Economic Theory
https://doi.org/10.1007/s00199-026-01723-2

RESEARCH ARTICLE

# Simple forward induction in monotonic multi-sender signaling games

**Péter Vida[1,2] · Takakazu Honryo[3] · Helmuts Āzacis[4]**

Received: 11 July 2025 / Accepted: 18 May 2026


**Abstract**
We introduce a new, powerful, and easy-to-use solution concept called simple forward induction which is implied by strategic stability in generic finite multi-sender signaling games. We apply this concept to infinite monotonic signaling games and show that a unique simple forward induction outcome exists and is non-distorted by asymmetric information. We also show that, in this class of games, the non-distorted outcome is the limit of stable outcomes of finite games. Hence, in this sense, we characterize stable outcomes in this class of games. The definition of our solution concept can be easily extended to arbitrary extensive-form games.

**Keywords** Multi-sender signaling · Forward induction · Strategic stability · Monotonic games

**JEL Classification** 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. We thank the associate editor, two anonymous referees, Malin Arve, Itzhak Gilboa, Srihari Govindan, Faruk Gul, Johannes Hörner, Navin Kartik, Donald Keenan, Sina Litterscheid, Pauli Murto, Kathleen Nosal, Luciano Pomatto, Dov Samet, Larry Samuelson, Joel Sobel, and Thomas Tröger for their helpful comments. All remaining errors are ours.

✉ Helmuts Āzacis
helmuts.azacis@uni-corvinus.hu

Péter Vida
vidapet@gmail.com

Takakazu Honryo
thonryo@mail.doshisha.ac.jp

1 CY Cergy Paris Université, THEMA, 33 boulevard du Port, 95011 Cergy-Pontoise Cedex, France

2 Corvinus University of Budapest, CCOR, CIAS, IODS, Budapest, Hungary

3 Doshisha University, Department of Economics, Kyoto, Japan

4 Corvinus University of Budapest, Institute of Economics, Budapest, Hungary

Published online: 10 June 2026

## 1 Introduction

Consider a finite extensive-form game with perfect recall, a (Nash) equilibrium outcome of this game, and the whole set of equilibria inducing this outcome, i.e., the equilibria that only differ off the equilibrium path. Forward induction restricts players' beliefs (and, hence, their sequentially rational actions) at off-the-path information sets to decision nodes that can be reached by other players' strategies that themselves are best replies to some equilibrium in that set (otherwise the deviator(s) would be strictly better off by not deviating in that way); see Govindan and Wilson (2009, page 1). If an equilibrium outcome cannot be supported by such beliefs, we say that it fails the test of forward induction and is rejected. Forward induction allows for decision nodes that can be reached by multilateral deviations. Therefore, testing forward induction is difficult in general because one has to take into consideration multilateral deviations before rejecting some equilibrium outcome. In this paper, we introduce the notion of *simple forward induction* which only requires to consider unilateral deviations. Namely, if the equilibrium outcome cannot be supported by beliefs restricted to decision nodes that can be reached by a unilateral deviation that is a best reply to some equilibrium in the set of equilibria inducing the given outcome, then we say that the outcome fails the test of simple forward induction and is rejected. It follows that more outcomes are rejected by simple forward induction than by forward induction, and the test also becomes simpler.

Forward induction is implied by strategic stability (see Proposition 6B in Kohlberg and Mertens (1986)), and because for generic games, a stable outcome exists, an outcome that passes the test of forward induction, also exists. However, it is unclear whether there always exists an equilibrium outcome that passes the test of simple forward induction. For this reason, we restrict attention to multi-sender signaling games, and prove that in this class of games, strategic stability indeed implies simple forward induction (Proposition 1).[1] Then, in the spirit of Cho and Sobel (1990), we apply our solution concept to infinite monotonic multi-sender signaling games and show that it selects a unique outcome (Theorem 1 and Corollary 1) which also turns out to be the limit of stable outcomes of approximating finite games (Proposition 2).[2] We also discuss how to extend the definition of our solution concept to arbitrary extensive-form games with perfect recall.

Banks and Sobel (1987), Cho and Kreps (1987), and Cho and Sobel (1990) successfully rule out unintuitive equilibria in a single-sender setting by applying solution concepts that restrict the support of the receiver's beliefs. These concepts, in the order of increasing strength, are: the intuitive, D1, D2, Universal Divinity, and Never Weak Best Response (NWBR) criteria. All of them can be thought of as some weak form of forward induction (as defined by Kohlberg and Mertens (1986) and Govindan and Wilson (2009)). However, we show on an example in Section 2.3 that, in a multi-sender

[1] Roughly speaking, a set of Nash equilibria is strategically stable if any strategy perturbation of the game has a Nash equilibrium close to this set (see the exact definition in Appendix B, in Kohlberg and Mertens (1986), or in Mertens (1989, 1991)). It is surprising that stability implies, among other desirable properties, (simple) forward induction as well.

[2] For example, the oligopoly limit pricing model of Bagwell and Ramey (1991) belongs to this class of games.

setting, highly unintuitive equilibria may still survive forward induction. Moreover, as we demonstrate in the example, testing forward induction is rather cumbersome, whereas finding the simple forward induction outcome is relatively simple.

The main weakness of the above-mentioned solution concepts, when applied in the multi-sender setting, is that when one of the senders was clearly deviating, they do not take into account the information conveyed by the equilibrium signals of the other senders who were (possibly) not deviating. As mentioned before, these concepts entertain the possibility that even the senders who have sent an equilibrium signal, are also deviating, i.e., that there are multilateral deviations.

