THEMA

théorie économique, modélisation et applications

THEMA Working Paper n°2022-08
CY Cergy Paris Université, France

Strong Forward Induction in Monotonic Multi-Sender Signaling Games

Peter Vida, Takakazu Honryo, Helmuts Azacis

March 2022

Strong Forward Induction in Monotonic Multi-Sender Signaling Games*

Péter Vida†, Takakazu Honryo‡ and Helmuts Āzacis§

March 31, 2022

Abstract

We introduce a new solution concept called strong forward induction which is implied by strategic stability in generic finite multi-sender signaling games (Proposition 1) and can be easily extended to and applied in arbitrary extensive form games with perfect recall. We apply this notion to infinite monotonic signaling games and show that a unique pure strong forward induction equilibrium exists and its outcome is necessarily non-distorted (Theorem 1). Finally, we show that in this class of games the non-distorted equilibrium outcomes are limits of stable outcomes of finite games (Proposition 2).

Keywords: multi-sender signaling, forward induction, strategic stability, monotonic games

JEL Classification Numbers: C72, D82

*The first two authors received financial support from SFB/TR 15, which is gratefully acknowledged. Péter Vida also received financial support from the Cowles Foundation and from Labex MME-DII, which is also gratefully acknowledged. The third author gratefully acknowledges hospitality and financial support of Matthias Corvinus Collegium through their Visiting Fellowship Program. We thank Malin Arve, Itzhak Gilboa, Srihari Govindan, Johannes Hörner, Navin Kartik, Donald Keenan, Sina Litterscheid, Pauli Murto, Kathleen Nosal, Luciano Pomatto, Larry Samuleson, Joel Sobel, and Thomas Tröger. All remaining errors are ours.

†Corresponding author: CY Cergy Paris Université, CNRS, THEMA, F-95000 Cergy, France and Corvinus University of Budapest, Corvinus Institute for Advanced Studies, E-mail: vidapet@gmail.com

‡Doshisha University, Department of Economics, Kyoto, Japan. E-mail: thonryo@mail.doshisha.ac.jp

§Cardiff Business School, Cardiff University, Aberconway Building, Colum Drive, Cardiff CF10 3EU, Wales, UK. E-mail: azacish@cf.ac.uk

1 Introduction

This paper studies a class of infinite monotonic multi-sender signaling games in the spirit of Cho and Sobel (1992). We introduce a strong form of belief restriction (dubbed strong forward induction) and, as our main contribution, we select a unique equilibrium in these games. We also show that strong forward induction is implied by strategic stability (as defined in Kohlberg and Mertens (1986)) in finite generic games and that our selected outcome is a limit of stable outcomes of approximating finite games.

Banks and Sobel (1987), Cho and Kreps (1987), and Cho and Sobel (1990) successfully rule out unintuitive equilibria in the single sender setting by the application of solution concepts which restrict the support of the receiver's belief (these concepts are: the intuitive, the D1, the Divinity, the D2, the Universal Divinity or the NWBR criteria). We show by an example in section 2.3, that in the multi-sender setting highly unintuitive equilibria may survive even forward induction (as defined in Kohlberg and Mertens (1986) or as in Cho (1987)).

The main weakness of these well known solution concepts when applied in the multi-sender setting is that they are not taking into account the information conveyed by the equilibrium signal of the senders who were (possibly) not deviating. This information could obviously further restrict the support of the receiver's belief and yield stronger predictions (see our example in section 2.3). Such a belief restriction, called unprejudiced beliefs, is analyzed in Vida and Honryo (2021). They assert that for generic finite multi-sender signaling games there is always an equilibrium outcome which can be supported by beliefs which are unprejudiced and satisfy forward induction at the same time. Roughly speaking, we say that such an equilibrium outcome satisfies strong forward induction (henceforth: SFI). We prove their assertion now in Proposition 1 which states that for generic finite games strategically stable outcomes satisfy SFI and hence, such an outcome generically exists.

The main contribution of this paper is that we apply the notion of SFI to a class of infinite monotonic multi-sender signaling games and characterize the unique pure SFI outcome (Theorem 1).

In our games the senders are assumed to have complete information and this information is unknown for the receiver. This assumption about the information structure is common and frequent in the literature with multiple senders, including a huge part of the implementation literature, mechanism design, social choice and many other settings in finance, industrial organization, law and economics and political economy.[1] Our monotonicity assumptions are exact parallels of those in Cho and Sobel (1992) tailored to the multi-sender setting.

The selected outcome is fully separating and the equilibrium signals are non-distorted by the a priori asymmetric information (as opposed to the single sender case), i.e. it is as if the senders and the receiver were playing a subgame perfect equilibrium of the associated complete information games where the receiver also knows what the senders know. In Proposition 2 we state and prove that any such equilibrium outcome is strategically stable in the finitistic sense (see the precise definition of the finitistic approach in the Appendix). Hence, in this sense, in the spirit of Proposition 3.2 in Cho and Sobel (1992), we have also shown that in our class of games pure SFI outcome(s) are also (finistically) stable outcome(s).

The paper is structured as follows. In section 2 we set up the basic game, we define our solution concept and provide an elaborated example. In section 3 we define our class of infinite monotonic games, the notion of non-distorted outcome and state and prove our main Theorem 1. In section 4 we discuss the advantages of the finitistic approach with the help of which some of our assumptions can be significantly relaxed and state our Propositions 1 and 2. In section 5 we conclude. Some the definitions and proofs are relegated to the Appendix.

2 The Model

First we define multi-sender signaling games. The formulation is based on Banks and Sobel (1987) and Cho and Sobel (1990). There are finitely many senders and the set

[1]Just to mention a few, see for example Bagwell and Ramey (1991), Battaglini (2002), Emons and Fluet (2009), Bester and Demuth (2015), Schultz (1996), (1999), Zhang (2020) and Hartman-Glaser and Hébert (2019).

of senders is denoted by S with $|S| > 1$. There is also a single receiver. A generic sender is denoted by $i \in S$ and the other senders by $-i$. At the beginning of the game, senders learn their common type which is unknown to the receiver. Namely, senders' types are perfectly correlated. This information is the senders' type t, an element of the set of the first T integers, and we also denote the set of types by T. Senders' type is drawn according to some probability distribution $\pi \in \Delta^{\epsilon} T$, where π is common knowledge among the players and $\pi(t)$ is the probability of t.[2] We denote type t sender i by (i, t). After the senders learn their type, each sender i simultaneously sends a signal m_i to the receiver. The set of possible signals for sender i is M_i, and we denote $\Pi_{i \in S} M_i$ by M. A generic signal profile is $m \in M$. The receiver responds to the senders' signals by taking an action a from a set A. Sender i's payoff function is $u_i(t, m, a)$, and the receiver's payoff function is $v(t, m, a)$.

