{x}, \varphi^*)$.
By definition of φ^*, the strategy profile $(\bar{x}, \varphi^*)$ satisfies (SPE1) of Definition 2.1. Let us show that (SPE2) holds.
Let $x \in X$. In light of assumption *($\mathcal{F}_3$)*, the sequence $(\bar{\varphi}_n(x))_n$ is contained in the compact set Ω, so there is a sequence of integers $(n_{k(x)})_{k(x)} \subseteq \mathbb{N}$ such that the subsequence $(\bar{\varphi}_{n_{k(x)}}(x))_{k(x)}$ is convergent. Denoting with $\bar{y}_x$ the limit of such a subsequence, by Lemma 3.1(i) we get $\bar{y}_x \in B(x)$ and so $y_x = \varphi^*(x)$. Since this holds for each convergent subsequence of $(\bar{\varphi}_n(x))_n$ and for each $x \in X$, the sequence $(\bar{\varphi}_n)_n$ pointwise converges to φ^*. Therefore, in light of Lemma 3.1(iii), $l(\bar{x}, \bar{y}) \geq l(x, \varphi^*(x))$ for each $x \in X$ and so $(\bar{x}, \varphi^*)$ satisfies (SPE2). □

The follower's optimal reaction is not always unique: positive answer only for stability of SPNE outomes

In this case the map B is not single valued. Unfortunately, the pointwise limit of the sequence $(\bar{x}_n, \bar{\varphi}_n)_n$ is not ensured to be an SPNE of Γ, even when the perturbation affects only the payoff function of just one player (that is, the action sets and the payoff function of the other player are not perturbed) and both the original game and the perturbed games have a unique SPNE. This is shown in the following illustrative example.

Illustrative example. In this example we use a perturbation, namely the perturbation via the so-called Tikhonov regularization, based on a classical regularization technique in optimization that we recall below.
Let $g\colon U \to \mathbb{R}$ be a real-valued function defined on a subset U of an Euclidean space $\mathbb{U}$. The

sequence of Tikhonov-perturbed functions $(g^{\mathcal{T}}_{\lambda_n})_n$ related to g is defined by

$$g^{\mathcal{T}}_{\lambda_n}(u) = g(u) - \lambda_n \|u\|^2_{\mathbb{U}},$$

where $\lambda_n > 0$ for each $n \in \mathbb{N}$. The quadratic function appearing in $g^{\mathcal{T}}_{\lambda_n}$ represents a penalty term that increases as the size of the variable increases; for example, in pollution models it can be interpreted as the damage caused by emissions or a discomfort occurring from consumption. A key-result on the relation between the maximizers of $g^{\mathcal{T}}_{\lambda_n}$ and of g is mentioned below (see [36] and also, for example, [13, Chapter 1]).

If the set U is compact and convex, the function g is upper semicontinuous and concave over U and $\lim_{n\to+\infty} \lambda_n = 0$, then

- $g^{\mathcal{T}}_{\lambda_n}$ *has a unique maximizer $\bar{u}^{\mathcal{T}}_n \in U$ for each $n \in \mathbb{N}$;*
- *the sequence $(\bar{u}^{\mathcal{T}}_n)_n$ converges to the minimum norm element in the set of maximizer of g.*

In Stackelberg games framework, the sequence of Tikhonov-perturbed Stackelberg games has been introduced in [28] as follows:

$$\Gamma^{\mathcal{T}}_{\beta_n} := \langle X, Y, l, f^{\mathcal{T}}_{\beta_n} \rangle,$$

where $\beta_n > 0$, $\lim_{n\to+\infty} \beta_n = 0$ and

$$f^{\mathcal{T}}_{\beta_n}(x, y) := f(x, y) - \beta_n \|y\|^2_{\mathbb{Y}}.$$

Note that the function $f^{\mathcal{T}}_{\beta_n}$ is obtained by perturbing f via the Tikhonov regularization with respect to the follower's variable only. Under the concavity of $f(x, \cdot)$, such a perturbation allows to bypass the non-single-valuedness of the follower's best reply correspondence (since the optimal reaction of the follower becomes always unique in $\Gamma^{\mathcal{T}}_{\beta_n}$) and the limit of the sequence of Stackelberg equilibria related to $\Gamma^{\mathcal{T}}_{\beta_n}$ has been shown to be a so-called *lower Stackelberg equilibrium pair*, see [28, Proposition 3.6]. Moreover, starting from the sequence of Stackelberg equilibria related to $\Gamma^{\mathcal{T}}_{\beta_n}$, in [30] a constructive method to approach an SPNE of Γ has been defined. Finally, we mention that a numerical investigation on the Stackelberg solutions of $\Gamma^{\mathcal{T}}_{\beta_n}$ as n goes to infinity has been provided in [12].