As an example, think of a pair of twins who always have the same mood which can be either good or bad. When they have a good mood, they wear their red hoodies, and when they have a bad mood, they wear their blue hoodies. One day, their mother sees that one of them is wearing a green hoodie and the other is wearing a blue one. Forward induction, in principle, allows their mother to believe that the twins have a good mood, i.e., that they both are deviating. Clearly, this issue does not arise in the single-sender setting.

Restricting the support of the receiver's beliefs to unilateral deviations yields stronger predictions (see our example in Section 2.3). Such a belief restriction, in the context of multi-sender signaling games, is called unprejudiced beliefs, and it was introduced in Bagwell and Ramey (1991) and further analyzed in Vida and Honryo (2021) who showed that unprejudiced beliefs are implied by strategic stability. In the above example, if the mother's belief is unprejudiced, then she must believe that the twins have a bad mood and that only the one wearing the green hoodie is deviating.

An equilibrium outcome which can be supported by unprejudiced beliefs and which satisfies forward induction at the same time is what we call a simple forward induction outcome (henceforth: SFI) and we show in Proposition 1 that such an outcome generically exists. Testing SFI is much simpler than testing forward induction because one has to test only whether unprejudiced beliefs satisfy forward induction.

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 (Corollary 1). Our monotonicity assumptions are natural and exact parallels of those in Cho and Sobel (1990) but tailored to the multi-sender setting. In our games, the senders are assumed to have common knowledge about the state, but this information is unknown to the receiver. This assumption about the information structure is frequent in the literature with multiple senders, including a large part of the implementation literature (in particular, when the designer is a player), mechanism design, social choice, and many other settings in finance, industrial organization, law and economics, and political economy.[3] In particular, our paper extends the setup (and the results) of Bagwell and Ramey (1991) to the case where the senders have more than two types. We hasten to say that the assumption of perfectly correlated types is only used to prove our main theorem. Defining the monotonicity properties would be hard if not impossible with imperfectly correlated types. The definition of

[3] Just to mention a few, see for example, Au and Kawai (2021), Bagwell and Ramey (1991), Baliga et al. (1997), Battaglini (2002), Bester and Demuth (2015), Chen et al. (2023), Emons and Fluet (2009), Hartman-Glaser and Hébert (2019), Hong (2024), Laslier and Van der Straeten (2004), Schultz (1996, 1999), Vaccari (2023), Qiu and Rao (2023), Zhang (2020).

SFI, however, can be easily adjusted to the case of imperfectly correlated types or to arbitrary extensive-form games.[4]

To simplify exposition, we initially prove Theorem 1 using a stronger solution concept that we call open-minded SFI, after Bagwell and Ramey (1991), or OSFI for short. Both SFI and OSFI restrict out-of-equilibrium beliefs to decision nodes that can be reached by a unilateral deviation. However, while SFI requires that the deviation is a best reply to some equilibrium in the set of equilibria inducing the given outcome, OSFI puts a small probability on deviations that are not best responses. Theorem 1 shows that there is a unique OSFI outcome in infinite monotonic multi-sender signaling games. The theorem is proven by making additional assumptions directly about the equilibria of the complete information version of the infinite monotonic games where the receiver also knows the state. Later we provide assumptions just about the incomplete information games and establish the result of Corollary 1 with SFI as the solution concept.

The SFI outcome in Corollary 1 (or, equivalently, OSFI outcome in Theorem 1) is fully separating and, unlike the single-sender case, the equilibrium signals are non-distorted by the ex ante asymmetric information, i.e., it is as if the senders and the receiver are playing subgame perfect equilibria 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. Hence, similar to Proposition 3.2 in Cho and Sobel (1990), we have also shown that, in our class of games, pure SFI outcomes are (finitistically) stable outcomes. Essentially, we characterize stable outcomes in this class of games using two restrictions on beliefs: forward induction and unprejudiced beliefs.

The paper is structured as follows. In Section 2, we set up the basic game, define our solution concepts, and provide an elaborate example. In Section 3, we define our class of infinite monotonic games, the notion of non-distorted outcome, and state and prove Theorem 1. Section 4 focuses on the following things. First, we discuss the role of some of our assumptions and we state Corollary 1 here. Next, we provide a thorough comparison with the results in Cho and Sobel (1990). In particular, similar to Cho and Sobel (1990), we show that D1 and unprejudiced beliefs together are equivalent to SFI in our class of games. Propositions 1 and 2 are also stated here. In Section 5, we explain how to extend the definition of our solution concept to arbitrary extensive-form games with perfect recall. In Section 6, we conclude. Some of the definitions and proofs are relegated to the Appendix.

## 2 The Model

We start by defining 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 of senders is denoted by $S$ with $|S| > 1$. There is also a single receiver (she). A generic sender (he) is denoted by $i \in S$ and the other senders by $-i$. At the beginning of

[4] In fact, when the type distribution has full support, SFI is equivalent to forward induction; see Proposition 1 in Vida and Honryo (2021) for a precise statement.