2.1 Strategies and Equilibria

We concentrate on pure strategies. We represent a pure strategy of sender i by $m_i(\cdot)$, where for each t, $m_i(t) \in M_i$ and we write $m(\cdot)$ for a profile of the senders' strategies. We represent a pure strategy of the receiver by $e(\cdot)$, where for each m, $e(m) \in A$.

Any combination of pure strategies $m(\cdot)$ and $e(\cdot)$, together with π, induce a probability distribution over the terminal nodes of the game, which we identify with $T \times M \times A$. This probability distribution over $T \times M \times A$ is called the outcome of the game induced by the strategies $m(\cdot)$ and $e(\cdot)$ and is denoted by $[m(\cdot), e(\cdot)]$.

The receiver's beliefs about the type of the senders after signal profiles is a collection of probability distributions $\mu = (\mu_m)_{m \in M}$ over T and let us write $\mu = (\mu_m, \mu_{-m})$ where $\mu_{-m} = (\mu_{m'})_{m' \in M \setminus \{m\}}$. Our first simplifying assumption is:

A1. $A = [\underline{a}, \overline{a}]$ and for all t, m: $v(t, m, a)$ is a strictly concave and differentiable function of a.

We let $e(\mu)(m)$ be the unique best response to $m \in M$ given the assessment μ and

[2] For any finite set X, $\Delta^{\epsilon} X$ denotes the set of probability distributions for which we have that $\pi(x) \geq \epsilon$ for all $x \in X$.

we denote by $e(\mu)(\cdot)$ the corresponding strategy. That is:

$$e(\mu)(m) := \arg\max_{a \in A} \sum_{t=1}^{T} v(t, m, a)\mu_m(t).$$

A Perfect Bayesian Equilibrium (PBE) is a triple of strategies and assessments $(m(\cdot), e(\cdot), \mu)$ that satisfies:

(1) Sequential Rationality: (a) $e(\cdot) = e(\mu)(\cdot)$ and (b) for all t and $i : m_i(t) \in \arg\max_{m_i} u_i(t, m_i, m_{-i}(t), e(m_i, m_{-i}(t)))$,

(2) Bayes rule: for every t: $\mu_{m(t)}(t) = \frac{\pi(t)}{\sum_{t':m(t')=m(t)} \pi(t')}$.

2.2 The Solution Concept: Strong Forward Induction

Fix a PBE $(m(\cdot), e(\cdot), \mu)$ and let us denote by $u_i^*(t)$ the equilibrium utility of (i, t) in this PBE. We impose the following restriction on μ for certain out-of-equilibrium signal profiles m. Consider any m for which there is an (i, t) such that $m_{-i}(t) = m_{-i}$ and there is no t' such that $m_i = m_i(t')$, and let us denote by $T_m = \{t | m_{-i}(t) = m_{-i}\} \neq \emptyset$ for such an m. At such an m, the receiver knows for sure that sender i was deviating. We are going to require that the receiver believes at such signal profiles that sender i was deviating unilaterally (unprejudiced beliefs). Moreover, we are going to require that the receiver cannot exclude the possibility of any such unilateral deviation of sender i (open-mindedness), however most of the weight must be put on types for which the deviation is a weak best response (forward induction).[3]

To this end, let us set $F_m = \{t \in T_m | \exists a \in A : \forall t' \in T_m : u_i(t', m, a) \leq u_i^*(t'), u_i(t, m, a) = u_i^*(t)\}$. That is, F_m is the set of types $t \in T_m$ such that for (i, t) sending the signal m_i is a weak best response.[4]

Fix an $\varepsilon \geq 0$. We say that $(m(\cdot), e(\cdot), \mu)$ satisfies $\varepsilon-$*strong forward induction* if for all m at which the receiver knows for sure that some sender i was deviating, we

[3] We could have restricted beliefs for more out-of-equilibrium signal profiles, but it is not needed for our result. In fact, we do not want to require open-mindedness at out-of-equilibrium signal profiles where there is no sender about which the receiver knows for sure that he was deviating. Compare this with footnote 17.

[4] We could restrict even more the set F_m by requiring that the action a in its definition must be a sequentially rational action of the receiver for some belief.

have that:

1. if $F_m = \emptyset$ or $F_m = T_m$ then $\mu_m \in \Delta^\varepsilon T_m$

2. if $F_m \neq \emptyset$ and $F_m \neq T_m$ then $\mu_m(F_m) = 1 - \varepsilon$ and $\mu_m(T_m \setminus F_m) = \varepsilon$,

Point (1) and point (2) require that μ_m is concentrated on T_m, that is, the receiver must believe that only sender i was deviating (unprejudiced beliefs). Point (2) requires that if possible, the belief puts a total weight of $1 - \varepsilon$ on those types of i for whom it is a weak best response to send the signal m_i (forward induction) and puts a total weight of ε on those types of i for whom it is never a weak best response to send the signal m_i (open-mindedness). Otherwise by point (1), each type in T_m must get a weight of at least ε.