Now let $\Gamma = \langle X, Y, l, f \rangle$ be the game where $X = Y = [-2, 2]$, $l(x, y) = x + y$ and $f(x, y) = -xy$. The best reply correspondence B is not single-valued and it is defined on $[-2, 2]$ by

$$B(x) = \begin{cases} \{2\}, & \text{if } x \in [-2, 0[\\ [-2, 2], & \text{if } x = 0 \\ \{-2\}, & \text{if } x \in]0, 2]. \end{cases}$$

Consider the perturbed Stackelberg games $\Gamma_n = \langle X_n, Y_n, l_n, f_n \rangle$ with $X_n = X$, $Y_n = Y$, $l_n \equiv l$ for each $n \in \mathbb{N}$ and $f_n(x, y) = f(x, y) - y^2/n$, that is $(\Gamma_n)_n$ is the sequence of Tikhonov-perturbed Stackelberg games $(\Gamma^{\mathcal{T}}_{\beta_n})_n$ with $\beta_n = 1/n$ for each $n \in \mathbb{N}$. The follower's best reply correspondence in each perturbed game is single-valued and $\Gamma^{\mathcal{T}}_{\beta_n}$ has a unique SPNE, namely the strategy profile $(\bar{x}^{\mathcal{T}}_n, \bar{\varphi}^{\mathcal{T}}_n)$ defined by

$$\bar{x}^{\mathcal{T}}_n = -4/n \quad \text{and} \quad \bar{\varphi}^{\mathcal{T}}_n(x) = \begin{cases} 2, & \text{if } x \in [-2, -4/n[\\ -nx/2, & \text{if } x \in [-4/n, 4/n] \\ -2, & \text{if } x \in]4/n, 2] \end{cases}$$

for each $n \in \mathbb{N}$. The sequence $(\bar{x}_n^{\mathcal{T}}, \bar{\varphi}_n^{\mathcal{T}})_n$ pointwise converges to $(\bar{x}^{\mathcal{T}}, \bar{\varphi}^{\mathcal{T}})$, where

$$\bar{x}^{\mathcal{T}} = 0 \quad \text{and} \quad \bar{\varphi}^{\mathcal{T}}(x) = \lim_{n\to+\infty} \bar{\varphi}_n^{\mathcal{T}}(x) = \begin{cases} 2, & \text{if } x \in [-2,0[\\ 0, & \text{if } x = 0 \\ -2, & \text{if } x \in]0,2]. \end{cases} \tag{2}$$

The strategy profile $(\bar{x}^{\mathcal{T}}, \bar{\varphi}^{\mathcal{T}})$ satisfies the condition (SPE1) in Definition 2.1, as $\bar{\varphi}^{\mathcal{T}}(x) \in B(x)$ for each $x \in [-2,2]$, but it does not satisfy the condition (SPE2) in Definition 2.1, because

$$l(x, \bar{\varphi}^{\mathcal{T}}(x)) = \begin{cases} x+2, & \text{if } x \in [-2,0[\\ 0, & \text{if } x = 0 \\ x-2, & \text{if } x \in]0,2], \end{cases} \quad \text{and} \quad \arg\max_{x\in X} l(x, \bar{\varphi}^{\mathcal{T}}(x)) = \emptyset.$$

Therefore the pointwise limit $(\bar{x}^{\mathcal{T}}, \bar{\varphi}^{\mathcal{T}})$ is not an SPNE of Γ.
Note that the original game Γ has a unique SPNE. In fact, let φ_α with $\alpha \in [-2,2]$ be a selection of the set-valued map B, that is

$$\varphi_\alpha(x) = \begin{cases} 2, & \text{if } x \in [-2,0[\\ \alpha, & \text{if } x = 0 \\ -2, & \text{if } x \in]0,2]. \end{cases}$$

Since the function $l(\cdot, \varphi_\alpha(\cdot))$ has a maximum only if $\alpha = 2$, the strategy profile $(\bar{x}, \bar{\varphi})$ with

$$\bar{x} = 0 \quad \text{and} \quad \bar{\varphi}(x) = \begin{cases} 2, & \text{if } x \in [-2,0] \\ -2, & \text{if } x \in]0,2] \end{cases} \tag{3}$$

is the unique SPNE of Γ.
Moreover, we point out that the sequence of action profiles $(\bar{x}_n^{\mathcal{T}}, \bar{\varphi}_n^{\mathcal{T}}(\bar{x}_n^{\mathcal{T}}))_n$ converges to $(\bar{x}, \bar{\varphi}(\bar{x}))$, i.e. the limit of the sequence of the SPNEs outcomes related to $(\Gamma_{\beta_n}^{\mathcal{T}})_n$ is the SPNE outcome of Γ.