the game, senders learn their common type which is unknown to the receiver. Namely, senders' types $t \in \{0, \dots, T\} = \bar{T}$ are perfectly correlated and are drawn according to some probability distribution $\pi \in \Delta\bar{T}$, where $\pi$ is common knowledge among the players and $\pi(t) > 0$ is the probability of $t$ for all $t \in \bar{T}$. We denote sender $i$ of type $t$ 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.[5] 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 $\pi$, induces a probability distribution over the terminal nodes of the game, which we identify with $\bar{T} \times M \times A$. This probability distribution over $\bar{T} \times M \times A$ 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 (assessment) about the types of the senders after signal profiles is a collection of probability distributions $\mu = (\mu_m)_{m \in M}$, where $\mu_m \in \Delta\bar{T}$ for all $m$. 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 $\mu$ and we denote by $e(\mu)(\cdot)$ the corresponding strategy. That is,

$$e(\mu)(m) := \arg\max_{a \in A} \sum_{t=0}^{T} v(t, m, a)\mu_m(t).$$

A perfect Bayesian equilibrium (PBE) is a triple of strategies and assessment $(m(\cdot), e(\cdot), \mu)$ that satisfies:
*Sequential rationality for the receiver:* $e(\cdot) = e(\mu)(\cdot)$,
*Sequential rationality for the senders:* for all $i$ and $t$:

$$m_i(t) \in \arg\max_{m_i} u_i(t, m_i, m_{-i}(t), e(m_i, m_{-i}(t))),$$

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

Throughout, unless otherwise stated, by equilibrium, we mean a PBE.

[5] We allow for mixed strategies in section 4.2.3 when considering finite games. We could allow for mixed strategies from the outset; however, in that case, we would need a stronger (and much more complicated) version of the single crossing condition (see A4 below) for our theorem to hold.

### 2.2 The Solution Concept: Simple Forward Induction

Fix a PBE $(m(\cdot), e(\cdot), \mu)$ and let $u_i^*(t) := u_i(t, m(t), e(m(t)))$ denote the equilibrium payoff of $(i, t)$ in this PBE. We will impose restrictions on $\mu$ for certain out-of-equilibrium signal profiles that can be reached by unilateral deviations. We focus on signal profiles at which the receiver can identify a sender about whom she knows *for sure* that this sender has deviated. Formally, let $m$ be such that there is an $(i, t)$ such that $m_{-i}(t) = m_{-i}$ and there is no $t'$ such that $m_i(t') = m_i$, and denote the set of all such signal profiles by $M(m(\cdot))$. We now introduce a list of restrictions on $\mu_m$ that we require to hold for all $m \in M(m(\cdot))$. Fix an $m \in M(m(\cdot))$ and call the corresponding $i$ the identified deviator.[6]

*The Definition of Unprejudiced Beliefs.* While, in principle, $m$ can also result from multilateral deviations, necessarily including that of the identified deviator $i$, we are going to require that the receiver thinks that sender $i$ has deviated *unilaterally* while senders $-i$ are following their equilibrium strategies. The receiver, however, might still be uncertain about the type of the identified deviator even if she thinks that a unilateral deviation has occurred. Hence, let us denote by $T_m = \{t | m_{-i}(t) = m_{-i}\}$ the set of possible types of $i$ who could reach $m$ by unilateral deviation. (We do not index $T_m$ by $i$ because sender $i$ is uniquely identified by $m$ given $m \in M(m(\cdot))$.) Our first restriction is that $\mu_m \in \Delta T_m$. Using the terminology of Bagwell and Ramey (1991), we refer to this property of beliefs as *unprejudiced beliefs*. The requirement that the belief is supported on $T_m$ implies that the receiver believes in and uses the information provided by senders $-i$. That is, she believes that senders $-i$ did not deviate and signal their types according to the equilibrium. Given that the types are perfectly correlated, this is also a signal about the deviator's possible types.

We further require that the belief must survive *forward induction* (FI) and, hence, must be concentrated on types of the identified deviator $i$ for which the deviation is a weak best response (Cho and Kreps, 1987). We note that while D1 or – with a slight qualification – the intuitive criterion would also suffice for our result to hold, we have chosen to work with the weak best response property because it has the simplest definition. Furthermore, in general, forward induction allows for weakly best responding types outside of $T_m$. However, given that unprejudiced beliefs already rule out the types outside of $T_m$, we only have to consider the incentives of sender $i$ of type $t \in T_m$ and, hence, our definition simplifies further. We provide a formal definition of forward induction and related concepts in Section 4.2.2 where one can immediately see how unprejudiced beliefs help to simplify the definition of forward induction.

[6] Given that our game is "short", we are only interested in out-of-equilibrium signal profiles or information sets that can be reached by a unilateral deviation. For simplicity, we restrict beliefs only at information sets belonging to $M(m(\cdot))$. However, all of our results extend if one also imposes similar restrictions on other out-of-equilibrium information sets. Namely, on those information sets where the receiver cannot identify a sender about whom she knows for sure that this sender has deviated, i.e., when all the signals are equilibrium signals but belong to different types.

We say that for $(i, t)$, the signal $m_i$ is a *weak best response* given that senders $-i$ send the signal profile $m_{-i}$ if there is an $a \in A$ such that $u_i(t, m, a) = u_i^*(t)$ and, for all $t' \in T_m$, we have that $u_i(t', m, a) \leq u_i^*(t')$. [7]

If $m_i$ is a weak best response for $(i, t)$ given that senders $-i$ send the signal profile $m_{-i}$, then there is an action $a$ of the receiver after the signal profile $m$ that maintains the equilibrium and $(i, t)$ is just indifferent between $m_i$ and his equilibrium signal. Namely, given $a$, no type of sender $i$ can profitably deviate to the signal $m_i$ while type $t$ gets exactly his equilibrium payoff $u_i^*(t)$ by deviating to $m_i$.