A PBE outcome $[m(\cdot), e(\cdot)]$ satisfies *strong forward induction* if there is an $\bar{\varepsilon} > 0$ such that for all $\varepsilon < \bar{\varepsilon}$ there is a μ^ε such that $(m(\cdot), e(\mu^\varepsilon)(\cdot), \mu^\varepsilon)$ satisfies $\varepsilon-$strong forward induction.[5]

2.3 An Example

In this example we show that there is a distorted equilibrium which survives forward induction but the equilibrium does not survive strong forward induction.[6] Consider the following setup inspired by Bagwell and Ramey (1991). Two firms learn their common type $t \in \{0, 1\}$. For all $i = 1, 2$ firm i chooses a price $m_i \in [0, \overline{m}_i]$ and receives profit according to $(m_i - t)(d - m_i + m_{-i})$. Finally, the entrant chooses an action $a \in [0, 1]$, which can depend on the observed price vector $m = (m_1, m_2)$. The entrant's utility is described by $v(t, a) = -(t - 1 + a)^2$. Firm i's overall utility is given

[5]Given that $m(\cdot)$ is fixed along the sequence as ε converges to 0, we have that $\mu^\varepsilon_{m(t)} = \frac{\pi(t)}{\sum_{t':m(t')=m(t)} \pi(t')}$ is also fixed by Bayes rule for all $t \in T$ and hence $e(\mu^\varepsilon)(m(t)) = e(m(t))$ and $[m(\cdot), e(\cdot)] = [m(\cdot), e(\mu^\varepsilon)(\cdot)]$ are also fixed, i.e. the outcome is fixed along the sequence. Open-mindedness and ε play an important role and have extra bite only in infinite games and their role becomes clear in the proof of Theorem 1 (see also footnote 17 and remark 2 in the Appendix for further clarification).

[6]By forward induction we mean the definition of Never Weak Best Response of Cho and Kreps (1987) or the definition of forward induction in Kohlberg and Mertens (1986). Strong forward induction mainly differs from forward induction in that additionally to forward induction, the belief is also restricted to be supported in T_m, i.e. unprejudiced beliefs are also required.

by:

$$u_i(t, m, a) = (m_i - t)(d - m_i + m_{-i}) + ak,$$

where $k < 0$ is a constant. The interpretation is that d is a demand parameter, t is a cost parameter which is unknown by the entrant but it is publicly known by the firms. Firms set prices, that is, they send signals and receive the corresponding profits. After observing the prices, the receiver, who is the entrant, chooses an effort level of entry a. Firms like small effort levels, while the entrant chooses higher effort when his belief about the industry cost is lower. Let us fix $d = 2$ and $k = -2$.

In any separating PBE outcome it must be that $m(0) = (2, 2), e(2, 2) = 1$ resulting in payoffs 2 for the firms of type 0. This is because type 0 firms in a separating equilibrium face maximal effort by the entrant, their prices should be mutual best responses to one another in the profit game as otherwise deviations to larger profit levels could not be deterred with higher effort. Suppose, however, that $m(1) \neq (3, 3)$ that is, the equilibrium is distorted in the sense that the equilibrium signal profile of type 1 firms is different from (3,3) which is the unique price profile where firms of type 1 mutually best respond to each other in the profit game. Say $m_1(1) = 2.8 \neq (m_2(1)+2+1)/2, m_2(1) = (2.8+2+1)/2 = 2.9$ and $e(2.8, 2.9) = 0$ because the entrant knows that $t = 1$ in any separating equilibrium. That is firm 2 is best responding to firm 1's price but firm 1 could increase its profit. This can be easily maintained as a PBE outcome. To see this, suppose that firm 1 deviates to some $m'_1 \in (2.8, 3.1)$ which increases its profit. To deter such a deviation we can set the entrant's belief after the signal pairs $(m'_1, 2.9)$ in such a way that the sequentially rational effort level of the entrant is sufficiently high. Say, the beliefs are concentrated on $t = 0$ and the corresponding actions are 1. It is easy to see that given such beliefs firm 1 of type 1 has no incentives to deviate and increase its profit because that would induce too high (maximal) effort from the entrant.

Notice that any belief for which the sequentially rational action of the entrant maintains the equilibrium must put a positive probability on the event that the firms are of type 0, i.e. $\mu_{(m'_1,2.9)}(0) > 0$. Such a belief, and hence the equilibrium, is ruled out by strong forward induction, because the belief of the entrant must be concentrated on $t = 1$ given firm 2's signal is $m_2(1) = 2.9$ since the entrant knows for sure that firm 1 was deviating and $T_{(m_1',2.9)} = \{1\}$.

However, such beliefs, i.e. those having a positive probability on the event that the firms are of type 0, are consistent with forward induction. Simple calculation shows that for firm 2 of type 0 it is a weak best response to send the signal 2.9 and for firm 1 of type 0 it is also a weak best response to send any signal $m_1' \in (2.8, 3.1)$.[7] It follows that using only forward induction, the entrant cannot exclude the possibility that the out-of-equilibrium signal profile is a consequence of the simultaneous deviations of firms 1 and 2 of type 0 and hence the entrant is allowed to put positive weight on $t = 0$.

We show in the sequel that this game has a unique equilibrium which survives strong forward induction in which type 0 firms set prices (2,2) and type 1 firms set prices (3,3), i.e. equilibrium signals are non-distorted.

3 Monotonic Signaling Games, The Theorem

We consider a class of infinite multi-sender signaling games that satisfy certain monotonicity conditions, and which is a natural extension of the monotonic single sender games considered in Cho and Sobel (1990). We show that in strong forward induction equilibria (henceforth, SFI) of these games all types separate and equilibrium signals are non-distorted. We state in section 4 and prove it in the Appendix that the finitistic limits of stable outcomes are non-distorted. In this sense SFI implies stability in this class of games.

For simplicity we assume that the signal spaces are open real intervals.[8]

A0. For all $i \in S$ and for all $t \in T : u_i(t, m, a)$ and $v(t, m, a)$ are continuous in (m, a).[9]

A2. For all $a', a : a' > a$ implies that $u_i(t, m, a') > u_i(t, m, a)$ for all i, t, m.

[7] Set $\mu_{(2,2.9)}(0) = e(2, 2.9) = 0.595, \mu_{(m_1',2)}(0) = e(m_1', 2) = m_1'(4 - m_1')/2 - 1 \in [0, 1]$ for all $m_1' \in [2.8, 3.1]$, resulting in a payoff $u_2(0, (2, 2.9), 0.595) = u_1(0, (m_1', 2), m_1'(4 - m_1')/2 - 1) = 2$ for both firms of type 0 from the deviation to 2.9 and to m_1' respectively, which is just their equilibrium payoff. Clearly, firm 1 of type 1 has no incentives to deviate to $m_1'' = 2$ given $e(2, 2.9) = 0.595$. For other out-of equilibrium message pairs the beliefs can also be chosen to satisfy forward induction.

[8] We discuss the role of this simplifying assumption and how to dispense with it in footnote 14 in the proof of our theorem.