The bad behavior illustrated above is not an isolated incident, but it appears also in other examples (e.g., the games of [28, Example 2.1] and [30, examples 3.2 to 3.4]).
Hence stability is not guaranteed for the SPNE in Stackelberg games under perturbations when the follower's best reply correspondence is not single valued, differently from the case where B is single valued.

Remark 3.2 (related literature) This lack of stability concerning the SPNEs was already noted in literature. In fact, in [20] an example where the pointwise limit of the SPNEs of perturbed games is not an SPNE of the original game is showed when the perturbation involves only the action sets (more precisely, perturbations made of finite-action games are considered). Our illustrative example proves that stability is lacking even when we keep the action sets unperturbed and we perturb only the payoff functions.

At the end of the *Illustrative example*, we displayed that the limit of the sequence of SPNE outcomes of the Tikhonov-perturbed Stackelberg games $(\Gamma^{\mathcal{T}}_{\beta_n})_n$ is an SPNE outcome of the original game Γ. The next result proves that there is actually stability for the SPNEs outcomes under general assumptions, although stability is not ensured for the SPNEs.

Theorem 3.2 (SPNE outcomes stability)**.** *Assume that $(\mathcal{A})$ holds.*
Let $(\bar{x}_n, \bar{\varphi}_n)$ be an SPNE of Γ_n for each $n \in \mathbb{N}$.
If the sequence of outcomes $(\bar{x}_n, \bar{\varphi}_n(\bar{x}_n))_n$ converges to $(\bar{x}, \bar{y})$, then $(\bar{x}, \bar{y})$ is an SPNE outcome of Γ. Therefore, "stability of SPNE outcomes" holds.

Proof. As in the proof of Theorem 3.1, we have that $l_n(x, y)$, $f_n(x, y)$, $\bar{\varphi}_n(x)$ are well-defined for each $x \in X$, $y \in Y$, $n \in \mathbb{N}$ and that $(\bar{x}, \bar{y})$ belongs to $X \times Y$.
Let $x \in X$. In light of assumption $(\mathcal{F}_3)$, the sequence $(\bar{\varphi}_n(x))_n$ is contained in the compact set Ω, so there is a sequence of integers $(n_{k(x)})_{k(x)} \subseteq \mathbb{N}$ such that the subsequence $(\bar{\varphi}_{n_{k(x)}}(x))_{k(x)}$ is convergent. Denoting by $\bar{y}_x$ the limit of such a subsequence, then assumption $(\mathcal{F}_1)$ guarantees that $\bar{y}_x \in Y$. Hence, the function $\widetilde{\varphi}$ defined by

$$\widetilde{\varphi}(x) := \begin{cases} \bar{y}_x, & \text{if } x \neq \bar{x} \\ \bar{y}, & \text{if } x = \bar{x} \end{cases}$$

is well-defined and belongs to Y^X. Clearly, the action profile $(\bar{x}, \bar{y})$ is an outcome of the strategy profile $(\bar{x}, \widetilde{\varphi})$. Let us show that $(\bar{x}, \widetilde{\varphi})$ is an SPNE of Γ.
We start by proving (SPE1) of Definition 2.1, that is $\widetilde{\varphi}(x) \in B(x)$ for each $x \in X$.
Let $x \in X \setminus \{\bar{x}\}$ and $(n_{k(x)})_{k(x)}$ be the sequence of integers such that $(\bar{\varphi}_{n_{k(x)}}(x))_{k(x)}$ converges to $\bar{y}_x$. By definition of $\widetilde{\varphi}$ and Lemma 3.1(i) we get $f(x, \widetilde{\varphi}(x)) = f(x, \bar{y}_x) \geq f(x, y)$ for each $y \in Y$. So $\widetilde{\varphi}(x) \in B(x)$ for each $x \in X \setminus \{\bar{x}\}$.
Consider $x = \bar{x}$. Since $(\bar{x}_n, \bar{\varphi}_n(\bar{x}_n))_n$ converges to $(\bar{x}, \bar{y})$, in light of the definition of $\widetilde{\varphi}$ and Lemma 3.1(ii) we have $f(\bar{x}, \widetilde{\varphi}(\bar{x})) = f(\bar{x}, \bar{y}) \geq f(\bar{x}, y)$ for each $y \in Y$. Hence $\widetilde{\varphi}(\bar{x}) \in B(\bar{x})$ and the strategy profile $(\bar{x}, \widetilde{\varphi})$ satisfies the condition (SPE1) of Definition 2.1.
Now, let us show (SPE2) of Definition 2.1, that is $\bar{x} \in \arg\max_{x \in X} l(x, \widetilde{\varphi}(x))$. Since $\widetilde{\varphi}(\bar{x}) = \bar{y}$, it is sufficient to prove that $l(\bar{x}, \bar{y}) \geq l(x, \bar{y}_x)$ for each $x \in X \setminus \{\bar{x}\}$.
Fix $x \in X \setminus \{\bar{x}\}$ and let $(n_{k(x)})_{k(x)}$ be the sequence of integers such that $(\bar{\varphi}_{n_{k(x)}}(x))_{k(x)}$ converges to $\bar{y}_x$. Moreover, $(\bar{\varphi}_{n_{k(x)}}(\bar{x}_{n_{k(x)}}))_{k(x)}$ converges to $\bar{y}$ and, thus, Lemma 3.1(iii) guarantees that $l(\bar{x}, \bar{y}) \geq l(x, \bar{y}_x)$. Therefore, $(\bar{x}, \widetilde{\varphi})$ satisfies condition (SPE2) of Definition 2.1. □