Let $F_m \subseteq T_m$ be the set of types $t$ for which $m_i$ is a weak best response for $(i, t)$ given that senders $-i$ send the equilibrium signal profile $m_{-i}(t) = m_{-i}$. We also say that $T_m \setminus F_m$ is the set of types of $i$ for whom sending the signal $m_i$ is *never a weak best response* (NWBR) given that senders $-i$ send $m_{-i}(t) = m_{-i}$. With this notation, we can finally define simple forward induction.

**Definition 1** A PBE $(m(\cdot), e(\cdot), \mu)$ satisfies *simple forward induction* (SFI) if for all $m \in M(m(\cdot))$, $\mu_m \in \Delta F_m$ if $F_m \neq \emptyset$ and $\mu_m \in \Delta T_m$ otherwise.

If $(m(\cdot), e(\cdot), \mu)$ satisfies SFI, then we say that the equilibrium outcome $[m(\cdot), e(\cdot)]$ survives SFI or that it is an SFI outcome. To emphasize, forward induction is simple once it is coupled with the requirement of unprejudiced beliefs. We illustrate this and other differences between forward induction and simple forward induction on the example in the next section and also in Section 4.2.2.

To simplify the proof of our main result, we also introduce a third requirement which we call *open-mindedness* (after Bagwell and Ramey (1991)). Later, in Section 4.1, we show how to dispense with it. Thus, we require that the receiver puts yet a small weight on some types in $T_m \setminus F_m$ if this set is not empty. Formally, fix an $\varepsilon \geq 0$. We say that $(m(\cdot), e(\cdot), \mu)$ satisfies $\varepsilon-$*simple forward induction* if for all $m \in M(m(\cdot))$, we have the following:

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

Note that 0-simple forward induction is equivalent to simple forward induction. Points (1) and (2) require that $\mu_m$ is concentrated on $T_m$, that is, the receiver must believe that the identified deviator was deviating unilaterally (unprejudiced beliefs). Point (2) additionally requires that if possible, the belief puts a total weight of $1 - \varepsilon$ on those types of the identified deviator $i$ for whom it is a weak best response to send the signal $m_i$ given that others send $m_{-i}$ (forward induction) and puts a total weight of $\varepsilon$ on those types of $i$ in $T_m$ for whom it is never a weak best response to send the signal $m_i$ (open-mindedness).

**Definition 2** A PBE outcome $[m(\cdot), e(\cdot)]$ satisfies *open-minded simple forward induction* (OSFI) if there is an $\bar{\varepsilon} > 0$ such that for all $\varepsilon \in [0, \bar{\varepsilon})$, there is an assessment $\mu^\varepsilon$ such that $(m(\cdot), e(\mu^\varepsilon)(\cdot), \mu^\varepsilon)$ satisfies $\varepsilon-$simple forward induction.

[7] Cho and Kreps (1987) require that $a \in A$ is a best response of the receiver given some beliefs and the signal $m$. This would make the definition of weak best response only stronger. For us, the weaker definition already suffices and it also keeps the notation simpler.

If a PBE outcome satisfies OSFI, then we say that it is an OSFI outcome.

Given that $m(\cdot)$ is fixed along the sequence as $\varepsilon$ 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 the Bayes rule for all $t \in \bar{T}$. Therefore, $e(\mu^{\varepsilon})(m(t)) = e(m(t))$ and $[m(\cdot), e(\cdot)] = [m(\cdot), e(\mu^{\varepsilon})(\cdot)]$ are also fixed, i.e., the outcome is fixed along the sequence.

Clearly, any OSFI outcome is also an SFI outcome. Moreover, in finite games, open-mindedness plays almost no role because once we further restrict PBE to perfect equilibria, OSFI outcomes coincide with SFI outcomes. We provide a comparison of OSFI and SFI in Section 4.3.

### 2.3 An Example

In this example, we show that there is an implausible equilibrium outcome that survives forward induction (in the sense of Kohlberg and Mertens (1986) or Govindan and Wilson (2009)) but does not survive simple forward induction. We also give the unique (O)SFI outcome of this game.

As in Bagwell and Ramey (1991), two incumbent firms signal the industry cost to a potential entrant via prices. The entrant, having observed the prices, decides how much effort to exert when entering the market. Formally, two incumbent firms learn their common type $t \in \{0, 1\}$. Next, each firm $i$, $i = 1, 2$, chooses a price $m_i \in [0, \overline{m}_i]$ and receives profit according to $(m_i - t)(d - m_i + m_{-i})$. We call this part of the game the profit game. 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 payoff is described by $v(t, a) = -(t - 1 + a)^2$. Firm $i$'s overall payoff is given 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 that is unknown by the entrant but is commonly known by the firms. Firms set prices, which serve as signals, and receive the corresponding profits in the profit game (the first term of $u_i(t, m, a)$). After observing the prices, the receiver, who is the entrant, chooses an effort level of entry $a$. Firms like small effort levels (the second term of $u_i(t, m, a)$), while the entrant chooses higher effort when her belief about the industry cost is lower. Let us fix $d = 2$ and $k = -2$.