[9] A0-A1 guarantees that $e(\mu)(\cdot)$ is a function, which is continuous in μ and m.

A3. For all $(t, m), \partial v(t, m, a)/\partial a$ is strictly increasing in t and for all m we have that $\arg\max_{a \in A} v(T, m, a) < \overline{a}$.[10]

A4. For each $i \in S$, for all $m_i, m_i', m_{-i}, a, a', t, t'$ such that $m_i < m_i'$ and $t < t'$:

$$u_i(t, m_i, m_{-i}, a) \leq u_i(t, m_i', m_{-i}, a') \text{ implies } u_i(t', m_i, m_{-i}, a) < u_i(t', m_i', m_{-i}, a').$$[11]

Before stating our theorem, we define the notion of non-distorted outcomes. Consider the degenerate incomplete information games indexed by t where the priors are concentrated on t and the utilities, the signal sets, and the action set is exactly the same as in the original game. For all $t \in T$, let us denote by $N(t) \subseteq M \times A$ the set of pure subgame perfect equilibrium outcomes $[m, e(m)]$ of these games. Now consider the original incomplete-information game and an outcome $[m(\cdot), e(\cdot)]$. We say that an outcome is *non-distorted* if for all types $t, t' \in T$ we have that $m(t) \neq m(t')$ and $[m(t), e(m(t))] \in N(t)$.[12]

To ensure uniqueness and existence of SFI we need an additional assumption about the complete-information games together with a technical assumption which significantly simplifies the exposition and the proof of our result. We discuss how to dispense with this seemingly demanding technical assumption in Section 4.

B1. For all $t \in T : |N(t)| = 1$.

B2. For all $t, t' \in T, [m(t), e(m(t))] \in N(t), [m(t'), e(m(t'))] \in N(t') : \exists i, j \in S : m_i(t) \neq m_i(t'), m_j(t) \neq m_j(t')$.[13]

[10]A2 and A3 are the reasons why we call our games monotonic signaling games.

A2 indicates that all types of all senders have identical preferences over the receiver's action and prefer higher actions. We could have assumed that all senders prefer lower actions. Also, all of our results hold with senders having monotonic preferences in different directions as long as the non-distorted outcome is a PBE. This is the case e.g. when the number of senders is at least 3 and types are separated at least by 3 senders (see the similar assumption B2 for two senders below, footnote 15, and how to dispense with this assumption using the finitistic approach in Section 4).

A3 together with A1 implies that $e(\mu)(m)$ is strictly increasing in μ_m in the sense of first-order stochastic dominance. We could have assumed as well that $\partial v/\partial a$ is strictly decreasing in t.

[11] A4 is a single crossing condition. It states that having fixed some pure signal of $-i$, if sender i of a certain type is indifferent between two signal-action pairs and one signal is greater than the other, then all higher types of sender i strictly prefer to send the greater signal and receive the corresponding action. The assumption guarantees that higher types are more willing to send higher signals than lower types. The single crossing condition can also hold in the other direction, i.e., with $t > t'$, and can hold in different directions for different senders.

[12] Notice that a non-distorted outcome is not necessarily a PBE outcome because deviations to signal profiles where it is not sure that a certain sender was deviating, could not be deterred.

[13]B1 requires that the complete-information games possess a unique pure subgame perfect Nash

We are now ready to state our main theorem.

Theorem 1. *1. Under B2, any non-distorted PBE outcome is an SFI outcome.*

2. Under assumptions A0-A4, any SFI outcome is non-distorted and satisfies B2.

3. Under assumptions A0-A4, and B1-B2 there exists a unique SFI outcome which is the non-distorted outcome.

Proof. Statement 1: Given that B2 is satisfied, T_m is always a singleton at $m-$s where it is sure that a certain sender i is deviating and hence, to comply with SFI, one must choose $\mu_m(t) = 1$ where t is such that $m_{-i} = m_{-i}(t)$. But then by non-distortion one can choose $e(m)$ to be the subgame perfect action of the receiver in the complete information game where the prior is concentrated on t. This $e(m)$ deters such a deviation. Hence the construction of the (constant) sequence of $\varepsilon-$SFI equilibria is trivial.

Statement 2: We prove the statement in 3 steps.

(1) All types separate: By contradiction, consider an $\varepsilon-$SFI outcome with a pooling signal profile m. Given that the signal spaces are open, there is an $m' = (m'_i, m_{-i})$ such that $m'_i > m_i$, there is no $t \in T$ for which $m_i(t) = m'_i$ and i prefers higher actions of the receiver by A2.[14] By A4 and by A0,A1 and A3, m'_i can be and is a weak best response only for the highest type t' of sender i in the pool if m'_i is sufficiently close to m_i, i.e. $F_{(m'_i, m_{-i})} = \{t'\}$. Hence, the belief at m' must put probability $1 - \varepsilon$ on t'. As a consequence, if ε is sufficiently small, just as in the single-sender case, the receiver's belief after m' jumps upward to t' relative to the one after the pooling signal profile m. By A0, A1, and A3, this induces an upward jump in the receiver's action resulting in a profitable deviation for any type of sender i in the pool if m'_i is close enough to m_i.

equilibrium outcome (see e.g. Rosen (1965) for assumptions on the primitives applying them to the Nash equilibrium of the game between the senders where the receiver's action after any signal profile is just his best reply given his belief is concentrated on any fixed $t \in T$.) B2 requires that in these equilibria for any pair of types there is a pair of senders having different non-distorted signals for this pair of types. Namely senders can separate all types with non-distorted signals, moreover each type is separated by at least two senders.

[14] If the signal spaces were compact, one can impose a technical assumption which is the multi-sender counterpart of A6 in Cho and Sobel (1990) and rules out pooling at the critical corners of M.