Remark 3.3 (literature on the convergence of the SPNE outcomes) The nice behavior of SPNE outcomes obtained via Theorem 3.2 has connections with previous results in literature. In N-player N-stage continuous game with perfect information where one player acts in each stage, [20, Theorem 1] showed the upper semicontinuity of the SPNE paths when perturbations affect only the action sets. Since their perturbation consists of a discretization of the sets (from inside) and this instance does not fit the assumptions $(\mathcal{L}_1)$-$(\mathcal{F}_1)$, our achievement represents a complementary result in the case of $N = 2$. Moreover, note that also [30, Theorem 3.1] and [8, Theorem 1] lead to the stability of SPNE outcomes for perturbations of the payoff functions via the Tikhonov regularization and via the proximal regularization, respectively. So we extend such results to general perturbations both on the action sets and on the payoff functions.

Hence, question **(Q)** is positively answered with respect to the SPNE outcomes (and non-positively answered with respect to the SPNE).
However, we mention that for a particular class of games and perturbations the limit of the sequence $(\bar{x}_n, \bar{\varphi}_n)_n$ can be shown to be an SPNE of Γ.

Proposition 3.1. *Assume that* $(\mathcal{A})$ *holds and that*

(a) $f(x,y) = g(x) + h(y)$ *for each* $(x,y) \in X \times Y$*, where* $g\colon X \to \mathbb{R}$ *and* $h\colon Y \to \mathbb{R}$*;*
(b) $f_n(x,y) = g_n(x) + h_n(y)$ *for each* $(x,y) \in X_n \times Y_n$ *and* $n \in \mathbb{N}$*, where* $g_n\colon X_n \to \mathbb{R}$ *and* $h_n\colon Y_n \to \mathbb{R}$*.*

Let $(\bar{x}_n, \bar{\varphi}_n)$ *be an SPNE of* Γ_n *for each* $n \in \mathbb{N}$*.*
If $(\bar{x}_n, \bar{\varphi}_n)_n$ *pointwise converges to* $(\bar{x}, \bar{\varphi})$*, then* $(\bar{x}, \bar{\varphi})$ *is an SPNE of* Γ*.*

Proof. All the functions and pointwise limits involved are well-defined (Remark 3.1).
Assumption *(b)* implies that the follower's best reply correspondence in Γ_n does not depend on the leader's variable. So, recalling that (by hypothesis) $\bar{\varphi}_n$ is the follower's strategy in an SPNE of Γ_n and $(\bar{\varphi}_n)_n$ pointwise converges to $\bar{\varphi}$, then there exists a sequence $(\bar{y}_n)_n \subseteq \mathbb{Y}$ converging to $\bar{y} \in Y$ such that $\bar{\varphi}_n(x) = \bar{y}_n$ for each $x \in X_n$ and $\bar{\varphi}(x) = \bar{y}$ for each $x \in X$.
Given the above, by using assumptions *(a)*-*(b)*, $(\mathcal{L}_2)$-$(\mathcal{F}_2)$ and the definition of SPNE for Γ_n, one can prove that conditions (SPE1) and (SPE2) of Definition 2.1 are satisfied. □

4 Stability but for a point

The previous section showed that, unless the follower's best reply correspondence is single valued (Theorem 3.1) or a special class of games and perturbations is involved (Proposition 3.1), the question **(Q)** of Section 2 cannot be fully positively addressed under the general assumption $(\mathcal{A})$ displayed in Section 3. However, by looking at the *Illustrative example*, the pointwise limit of $(\bar{x}_n^{\mathcal{T}}, \bar{\varphi}_n^{\mathcal{T}})_n$ found in (2) and the unique SPNE of the game Γ in (3) are very "close" to each other: they differ just in one value of the follower's strategy. Hence, it is natural to ask:

(Q') From the pointwise limit of a sequence $(\bar{x}_n, \bar{\varphi}_n)_n$ of SPNEs related to $(\Gamma_n)_n$ is it possible to achieve an SPNE of the original game Γ (even when such a limit is not an SPNE of Γ)? And how?