We start by making few observations. First, a PBE is said to be separating if the incumbents' two types choose different price vectors: $m(0) \neq m(1)$. Hence, by observing the equilibrium prices in a separating PBE, the entrant can infer the incumbents' type for sure: $\mu_{m(0)}(0) = 1$ and $\mu_{m(1)}(0) = 0$. Second, given the entrant's beliefs $\mu$ and the incumbents' prices $m$, the sequentially rational action of the entrant is $e(\mu)(m) = \mu_m(0)$, that is, the optimal effort by the entrant is equal to the probability that she assigns to the incumbent firms of being type 0. Consequently, in any separating PBE, $e(\mu)(m(0)) = 1$ and $e(\mu)(m(1)) = 0$. Third, the best response function of incumbent $i$ in the profit game is $m_i = b_i(t)(m_{-i}) = (m_{-i} + 2 + t)/2$. Hence, the Nash equilibrium of the profit game is (2, 2) when $t = 0$ and (3, 3) when $t = 1$ (see Figure 1). We will refer to $m(0) = (2, 2)$ and $m(1) = (3, 3)$ as non-distorted (price) signals because the firms play as if the entrant already knew their types. Finally, in any separating PBE, the firms must choose non-distorted signals when their type is 0, $m(0) = (2, 2)$. Given $m(0)$, the entrant chooses the maximal effort. Therefore, if $m(0)$ was different from $(2, 2)$, some firm would want to deviate and its deviation could not be deterred with the threat of higher effort by the entrant. Given $m(0) = (2, 2)$ and $e(2, 2) = 1$, the equilibrium payoffs of type 0 firms are $(2, 2)$.

Next we argue that there are multiple separating PBE outcomes that differ in $m(1)$. First of all, there is a separating equilibrium outcome in non-distorted signals $m(1) = (3, 3)$ (together with $m(0) = (2, 2)$, $e(2, 2) = 1$, and $e(3, 3) = 0$). This outcome can be supported with the following beliefs: $\mu_m(0) = 0$ if $m_i = 3$ and $m_{-i} \neq 2$ for some $i$, and $\mu_m(0) = 1$ otherwise. Note that given these beliefs, if $m_{-i} = 2$, then firm $i$ cannot induce the entrant to choose lower effort through a unilateral deviation. Therefore, firm $i$ does not have incentives to deviate from $m_i(0) = 2$ in state 0. It does not want to deviate from $m_i(1) = 3$ in state 1 either because that can only lead to higher effort by the entrant.

This PBE outcome also satisfies (O)SFI. For any $m = (m_i, 2) \in M(m(\cdot))$ (i.e., $m_i \neq 2$ or $m_i \neq 3$), we have $T_m = \{0\}$ and $\mu_m(0) = 1$ by the equilibrium construction. Likewise, for any $m = (m_i, 3) \in M(m(\cdot))$, we have $T_m = \{1\}$ and $\mu_m(1) = 1$. Hence, the beliefs are unprejudiced because $\mu_m \in \Delta T_m$ holds for any $m \in M(m(\cdot))$. In other words, after $m \in M(m(\cdot))$, the entrant assumes that incumbent $-i$ has not deviated and, therefore, forms beliefs based on the price signal of incumbent $-i$. Both simple forward induction and open-mindedness are also trivially satisfied because $T_m$ is a singleton for any $m \in M(m(\cdot))$.

There are also separating PBE outcomes in which type 1 firms send distorted price signals. For example, $m(1) = (2.8, 2.9)$ (together with $m(0) = (2, 2)$, $e(2, 2) = 1$, and $e(2.8, 2.9) = 0$). As seen in Figure 1, while firm 2 of type 1 is best responding to firm 1's price, $m_2(1) = (2.8 + 2 + 1)/2 = 2.9$, firm 1 is not, $m_1(1) = 2.8 \neq (m_2(1) + 2 + 1)/2$. Firm 1 could increase its profit by choosing $m_1 \in (2.8, 3.1)$. To deter any such deviation, we can set $\mu_{(m_1,2.9)}(0)$ to be sufficiently high (say, equal to 1). This will induce high effort from the entrant, which will cancel any gains in profit of firm 1 from sending $m_1$.

A positive probability $\mu_{(m_1,2.9)}(0) > 0$ means that, after observing the signal profile $(m_1, 2.9)$, the entrant believes that both type 0 firms could have deviated simultaneously from $m(0) = (2, 2)$, i.e., a multilateral deviation has occurred. We now argue that such a belief cannot be ruled out by FI alone, that is, without additionally imposing unprejudiced beliefs. Specifically, we show that it is a weak best response for firm 1 of type 0 to send any signal $m_1 \in (2.8, 3.1)$ *given* that firm 2 of type 0 sends signal $m_2(0) = 2$. It is also a weak best response for firm 2 of type 0 to send signal 2.9 *given* that firm 1 of type 0 sends signal $m_1(0) = 2$. Let $\mu_{(2,2.9)}(0) = e(2, 2.9) = 0.595$ and $\mu_{(m_1,2)}(0) = e(m_1, 2) = m_1(4 - m_1)/2 - 1 \in [0, 1]$ for all $m_1 \in (2.8, 3.1)$. Then, the payoff of both type 0 firms is $u_1(0, (m_1, 2), m_1(4 - m_1)/2 - 1) = u_2(0, (2, 2.9), 0.595) = 2$, which is exactly their equilibrium payoff. Simple calculation also verifies that 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 signal pairs, the beliefs can also be chosen to satisfy forward induction. It follows that the