(2) Any pair of types are separated at least by two senders: By contradiction, suppose there are types t, t' with $t < t'$ for which $\exists! i$ such that $m_i(t) \neq m_i(t')$ and $m_{-i}(t) = m_{-i}(t')$. Choose t to be the smallest type for which this is true. Fix the $\varepsilon > 0$ and consider a deviation m_i' of (i, t) sufficiently close to $m_i(t)$ such that there is no type t'' for which $m_i(t'') = m_i'$. Now by $\varepsilon-$SFI (i, t) can get an extra weight, in the belief of the receiver, of at least ε on some set of types $t'' > t$, with which senders $-i$ pool with t. Hence, the receiver's belief increases to $\mu_{m_i', m_{-i}(t)} >_{stoch.\ dom.} \mu_{m(t)}(t) = 1$. The receiver's action also increases to $e(\mu)(m_i', m_{-i}(t)) > e(\mu)(m(t))$ if m_i' is sufficiently close to $m_i(t)$, because ε is fixed, which results in a profitable deviation of (i, t) once m_i' is sufficiently close to $m_i(t)$ and so, it is not too costly for sender i.

(3) For all $t : [m(t), e(m(t))] \in N(t)$ (non-distortion): By contradiction, suppose that there is a t, i and m_i' such that $u_i(t, m(t), e(m(t))) < u_i(t, (m_i', m_{-i}(t)), e(\mu)(m_i', m_{-i}(t)))$, where $\mu_{(m_i', m_{-i}(t))}(t) = 1$, i.e. in the complete information game where the prior is concentrated on t, sender i has a profitable deviation given that the receiver acts sequentially rationally. Notice that m_i' can be chosen to be such that there is no t' for which $m_i(t') = m_i'$. But then given that $T_m = \{t\}$, since every pair of types is separated by at least two senders (see (2) above), SFI requires that $\mu_{(m_i', m_{-i}(t))}(t) = 1$ and hence it must be that $e(m_i', m_{-i}(t)) = e(\mu)(m_i', m_{-i}(t))$ which together with the above strict inequality indicates a profitable deviation for (i, t).

Statement 3: B1-B2 ensures that the non-distorted outcome is unique and all types are separated by at least two senders. Hence, by statement 2, this can be the only SFI outcome. A2 ensures that it is a PBE outcome, i.e. after out-of-equilibrium signal profiles, where there is no sender about whom the receiver knows that he was surely deviating, the actions of the receiver can be chosen to deter deviations. It is because all the senders prefer higher actions of the receiver and one can set beliefs concentrated on the lowest type of the possible deviators.[15] By the 1st statement this non-distorted PBE outcome is indeed an SFI outcome which concludes the proof of the 3rd statement. □

[15] If each pair of types are separated at least by 3 senders, the deviator can always be identified after unilateral deviations and one could allow the senders to have different preferences about the receiver's action (see also footnote 10). In this case it is also true that all non-distorted outcomes are PBE outcomes and they are also SFI outcomes (see also footnote 12).

4 The finitistic approach

In this section we informally introduce the finitistic approach which allows us to tackle the problems arose when dealing with infinite games and also allows us to connect our solution concept to those available for finite games.

The finitistic approach allows us to think about outcomes of the infinite games in a broader sense, i.e. as the limit of outcomes of finite approximation of the infinite game as the signal and action spaces become richer. Clearly, not every limit of outcomes of approximating finite games is an outcome (in the usual sense) in the limit game because types which may separate along the sequence may pool at the limit.[16]

Assumption B2 seems rather restrictive, moreover it is implied by SFI in our class of games. We argue in section 4.1 that the finitistic approach solves this problem and we can dispense with B2. We also argue that the finitistic approach solves an even more severe situation when for two types the corresponding subgame perfect signal profiles completely coincide and these types can only separate with distorted signals.

The finitistic approach also allows us to connect our solution concept to those used in finite games. First, in section 4.2 in Proposition 1, we state that in generic finite games, SFI is implied by strategic stability. Second, in Proposition 2 we state, that in our class of games, non-distorted outcomes are limits of stable outcomes. Hence, given statement 2 of Theorem 1, SFI implies stability in our class of games in the finitistic sense. In the Appendix, we give a formal definition of the finitistic approach, and therein in Remark 2 we discuss the relation between the various limit and infinite solution concepts and suggest a limit solution concept for infinite games in which B2 is not satisfied or in which the subgame perfect signals cannot separate certain types at all.

4.1 Pooling only in the limit, dispensing with B2

Non-distorted outcomes are separated outcomes by definition. B2 requires a rather strong form of separation. We argue now with the help of an example that we can

[16] The finitistic approach was introduced by Simon and Stinchcombe (1995).

dispense with this strong form of separation once we allow outcomes in the infinite game to be also the limits of outcomes of approximating finite games.

Consider the two-sender, two-type version of the job market signaling model of Spence (1973), where education does not increase the marginal products, types are perfectly positively correlated and all of our assumptions but B2 are satisfied. In this game, the corresponding subgame perfect equilibrium education levels are 0 for both types and both senders and the situation is even more severe. Not only that the types cannot be separated by the corresponding subgame perfect equilibrium signals of at least two senders, as it is required by B2, but they cannot be separated at all by these signals. It is easy to see that in the finite versions of the multi-sender Spence model there are $\varepsilon-$SFI outcomes in which the low types choose 0 education level and the high types choose the lowest feasible education level, denoted by 0^+, which is different from 0 and hence both senders separate.[17]

In the limit, types pool on the non-distorted 0 education level yet the action of the receiver is different for the low and for the high types as they separate for both senders along the equilibria of the approximating sequence.

4.2 Connection with stability

Consider the obvious generalization of SFI to mixed strategies in finite games. The following proposition further justifies our solution concept using the notion of stability

[17] The only problematic possible deviation is for example that the high type of sender one chooses 0 education level. From the receiver's perspective, after observing the signal profile $(0, 0^+)$, it is also possible that the low type of sender two deviated to 0^+. The belief of the receiver after this type of out-of-equilibrium signal profiles is not restricted by $\varepsilon-$SFI because there is no sender about which the receiver knows for sure that he has deviated. The receiver is allowed to believe that the low type of sender two deviated to 0^+ with probability 1. Hence, no matter how rich the signal space is and how close 0^+ is to 0, neither the low type of sender two nor the high type of sender one will have incentives to deviate to 0^+ or to 0 respectively.

We note that there is another (distorted) equilibrium of interest in which one of the senders chooses the Riley outcome while the other sender pools on 0 education level. This outcome is also a limit of SFI outcomes and it is also a 0-SFI but it is not an SFI as it is distorted and also violates B2 (see statement 2 in Theorem 1). In fact, it is also a limit of stable outcomes. To see this notice that the incentives of the low type of the pooling sender to deviate can always be stabilized and the argument for the separating sender is the same as in the single sender case.