The first positive answer to question **(Q')** has been shown in [30] in the case of the perturbation via the Tikhonov regularization. Then, an analogous result has been proved in [8] when both players' payoff functions are perturbed via the so-called proximal regularization. The next key-result does not involve a specific perturbation and provides sufficient conditions on the game Γ and on the perturbed games $(\Gamma_n)_n$ that allow to answer to question **(Q')** under general assumptions.

Theorem 4.1 (SPNEs stability but for a point)**.** *Assume that* $(\mathcal{A})$ *holds.*
Let $(\bar{x}_n, \bar{\varphi}_n)$ *be an SPNE of* Γ_n *for each* $n \in \mathbb{N}$.
If $(\bar{x}_n, \bar{\varphi}_n)_n$ *pointwise converges to* $(\bar{x}, \bar{\varphi})$, *then there exists* $\bar{y} \in Y$ *such that the strategy profile* $(\bar{x}, \widehat{\varphi})$ *where*

$$\widehat{\varphi}(x) := \begin{cases} \bar{\varphi}(x), & \text{if } x \neq \bar{x} \\ \bar{y}, & \text{if } x = \bar{x} \end{cases}$$

is an SPNE of Γ.
Therefore, "stability but for a point" holds for the SPNEs.

Proof. As in the proof of Theorem 3.1, we have that $l_n(x, y)$, $f_n(x, y)$, $\bar{\varphi}_n(x)$ are well-defined for each $x \in X$, $y \in Y$, $n \in \mathbb{N}$ and that $(\bar{x}, \bar{\varphi})$ belongs to $X \times Y^X$.
In light of assumption $(\mathcal{F}_3)$, the sequence $(\bar{\varphi}_n(\bar{x}_n))_n$ is contained in the compact set Ω, so it has a subsequence $(\bar{\varphi}_{n_k}(\bar{x}_{n_k}))_k$ converging to $\bar{y}$. By assumption $(\mathcal{F}_1)$, then $\bar{y} \in Y$ and the strategy profile $(\bar{x}, \widehat{\varphi})$ belongs to $X \times Y^X$.
Now, let us show that $(\bar{x}, \widehat{\varphi})$ is an SPNE of Γ.
We start by proving (SPE1) of Definition 2.1, that is $\widehat{\varphi}(x) \in B(x)$ for each $x \in X$.
Let $x \in X \setminus \{\bar{x}\}$. Since $(\bar{\varphi}_n(x))_n$ converges to $\bar{\varphi}(x)$, Lemma 3.1(i) and the definition of $\widehat{\varphi}$ implies that $f(x, \widehat{\varphi}(x)) = f(x, \bar{\varphi}(x)) \geq f(x, y)$ for each $y \in Y$. Hence, $\widehat{\varphi}(x) \in B(x)$ for each $x \in X \setminus \{\bar{x}\}$.
Let $x = \bar{x}$. As $(\bar{\varphi}_{n_k}(\bar{x}_{n_k}))_k$ converges to $\bar{y}$, by Lemma 3.1(ii) we have $f(\bar{x}, \widehat{\varphi}(\bar{x})) = f(\bar{x}, \bar{y}) \geq f(\bar{x}, y)$ for each $y \in Y$. So $\widehat{\varphi}(\bar{x}) \in B(\bar{x})$ and $(\bar{x}, \widehat{\varphi})$ satisfies the condition (SPE1) of Definition 2.1.
Now, let us show (SPE2) of Definition 2.1, that is $\bar{x} \in \arg\max_{x \in X} l(x, \widehat{\varphi}(x))$. Since $\widehat{\varphi}(\bar{x}) = \bar{y}$, it is sufficient to prove that $l(\bar{x}, \bar{y}) \geq l(x, \bar{\varphi}(x))$ for each $x \in X \setminus \{\bar{x}\}$.
Fix $x \in X \setminus \{\bar{x}\}$. In light of the convergence of $(\bar{\varphi}_{n_k}(x))_k$ and of $(\bar{\varphi}_{n_k}(\bar{x}_{n_k}))_k$ to $\bar{\varphi}(x)$ and $\bar{y}$, respectively, Lemma 3.1(iii) ensures that $l(\bar{x}, \bar{y}) \geq l(x, \bar{\varphi}(x))$. Therefore, $(\bar{x}, \widehat{\varphi})$ satisfies condition (SPE2) of Definition 2.1. □

Remark 4.1 (weakening of the assumptions) The assumptions on the payoff functions in Theorem 4.1 can be weaken. More precisely, $(\mathcal{L}_2)$ and $(\mathcal{F}_2)$ can be replaced with the weaker conditions used in [27, Propositions 3.1 and 3.2]; see also [27, Remark 2.3] for further discussion on the connections between such conditions and existing notions of convergence. The same holds also for Theorems 3.1 and 3.2. We preferred to make stronger assumptions only for reasons of readability.