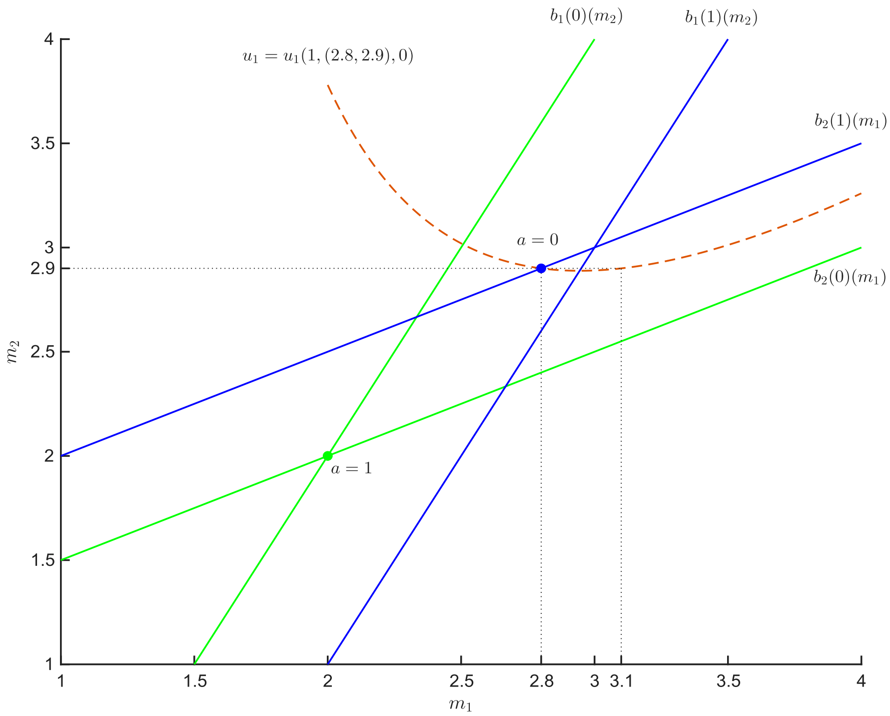


**Fig. 1** PBE that survives FI, but not SFI.

separating PBE in which type 1 firms send distorted signals $m(1) = (2.8, 2.9)$ cannot be eliminated by FI.

Unprejudiced beliefs instead require that after observing the signal profile $(m_1, 2.9)$, the entrant believes that only firm 1 has deviated, which means that the entrant can infer from the signal of firm 2 that the state is 1 and, correspondingly, chooses zero effort: $\mu_{(m_1,2.9)}(0) = e(m_1, 2.9) = 0$. But then, the deviation from $m_1(1) = 2.8$ to $m_1$ cannot be deterred. Consequently, the separating PBE outcome in which type 1 firms send distorted signals $m(1) = (2.8, 2.9)$ does not survive SFI. In fact, Theorem 1 and Corollary 1 will tell us later that the unique PBE outcome, among both pooling and separating equilibria, that survives (O)SFI in this example, is the one in which the firms send non-distorted signals.[8]

[8] In the above example, unprejudiced beliefs alone are sufficient to eliminate the separating outcome that survives forward induction. However, in the case of pooling outcomes, unprejudiced beliefs have no bite because $T_m = \{0, 1\}$ holds for any $m \in M(m(\cdot))$ and, consequently, only forward induction can eliminate them. Importantly, there are also equilibria that can only be eliminated by the joint power of unprejudiced beliefs and forward induction. Such an example is discussed in detail in Vida and Honryo (2021) but is also briefly mentioned in Section 4.2.2 below.

## 3 Monotonic Signaling Games and the Theorem

We consider a class of infinite multi-sender signaling games satisfying certain monotonicity conditions. This class is a natural extension of the monotonic single-sender games considered in Cho and Sobel (1990). We show that, in these games, the unique equilibrium outcome that survives OSFI, is the one in which all types separate and the equilibrium signals are non-distorted. Later we show in Section 4.2.3 that the selected outcome is also the limit of stable outcomes (Kohlberg and Mertens 1986) of approximating finite games.

For simplicity, we assume that the signal spaces are open intervals.[9] We also assume the following:

**A0** $u_i(t, m, a)$ and $v(t, m, a)$ are continuous in $(m, a)$ for all $i, t$.
**A2** For all $a', a : a' > a$ implies that $u_i(t, m, a') > u_i(t, m, a)$ for all $i, t, m$.
**A3** $\partial v(t, m, a)/\partial a$ is strictly increasing in $t$ for all $t, m, a$; $\arg\max_{a \in A} v(T, m, a) < \overline{a}$ for all $m$.
**A4** For all $i, m_i, m_i', m_{-i}, a, a', t, t'$ such that $m_i < m_i'$ and $t < t'$:

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

Assumptions A0-A4 are natural multi-sender counterparts of assumptions A0-A4 in Cho and Sobel (1990) and they play the same roles.[10]

We now define the notion of non-distorted outcome in the incomplete-information game. Consider degenerate incomplete-information games indexed by $t$, where the priors are fully concentrated on $t$ and the players, the payoffs, the signal sets, and the action set are exactly the same as in the original game. Namely, for every $t$, let us assume that the receiver knows that the type of the senders is $t$ and consider these as separate complete-information games as $t$ varies. For all $t \in \bar{T}$, let us denote by $N(t) \subseteq M \times A$ the set of pure subgame perfect equilibrium outcomes of these complete-information games. If we write $e(t)(m) = \arg\max_{a \in A} v(t, m, a)$, then $(m, a) \in N(t)$ if and only if $a = e(t)(m)$ and $m_i \in \operatorname{argmax}_{m_i'} u_i(t, (m_i', m_{-i}), e(t)(m_i', m_{-i}))$ for all $i \in S$. Clearly, $(m, e(t)(\cdot))$ is the corresponding subgame perfect equilibrium.

9 This simplifying assumption is innocuous in the presence of our other assumptions that guarantee the existence of equilibrium even with open signal spaces. Its role is to ensure that we can always consider both upward and downward deviations. In Section 4.1, we assume that the signal spaces are compact intervals.