See Remark 2 in the Appendix for a definition of a finitistic solution concept which selects the non-distorted outcome and rules out the distorted one.

á la Kohlberg and Mertens (1986).[18]

Proposition 1. *In generic finite games, stable outcomes satisfy strong forward induction.*

Proof. See the proof in the Appendix. □

According to statement 2 of Theorem 1, in monotonic games, SFI outcomes are non-distorted. Together with the following proposition we have that SFI implies stability in our class of games in the finitistic sense.

Proposition 2. *Under assumptions A0-A3, a non-distorted outcome is the limit of stable outcomes of the approximating finite games.*

Proof. See the proof in the Appendix. □

5 Conclusion

We have introduced a new and powerful solution concept which can be easily applied to any (even infinite) extensive form games with perfect recall. In generic finite multi-sender games it is implied by strategic stability and hence a solution generically exists. In our main Theorem 1 we have shown how powerful this selection is in monotonic infinite multi-sender games and demonstrated that the solution is non-distorted. Moreover, we have also shown that the selected equilibrium outcome is a limit of stable outcomes of approximating finite games.

References

[1] Bagwell, K., and G. Ramey 1991. Oligopoly Limit Pricing. *RAND Journal of Economics,* 22, 155-172.

[18]The formal definition of stable sets can be found in the Appendix.

[2] Banks, J., and J. Sobel 1987. Equilibrium Selection in Signaling Games. *Econometrica*, 55, 647-661.

[3] Battaglini, M. 2002. Multiple Referrals and Multidimensional Cheap Talk. *Econometrica*, 70/4, 1379-1401.

[4] Blume, L., and W. Zame 1994. The Algebraic Geometry of Perfect and Sequential Equilibrium. *Econometrica*, 62, 783-794.

[5] Cho, I. 1987. A Refinement of Sequential Equilibrium. *Econometrica*, 55,6 1367-1389.

[6] Cho, I., and D. Kreps 1987. Signaling Games and Stable Equilibria. *Quarterly Journal of Economics*, 102, 179-221.

[7] Cho, I., and J. Sobel 1990. Strategic Stability and Uniqueness in Signaling Games. *Journal of Economic Theory,* 50, 381-413.

[8] Hartman-Glaser, B. and B. Hebert 2019. The Insurance Is the Lemon: Failing to Index Contracts. *Journal of Finance*, Vol. 75, Issue. 1, 463-506.

[9] Kohlberg, R., and J.-F. Mertens 1986. On the Strategic Stability of Equilibria. *Econometrica*, 54, 1003-1038.

[10] Rosen, J. 1965. Existence and Uniqueness of Equilibrium Points for Concave N-Person Games. *Econometrica*, 33, 520-534.

[11] Schultz, C. 1996. Polarization and Inefficient Policies. *Review of Economic Studies,* 63, 331-344.

[12] Schultz, C. 1999. Limit Pricing When Incumbents Have Conflicting Interests. *International Journal of Industrial Organization,* 17, 801-825.

[13] Simon, L., and M. Stinchcombe 1995. Equilibrium Refinement for Infinite Normal-Form Games. *Econometrica*, 63, 1421-1443.

[14] Spence, M. 1973. Job Market Signaling. *The Quarterly Journal of Economics*, 87, 355-374.

[15] Vida, P., and T. Honryo 2021. Strategic Stability of Equilibria in Multi-Sender Signaling Games. *Games and Economic Behavior*, 127, 102-112.

[16] Zhang, Q. 2020. Vagueness in Multidimensional Proposals. *Games and Economic Behavior*, 121, 307-328.

Appendices

6 Proof of Proposition 1, Definition of Stability

First we define stable sets of equilibria à la Kohlberg and Mertens (1986) for multi-sender signaling games. Consider the (reduced) normal form Γ of a finite multi-sender signaling game. Let $\sigma = (\sigma_1, \ldots, \sigma_{|S|})$, where σ_i is a completely mixed-strategy of sender $i \in S$.[19] For $\delta > 0$, consider the set of all normal form games Γ' that have the same strategy space as Γ and for which for all $i \in S$ there exists $\delta_i \in (0, \delta)$, such that if some strategy profile $(\sigma^*, e(\cdot))$ is played in Γ', then the payoffs are the same as when each sender $i \in S$ plays $(1 - \delta_i)\sigma_i^* + \delta_i \sigma_i$ and the receiver plays $e(\cdot)$ in Γ. A game in this set is called a (σ, δ) perturbation of Γ.

Definition 1. *A set of Nash equilibria of* Γ *is* stable *if it is minimal with respect to the following property:* $\mathcal{N}$ *is a closed set of Nash equilibria of* Γ *satisfying: for each* $\epsilon > 0$, *there is a* $\delta > 0$ *such that for any completely mixed* σ *the* (σ, δ) *perturbations of* Γ *have a Nash equilibrium* ϵ*-close to* $\mathcal{N}$.

Proof of Proposition 1: Consider a stable set. By Proposition 6(B) in Kohlberg and Mertens (1986) this stable set contains a stable set of the game obtained by deleting strategies which are never weak best responses (inferior), e.g. those who are not in F_m. By Theorem 1 in Vida and Honryo (2021) in generic games this stable set then contains a PBE in which beliefs, after any m where there is a sender who was deviating for sure, are supported on F_m if F_m is not empty, and on T_m if $F_m = \emptyset$. Additionally the ε properties in point 1., point 2. can be also satisfied for some $\bar{\varepsilon}$ sufficiently small because the game is finite.

[19]For simplicity, we perturb only the strategies of the senders (just as in the literature of the single-sender case), as we are interested in the beliefs generated by the stabilization of these trembles. Abusing notation slightly, we can identify mixed and behavioral strategies.

7 The Finitistic Approach for Infinite Games

We extend the scope of solution concepts defined for finite games to infinite games following the finitistic approach introduced by Simon and Stinchcombe (1995) and in Remark 2 below we define a limit solution concept for infinite games in which B2 need not be satisfied or in which the subgame perfect signals might not separate certain types at all and which selects the non-distorted limit outcome and resolves the multiplicity problem discussed in footnote 17.