Therefore, even when the pointwise limit of $(\bar{x}_n, \bar{\varphi}_n)_n$ is not an SPNE of Γ, the asymptotic behavior of the sequence $(\bar{x}_n, \bar{\varphi}_n)_n$ allows to recover an SPNE of Γ. In fact, according to Theorem 4.1, an SPNE of Γ can be defined from the sequence $(\bar{x}_n, \bar{\varphi}_n)_n$ by possibly modifying at just one point the pointwise limit $(\bar{x}, \bar{\varphi})$. Moreover, the proof of Theorem 4.1 tells also how an SPNE can be defined, that is how the point in the limit should be modified. In fact, taking into account each cluster point of the sequence $(\bar{\varphi}_n(\bar{x}_n))_n$, one can obtain different SPNEs by replacing the value of the follower's strategy $\bar{\varphi}(\bar{x})$ with the limit of a convergent subsequence of $(\bar{\varphi}_n(\bar{x}_n))_n$. So question **(Q')** is addressed.

Remark 4.2 (on the continuous convergence of $(\bar{\varphi}_n)_n$) We point out that, if the sequence $(\bar{\varphi}_n)_n$ is continuously convergent to $\bar{\varphi}$, then in Theorem 4.1 we have $\bar{y} = \bar{\varphi}(\bar{x})$ and so $(\bar{x}, \bar{\varphi})$ is an SPNE of Γ.

However, even when Γ is a "nice" game and $(\Gamma_n)_n$ is a "nice" perturbation (meaning that the action sets are compact and convex, the payoff functions are smooth and $(l_n)_n$ and $(f_n)_n$ continuously converge to l and f), the sequence $(\bar{\varphi}_n)_n$ is not guaranteed to be continuously convergent. This has been shown in [28, Remark 3.1] for perturbations via the Tikhonov regularization. In fact, consider in *Illustrative example* the sequence $(\bar{x}_n^{\mathcal{T}})_n = (-4/n)_n$, which converges to $\bar{x}^{\mathcal{T}} = 0$. Then, we have

$$\lim_{n\to+\infty} \bar{\varphi}_n^{\mathcal{T}}(\bar{x}_n^{\mathcal{T}}) = 1 \neq 0 = \lim_{n\to+\infty} \bar{\varphi}_n^{\mathcal{T}}(\bar{x}^{\mathcal{T}}).$$

Hence, the sequence $(\bar{\varphi}_n^{\mathcal{T}})_n$ does not continuously converge (even though the action sets and the leader's payoff function are not perturbed and the sequence $(f_{\beta_n}^{\mathcal{T}})_n$ continuously converge to f).
Obviously, the continuous convergence of $(\bar{\varphi}_n)_n$ to $\bar{\varphi}$ is guaranteed in the trivial situation where $\bar{\varphi}_n$ and $\bar{\varphi}$ are constant functions, which happens for example when f_n and f are additively separable (see Proposition 3.1). For further discussion on achievements involving the continuous convergence of $(\bar{\varphi}_n)_n$, see [14, Section 5].

5 Conclusions

In this paper we investigated the asymptotic behavior of the SPNE in Stackelberg games when perturbations affect both the action sets and the payoff functions. Our analysis focused on the following questions:
Is the limit of a sequence of SPNEs of perturbed games an SPNE of the original game? If not, how far is such a limit from an SPNE? And what about the limit of the sequence of SPNE outcomes?
In order to address these questions, we took into account a set of assumptions on the perturbed games which easily accommodate broadly used kinds of perturbations of sets and of functions (assumption *$(\mathcal{A})$* in Section 3). Under such assumptions we showed that

- when the follower's best reply correspondence is single valued, the limit of a sequence of SPNEs of perturbed Stackelberg games is ensured to be an SPNE of the original game and the limit of a sequence of SPNE outcomes of perturbed Stackelberg games is ensured to be an SPNE outcome of the original game (Theorem 3.1);
- when the follower's best reply correspondence is not single valued,
  - ◇ stability result for SPNEs does not hold, even when the action sets are unperturbed, only one payoff function is perturbed and both the original and the perturbed games have a unique SPNE (*Illustrative example*);
  - ◇ stability result for SPNEs outcomes holds (Theorem 3.2);
  - ◇ stability result for SPNEs holds if a particular class of games and perturbations is involved (Proposition 3.1);
- if the limit of a sequence of SPNEs of perturbed games is not an SPNE of the original game, an SPNE can be achieved by modifying the value of the follower's strategy at just one point (Theorem 4.1).