10 A0-A1 guarantees that $e(\mu)(\cdot)$ is a function, which is continuous in $\mu$ and $m$. A2 and A3 are the reasons why we can 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. Alternatively, we could have assumed that all senders prefer lower actions as in the example of Section 2.3. Even more, all of our results hold with senders having monotonic preferences in different directions as long as the outcome with non-distorted signals is a PBE outcome. A3, together with A1, implies that $e(\mu)(m)$ is strictly increasing in $\mu_m$ in the sense of first-order stochastic dominance. Instead, we could have assumed that $\partial v/\partial a$ is strictly decreasing in $t$ as in the example of Section 2.3. 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 higher than the other, then all higher types of sender $i$ strictly prefer to send the higher 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 could also hold in the other direction, i.e., with $t > t'$, or even in different directions for different senders.

Consider now again the original incomplete-information game. We say:

**Definition 3** An outcome $[m(\cdot), e(\cdot)]$ is *non-distorted* if $(m(t), e(m(t))) \in N(t)$ for all $t \in \bar{T}$.

We refer to $m(t)$ for all $t \in \bar{T}$ in the above definition as non-distorted signals.

Additionally, we make the following two assumptions about the associated complete-information games:

**B1** $|N(t)| = 1$ for all $t$.
**B2** For all $t, t' \in \bar{T}$ such that $t \neq t' : \exists (m, a) \in N(t), (m', a') \in N(t'), i, j \in S : i \neq j, m_i \neq m'_i, m_j \neq m'_j$.

For instance, in the example of Section 2.3, the set of subgame perfect equilibrium outcomes is $N(0) = \{((2, 2), 1)\}$ if, in addition to the incumbents, the entrant also knows the state is 0, and it is $N(1) = \{((3, 3), 0)\}$ if the entrant knows the state is 1; thus, B1 and B2 are both satisfied. It can also be seen in Figure 1 where the signal profiles in $N(0)$ and $N(1)$ correspond to the unique intersection of the two green and the two blue lines, respectively. Clearly, neither of the coordinates of these two intersections coincide, meaning that the two types are separated by both incumbents.

These assumptions help us to simplify the exposition. In Section 4.1, we give conditions on the primitives of the game that imply B1 and B2. B1 ensures that a non-distorted outcome, if it exists, is unique. In the absence of B1, the uniqueness of OSFI outcome in Theorem 1 below does not hold. However, the possible multiplicity of OSFI outcomes would be due to the multiplicity of subgame perfect equilibria in the complete information games which is orthogonal to the well-known multiplicity arising from out-of-equilibrium beliefs in incomplete information games that we seek to address.

Note, however, B1 does not guarantee that the non-distorted outcome necessarily exists because there can be $t, t'$ such that the non-distorted signals fully overlap, $m(t) = m(t')$, but $e(t)(m(t)) \neq e(t')(m(t'))$. For example, the non-distorted outcome of the Spence (1973) model would require that both low and high productivity workers acquire no education, but high productivity workers are paid more, which is impossible under incomplete information. Said simply, this is not a feasible outcome in the incomplete information game. B2 ensures the existence of non-distorted outcome because each pair of types can be separated by non-distorted signals. Furthermore, for any pair of types, these non-distorted signals differ for at least two senders. This is important because we prove in the theorem that under A0-A4, any OSFI outcome is non-distorted and each pair of types must be separated by at least two senders. Thus, B2 together with A0-A4 also guarantees the existence of OSFI outcome. In Section 4.3, by using the Spence (1973) model, we explain that when assumption B2 fails, we can still restore the existence by considering the limit of OSFI outcomes of finite games that approximate the infinite game.

**Theorem 1** *Under assumptions A0-A4, an outcome is an OSFI outcome if and only if it is non-distorted and each pair of types is separated by at least two senders. If B1 and B2 also hold, then there exists a unique OSFI outcome which is the non-distorted outcome.*

We show later in Corollary 1 that with additional assumptions, Theorem 1 also holds for SFI, i.e., we can drop the requirement of open-mindedness on the receiver's beliefs. Those assumptions will also imply B1 and B2.

***Proof of Theorem 1:*** We prove the theorem in 4 steps.

**(1) All types separate (by forward induction and unprejudiced beliefs):** The proof is by contradiction. Suppose there is an (O)SFI outcome where some subset of types $T'$ pools by sending the same signal profile. Let $t'$ denote the highest type in $T'$. Thus, all types $t \in T'$ send $m(t')$. Now, take any sender $i$ of any type in $T'$. Given that the signal spaces are open, one can always find $m_i$ such that $m_i > m_i(t')$, $m_i$ is sufficiently close to $m_i(t')$, and there is no $t \in \bar{T}$ for which $m_i(t) = m_i$, i.e., $m = (m_i, m_{-i}(t')) \in M(m(\cdot))$. Note that $T' \subseteq T_m$ holds by unprejudiced beliefs. Let $t''$ be the highest type in $T_m$ such that $u_i^*(t'') = u_i(t'', m(t'), e(m(t')))$ holds. Clearly, $t'' \geq t'$. If $t'' > t'$, then sender $i$ of type $t''$ is indifferent between sending his equilibrium signal $m_i(t'')$ and "pooling" signal $m_i(t')$. Also note that there can be $t \in T_m$ such that $t > t''$ but then $u_i^*(t) > u_i(t, m(t'), e(m(t')))$.