Consider a multi-sender signaling game form G with $M_i = [\underline{m}_i, \overline{m}_i]$ for all $i \in S$ and $A = [\underline{a}, \overline{a}]$ being real compact intervals, with finite type space T, and fix the utility functions of the senders $u = (u_1, \ldots, u_{|S|})$ and of the receiver v, as defined in Section 2. A sequence of finite multi-sender game forms $(G^n)_{n\in\mathbb{N}}$ is a finite approximation of G if the corresponding sequence of set of signals M_i^n and set of responses A^n are subsets of M_i and A, respectively, and converge in the Hausdorff distance to M_i and A, respectively, for all $i \in S$. For any G^n, consider the point in $x^n \in \mathbb{R}^{\dim G^n}$ induced by (u, v), where $\dim G^n = (|S|+1)|T \times M^n \times A^n|$, where $M^n = \Pi_{i\in S} M_i^n$. Let $B(x^n, \epsilon^n)$ be the $\epsilon^n > 0$ ball around x^n, say, in the Euclidean metric, and let us choose open sets $D^n \subseteq B(x^n, \epsilon^n)$ for all n with $\epsilon^n \to 0$. $(D^n)_{n\in\mathbb{N}}$ is called a sequence of payoff perturbations. Let $\mathcal{R}$ denote some solution concept for finite multi-sender signaling games. Then:

Definition 2. *Fix an infinite multi-sender signaling game form G with (u, v). We say that $\lambda \in \Delta(T \times M \times A)$ is a (pure) $\mathcal{R}^*$ outcome of the infinite game if there is a finite approximation $(G^n)_{n\in\mathbb{N}}$ of G together with a sequence of payoff perturbations $(D^n)_{n\in\mathbb{N}}$, such that for any sequence $(u^n, v^n)_{n\in\mathbb{N}}$, for which $(u^n, v^n) \in D^n$ for all $n \in \mathbb{N}$, there is a corresponding sequence of (pure) $\mathcal{R}$ outcomes $(\lambda^n)_{n\in\mathbb{N}} : \lambda^n \in \Delta(T \times M^n \times A^n)$ of the games G^n, with (u^n, v^n), weakly converging (in the weak* topology) to λ.*

Remark 1. *The requirement that the D^n sets are open in the definition is necessary because Proposition 1 holds only for generic games and we wanted to be sure that an SFI* outcome always exists. We show in the supplementary material that a stable**

outcome always exists.[20] *It then simply follows from Proposition 1 and from Definition 2 that an SFI* outcome also always exists. The proof of Proposition 2 also exploits the fact that these sets are open.*

Remark 2. *In finite games any 0-SFI outcome is also an* ε^n-*SFI outcome for some* ε^n *sufficiently small which may clearly depend on the level of approximation* n *(i.e. on the richness of the signal spaces) and may converge to 0. It follows that SFI* and 0-SFI* outcomes are the same, however, in infinite games not all 0-SFI outcomes are SFI outcomes (see for example the one described in footnote 17).*

We suggest the following solution concept: A limit outcome is $\mathcal{SFI}^\star$ *if there is an* $\bar{\varepsilon} > 0$ *such that the limit outcome is* $\varepsilon-$*SFI* for all* $\varepsilon < \bar{\varepsilon}$.

The non-distorted limit outcome described in subsection 4.1 is $\mathcal{SFI}^\star$ *while the distorted outcome described in footnote 17 is not. It is because as the approximation becomes finer, the corresponding* $\varepsilon^n > 0$, *which makes the distorted finite outcome to be* ε^n-*SFI, must converge to 0.*

8 Proof of Proposition 2

We prove the proposition for simplicity under the additional assumption B2 (see footnote 21 below how to dispense with this assumption) and only with two senders (the proof directly generalizes to arbitrarily finitely many senders).

Consider a non-distorted outcome λ. Because of A2, it is a PBE outcome and one can easily see that is also a PBE* outcome. Consider the pair $(m(\cdot), e(\cdot))$ for which $[m(\cdot), e(\cdot)] = \lambda$. Look at the sequence of tuples $(G^n, D^n, \lambda^n, m^n(\cdot), e^n(\cdot))_{n\in\mathbb{N}}$ justifying λ such that $m^n(\cdot), e^n(\cdot)$ is a PBE of the game form G^n together with any utility point $u^n, v^n \in D^n$ generating λ^n. Notice that by B2 we can choose this sequence in a way that all the types are separated by both of the senders.[21] Fix a utility

[20] Srihari Govindan pointed out to us that the proof is simple by using the technique of Blume and Zame (1994).

[21] We exploit assumption B2 at this point. In the absence of B2, one can argue that the approximating sequence can be chosen such that types separate on both sides. This can be done in a way that any choice of a signal, which is a better reply in the complete information game, result in an off signal pair where the deviator is not known. Then beliefs can be chosen freely and by A2 can be set in such a way so as the sequentially rational action of the receiver deters such deviations.

point $y^n = (u^n, v^n) \in D^n$ and a ϵ^n such that $m^n(\cdot), e^n(\cdot)$ is a strict PBE and that the non-distorted outcome is generated by strict subgame perfect equilibria for all t for all $y \in B(y^n, \epsilon^n) \subseteq D^n$.[22] Such a y^n and ϵ^n exists by the genericity of strict PBE and that the D^n sets are open. We show that λ is a pure stable* outcome for a subsequence of $(G^n)_{n\in\mathbb{N}}$ together with the sequence of open payoff perturbations $(B(y^n, \epsilon^n))_{n\in\mathbb{N}}$. To this end, fix an n and some $y \in B(y^n, \epsilon^n)$. Now we show stability of the whole component belonging to the outcome $[m^n(\cdot), e^n(\cdot)]$. From now on we suppress the superscript n. For simplicity, we perturb only the senders' strategy and with the same δ. The proof goes through for the case when these δ-s are different.

Fix an $\varepsilon > 0$. We design a $\bar{\delta}$ such that for all σ^* strategy perturbation of the senders for all $\delta < \bar{\delta}$ we have an equilibrium (σ, e) of the (σ^*, δ) perturbed game such that σ is ε close to $m(\cdot)$. Fix a $\xi < \varepsilon$. We are going to set this ξ later to be sufficiently small. Consider the following auxiliary game with the following strategy perturbations of the senders. For each $i \in S$ independently, nature chooses the signal of sender i as follows:

(1) with probability δ signals are chosen according to σ_i^*;

(2) with probability $(1-\delta)(1-\xi)$ the strategy $m_i(\cdot)$ is chosen.