In summary, we can state that there is *stability but for a point* for the SPNEs, there is stability for SPNEs under the single-valuedness of the follower's best reply correspondence or under a particular class of games and perturbations and there is stability for the SPNE outcomes. Note that the assumptions on the payoff functions $(\mathcal{L}_2)$-$(\mathcal{F}_2)$ can be replaced with weaker convergence requirements (as noted in Remark 4.1) and, since our analysis is carried out in continuous setting, all the results holds also for the mixed extension of Stackelberg games with finite action sets.
These achievements leave a first significant picture on the topic of asymptotic behavior of SPNEs in Stackelberg games under perturbation and so they contribute to fill the gap with respect to the richer literature on the asymptotic behavior of Nash equilibria in simultaneous-move games.

For a further step on this subject, some extensions could be taken into account for future research.
A first extension of our investigation could be to consider the more general case where the set of actions available for each player depends on the choice of the other player. This includes the case of constraints defined by parametric inequalities, parametrized by the actions of the other player. Furthermore, it is worth considering the extension to an infinite-dimensional framework which would allow to deal with differential games (see, e.g., [19]) and where a classical example of perturbation is a discretization associated to the Galerkin method.
Finally, another direction of research could be to consider more general classes of games, as the N-player N-stage perfect-information games analyzed in [20] or the multistage infinite-horizon games examined in [6].

References

[1] H. Attouch. *Variational convergence for functions and operators.* London: Pitman, 1984.

[2] H. Attouch and A. Soubeyran. "Inertia and reactivity in decision making as cognitive variational inequalities". In: *J. Convex Anal.* 13 (2006), pp. 207–224.

[3] J.-P. Aubin and H. Frankowska. *Set-valued analysis.* Boston, MA: Birkhäuser, 2009.

[4] B. Bank, J. Guddat, D. Klatte, B. Kummer, and K. Tammer. *Non-linear parametric optimization.* Basel: Springer, 1983.

[5] T. Börgers. "Perfect equilibrium histories of finite and infinite horizon games". In: *J. Econ. Theory* 47 (1989), pp. 218–227. DOI: 10.1016/0022-0531(89)90112-9.

[6] T. Börgers. "Upper hemicontinuity of the correspondence of subgame-perfect equilibrium outcomes". In: *J. Math. Econ.* 20 (1991), pp. 89–106. DOI: 10.1016/0304-4068(91)90019-P.

[7] G. Carmona. "On games of perfect information: equilibria, ϵ-equilibria and approximation by simple games". In: *Int. Game Theory Rev.* 07 (2005), pp. 491–499. DOI: 10.1142/S0219198905000661.

[8] F. Caruso, M.C. Ceparano, and J. Morgan. "Subgame perfect Nash equilibrium: a learning approach via costs to move". In: *Dyn. Games Appl.* 9 (2019), pp. 416–432. DOI: 10.1007/s13235-018-0277-3.

[9] F. Caruso, M. B. Lignola, and J. Morgan. "Regularization and approximation methods in Stackelberg games and bilevel optimization". In: *Bilevel Optimization: Advances and Next Challenges.* Ed. by S. Dempe and A. Zemkoho. Cham: Springer International Publishing, 2020, pp. 77–138. DOI: 10.1007/978-3-030-52119-6_4.

[10] E. Cavazzuti and S. D. Flåm. "Entropic penalties in finite games". In: *Ann. Oper. Res.* 137 (2005), pp. 331–348. DOI: 10.1007/s10479-005-2264-5.

[11] G. Dal Maso. *An introduction to Γ-convergence.* Vol. 8. Springer, 1993.

[12] S. Dempe. "A bundle algorithm applied to bilevel programming problems with non-unique lower level solutions". In: *Comput. Optim. Appl.* 15 (2000), pp. 145–166. DOI: 10.1023/A:1008735010803.

[13] A. L. Dontchev and T. Zolezzi. *Well-Posed Optimization Problems.* Springer, Berlin, Heidelberg, 1993.

[14] S.D. Flåm. "On games and cost of change". In: *Ann. Oper. Res.* 301 (2021), pp. 107–119. DOI: 10.1007/s10479-020-03585-w.

[15] D. Fudenberg and D. Levine. "Subgame-perfect equilibria of finite- and infinite-horizon games". In: *J. Econ. Theory* 31 (1983), pp. 251–268. DOI: 10.1016/0022-0531(83)90076-5.

[16] C. Harris. "A characterisation of the perfect equilibria of infinite horizon games". In: *J. Econ. Theory* 37 (1985), pp. 99–125. DOI: 10.1016/0022-0531(85)90032-8.