With the remaining probability, which is $(1-\delta)\xi$, sender i is free to choose any random signal in ΔM_i.

There must be a mixed-strategy equilibrium of this game. Denote the senders' strategy in this equilibrium by σ. We show that in this mixed equilibrium, it must be that for all the types t of each sender i we have $\sigma_i(m_i(t)|t) > 0$. Hence σ can be transformed in the obvious way into an equilibrium of the game where sender i can freely choose his signals with probability $(1-\delta)$ and with probability δ the signal is chosen according to σ_i^*. Moreover, it will be ε-close, in fact, ξ-close to the original component we have considered. We proceed by contradiction and assume that there is a t such that w.l.o.g. $\sigma_1(m_1(t)) = 0$. Hence there must be an $m_1 \in M_1$ such that $m_1 \neq m_1(t)$ and $\sigma_1(m_1|t) \geq 1/k$, where $k = |M_1| - 1$. Let $O_1 = \{m_1' \in M_1 | \forall t \in T, m_1(t) \neq m_1'\}$ be the set of signals of sender 1 in the finite game that are never sent

[22]In strict equilibria deviators are always strictly worse off.

in the original pure equilibrium. There are two cases to consider.

First, suppose that $m_1 \in O_1$. Then the probability that the signal profile $(m_1, m_2(t))$ is sent by senders of type t can be bounded from below by $(1-\delta)^2(1-\xi)\xi/k$ (independently of σ^*). The probability that the signal profile $(m_1, m_2(t))$ is sent by senders of type $t' \neq t$ can be bounded (independently of σ^*) from above by $((1-\delta)\xi+\delta)^2$. Notice, that only k depends on n so we can choose ξ for each k to be sufficiently small so that whenever $\delta < \xi$ the ratio:

$$\frac{((1-\delta)\xi+\delta)^2}{(1-\delta)^2(1-\xi)\xi/k},$$

gets arbitrarily close to 0. That is, the receiver puts weight arbitrarily close to 1 on the event that the signal pair arrived from senders of type t, and hence plays the action $e(\mu)(m_1, m_2(t))$, where $\mu_{(m_1,m_2(t))}(t) = 1$. Choose ϵ^n in such a way that the complete-information equilibria remain strict in the finite payoff perturbed complete-information game. But then sender 1 of type t gets strictly less, as opposed to sending the (complete-information) equilibrium signal $m_1(t)$ if ξ is small enough, which is a contradiction. Hence $\sigma_i(m_i(t)) > 0$ for all t for all i and then σ can be transformed into a part of an equilibrium strategy in the perturbed game $\xi-$close to the component.

Second, suppose that $m_1 \notin O_1$, that is, there is a $t' \neq t$ such that $m_1 = m_1(t')$. By B2 there is only one such t'. It is easy to see that if σ is an equilibrium then it must be that $\sigma_2(m_2(t)|t') > 0$. Otherwise, similar to the argument of the first case, the receiver puts weight arbitrarily close to 1 in the event that the signal pair arrives from senders of type t and we reach a contradiction. But by $A2$ and $A3$ it is impossible that $\sigma_2(m_2(t)|t') > 0$ and $\sigma_1(m_1(t')|t) > 0$ hold at the same time since then the receiver's belief will be arbitrarily close to a belief between t and t'. Assume that senders prefer higher actions of the receiver. Similar argument holds if senders prefer lower actions of the receiver. Suppose that $t < t'$, but then sender 2 of type t' is strictly worse off by sending $m_2(t)$ than $m_2(t')$. If $t > t'$, a similar contradiction holds for sender 1 of type t. Q.E.D.

9 Supplementary Material, Stable* Outcome Exists

For the existence of a stable* outcome one must show that if a stable outcome exists for given payoffs then it also exists for all payoffs in a neighborhood of the given payoffs. To this end, for any finite game form G^n call a payoff $(u^n, v^n) \in \mathbb{R}^{dimG^n}$ very-very nice if there is a neighborhood of it such that for all payoffs in the neighborhood the corresponding game has a stable outcome and the payoff-stable outcome correspondence $STO : \mathbb{R}^{dimG^n} \rightrightarrows \Delta(T \times M^n \times A^n)$, is continuous in that neighborhood.

Lemma 1. *The set of not very-very nice payoffs has Lebesgue measure 0.*

Proof of Lemma 1. It follows from the fact that the graph of the payoff-stable outcome correspondence can be defined by a first-order formula, and hence by the Tarski-Seidenberg Theorem it is semi-algebraic.[23] Then the statement follows from the Lemma in Blume and Zame (1994). □

Proposition 3. *For any multi-sender signaling game there exists a stable* outcome.*

Proof of Proposition 3. Pick any finite approximation, choose $y^n = (u^n, v^n)$ which are very-very nice from $B(x^n, \epsilon^n)$ (this can be done by Lemma 1), then there exists a sequence of stable outcomes λ'^n. Pick a converging subsequence, which exists by compactness of $\Delta(T \times M \times A)$ and consider the corresponding subsequence $(G^n)_{n \in \mathbb{N}}$ as a finite approximation. Fix and denote the limit of the convergent subsequence by λ. Since y^n is very-very nice, there exists ϵ'^n such that for any $(u^n, v^n) \in D^n = B(y^n, \epsilon'^n)$ one can find a stable outcome λ^n arbitrarily close (in the l_1 metric) to the stable outcome λ'^n, hence also weakly converging to λ. □

[23] Let us demonstrate this fact by using the notation of Blume and Zame (1994). For a multi-sender signaling game we can define the graph of STO by:

$$Graph(STO) = \{(u, \lambda) \in U \times \Delta(T \times M^n \times A^n) : \forall \varepsilon > 0 \forall \delta > 0 \forall \eta \in \mathbb{R}^C_{++}$$
$$\exists s \in SE(u) \wedge \lambda(z) > 0 \Leftrightarrow Pr\{z|s\} > 0 \wedge$$
$$\exists s' \in S(\eta) \wedge (u, \eta, s') \in Graph(PNE) \wedge$$
$$||\eta|| < \delta \wedge ||s - s'|| < \varepsilon\}.$$