[17] C. Harris. "Existence and characterization of perfect equilibrium in games of perfect information". In: *Econometrica* 53 (1985), pp. 613–628.

[18] J. C. Harsanyi and R. Selten. *A general theory of equilibrium selection in games.* Cambridge (MA): The MIT Press, 1988.

[19] A. Haurie, J.B. Krawczyk, and G. Zaccour. *Games and Dynamic Games.* World Scientific, 2012.

[20] M. Hellwig, W. Leininger, P. J. Reny, and A. J. Robson. "Subgame perfect equilibrium in continuous games of perfect information: An elementary approach to existence and approximation by discrete games". In: *J. Econ. Theory* 52 (1990), pp. 406–422. DOI: 10.1016/0022-0531(90)90039-M.

[21] E. Kohlberg and J.-F. Mertens. "On the strategic stability of equilibria". In: *Econometrica* 54 (1986), pp. 1003–1037. DOI: 10.2307/1912320.

[22] K. Kuratowski. *Topology.* Vol. 1. New York: Academic Press, 1966.

[23] M.B. Lignola and J. Morgan. "Further on Inner Regularizations in Bilevel Optimization". In: *J. Optim. Theory Appl.* 180 (2019), pp. 1087–1097. DOI: 10.1007/s10957-018-1438-7.

[24] M.B. Lignola and J. Morgan. "Inner regularizations and viscosity solutions for pessimistic bilevel optimization problems". In: *J. Optim. Theory Appl.* 173 (2017), pp. 183–202. DOI: 10.1007/s10957-017-1085-4.

[25] M.B. Lignola and J. Morgan. "Stability of Regularized Bilevel Programming Problems". In: *J. Optim. Theory Appl.* 93 (1997), pp. 575–596. DOI: 10.1023/A:1022695113803.

[26] J. L. Lions. *Optimal Control of Systems Governed by Partial Differential Equations.* Heidelberg: Springer-Verlag, 1971.

[27] P. Loridan and J. Morgan. "A theoretical approximation scheme for Stackelberg problems". In: *J. Optim. Theory Appl.* 61 (1989), pp. 95–110. DOI: 10.1007/BF00940846.

[28] P. Loridan and J. Morgan. "Regularizations for two-level optimization problems". In: *Advances in Optimization.* Ed. by W. Oettli and D. Pallaschke. Berlin, Heidelberg: Springer Berlin Heidelberg, 1992, pp. 239–255.

[29] L. Mallozzi and P. M. Pardalos. "Entropic regularization in hierarchical games". In: *Operations Research Forum* 3 (2022). DOI: 10.1007/s43069-022-00130-2.

[30] J. Morgan and F. Patrone. “Stackelberg problems: subgame perfect equilibria via Tikhonov regularization”. In: *Advances in Dynamic Games: Applications to Economics, Management Science, Engineering, and Environmental Management.* Ed. by A. Haurie, S. Muto, L.A. Petrosyan, and T.E.S. Raghavan. Boston, MA: Birkhäuser Boston, 2006, pp. 209–221. DOI: 10.1007/0-8176-4501-2_12.

[31] R. B. Myerson. “Refinements of the Nash equilibrium concept”. In: *Int. J. Game Theory* 7 (1978), pp. 73–80. DOI: 10.1007/BF01753236.

[32] A. Quarteroni and A. Valli. *Numerical approximation of partial differential equations.* Springer Berlin, Heidelberg, 1994.

[33] R. T. Rockafellar and R. J.-B. Wets. *Variational analysis.* Springer Science & Business Media, 2009.

[34] R. Selten. “Reexamination of the perfectness concept for equilibrium points in extensive games”. In: *Int. J. Game Theory* (1975), pp. 25–55. DOI: 10.1007/BF01766400.

[35] R. Selten. “Spieltheoretische behandlung eines oligopolmodells mit nachfrageträgheit: Teil i: Bestimmung des dynamischen preisgleichgewichts”. In: *Z. Gesamte Staatswiss.* (1965), pp. 301–324. URL: https://www.jstor.org/stable/40748884.

[36] A.N. Tikhonov. “Methods for the regularization of optimal control problems”. In: *Sov. Math. Dokl.* 6 (1965), pp. 761–763.

[37] E. van Damme. “A relation between perfect equilibria in extensive form games and proper equilibria in normal form games”. In: *Int. J. Game Theory* 38 (1984), pp. 1432–1270. DOI: 10.1007/BF01769861.

[38] E. van Damme. *Stability and perfection of Nash equilibria.* Springer Berlin, Heidelberg, 1991. DOI: 10.1007/978-3-642-58242-4.