First we claim that no type in $T'' = \{t \in T_m | t < t''\}$ belongs to $F_m$. Take any $t \in T''$. If there is no $a \in A$ such that $u_i(t, m, a) = u_i^*(t)$, then it is immediate that $t \notin F_m$. Therefore, suppose there is $a \in A$ such that $u_i(t, m, a) = u_i^*(t)$. Because $[m(\cdot), e(\cdot)]$ is a PBE and because $m_{-i}(t) = m_{-i}(t')$, $u_i^*(t) \geq u_i(t, m(t'), e(m(t')))$ (and with equality if $t \in T'$). Hence, $u_i(t, m, a) \geq u_i(t, m(t'), e(m(t')))$ holds. By A4, $u_i(t'', m, a) > u_i(t'', m(t'), e(m(t'))) = u_i^*(t'')$, implying that $m_i$ is never a weak best response for $t$ and so $t \notin F_m$.

Next we argue that $t'' \in F_m$. By definition of $t''$, $u_i^*(t) > u_i(t, m(t'), e(m(t')))$ for all $t \in (T_m \backslash T'') \backslash \{t''\}$. If $m_i$ is close enough to $m_i(t')$, then by A0, A2, and A3, one can find $a \in A$ such that $u_i(t'', m, a) = u_i^*(t'')$ and $u_i(t, m, a) < u_i^*(t)$ for all $t \in (T_m \backslash T'') \backslash \{t''\}$. By A4, $u_i(t, m, a) < u_i^*(t)$ also holds for all $t \in T''$. Hence, $t'' \in F_m$.

Because $F_m \subseteq T_m \backslash T''$, by unprejudiced beliefs and forward induction, the belief at $m$ first-order stochastically dominates the belief at $m(t')$. (Note that by the Bayes rule, $\mu_{m(t')}(t) \in (0, 1)$ for all $t \in T'$.) By A0, A1, and A3, this induces an upward jump in the receiver's action, resulting in a profitable deviation for any type $t \in T'$ of sender $i$ provided that $m_i$ is close enough to $m_i(t')$.

**(2) Any pair of types are separated by at least two senders (by open-mindedness and unprejudiced beliefs):** This is the only place where we need open-mindedness. By contradiction, suppose there are types $t, t'$ for which $\exists! i$ such that $m_i(t) \neq m_i(t')$ and $m_{-i}(t) = m_{-i}(t')$. Let $t$ be the smallest type for which this is true. Fix $\varepsilon > 0$ and consider a deviation $m_i < m_i(t)$ of $(i, t)$ sufficiently close to $m_i(t)$ such that there is no type $t''$ for which $m_i(t'') = m_i$, i.e., $m = (m_i, m_{-i}(t)) \in M(m(\cdot))$, meaning that unprejudiced beliefs must apply after $m$. Such a deviation is possible because the signal spaces are open.

We claim that $F_m = \{t\}$. We first show that $t \in F_m$. By A0, A2, and A3, there is an $a$ that makes $(i, t)$ indifferent between $m_i$ and $m_i(t)$, $u_i(t, m, a) = u_i(t, m(t), e(m(t)))$. By A4, $u_i(t', m, a) < u_i(t', m(t), e(m(t)))$ for any $t' \in T_m \backslash \{t\}$. Furthermore, if $u_i(t', m, a) > u_i(t', m(t'), e(m(t')))$ was true, then together with the previous inequality, it would imply that $u_i(t', m(t'), e(m(t'))) < u_i(t', m(t), e(m(t)))$, which contradicts that $[m(\cdot), e(\cdot)]$ is a PBE outcome (because sender $(i, t')$ can unilaterally reach $m(t)$ from $m(t')$ by sending $m_i(t)$ instead of $m_i(t')$). Thus, $u_i(t', m, a) \leq u_i(t', m(t'), e(m(t')))$ for all $t' \in T_m \backslash \{t\}$. It follows that $m_i$ is a weak best response for type $t$ given $m_{-i}(t)$ and thus $t \in F_m$.

Second, if $m_i$ was a weak best response for some type $t' \in T_m \backslash \{t\}$ (where $t$ is the lowest type in $T_m$), then $u_i(t', m, a') = u_i(t', m(t'), e(m(t')))$ and $u_i(t, m, a') \leq u_i(t, m(t), e(m(t)))$ for some $a'$. But then, again by A4, $u_i(t', m(t'), e(m(t'))) = u_i(t', m, a') < u_i(t', m(t), e(m(t)))$, which contradicts that $[m(\cdot), e(\cdot)]$ is a PBE outcome. Therefore, $F_m = \{t\}$.

By open-mindedness, the receiver assigns a probability of $\varepsilon$ to types in $T_m \backslash F_m$. Because $t' > t$ for all $t' \in T_m \backslash F_m$, the receiver's belief $\mu_{m_i, m_{-i}(t)}$ first-order stochastically dominates the belief at $m(t)$ which is $\mu_{m(t)}(t) = 1$. Because $\varepsilon$ is the same for all signal profiles in $M(m(\cdot))$, the receiver's action also increases to $e(\mu)(m_i, m_{-i}(t)) > e(\mu)(m(t)) = e(t)(m(t))$ if $m_i$ is sufficiently close to $m_i(t)$. But this implies a profitable deviation for $(i, t)$, provided again that $m_i$ is sufficiently close to $m_i(t)$ (so that it is not too costly for $(i, t)$ to send $m_i$), which is a contradiction.

**(3) For all $t$ : $[m(t), e(m(t))] \in N(t)$ (non-distortion by unprejudiced beliefs):** Suppose by contradiction that there is an SFI with $u_i(t, m(t), e(m(t))) < u_i(t, m_i, m_{-i}(t), e(t)(m_i, m_{-i}(t)))$ for some $t, i, m_i$, 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 and as if she knew that the senders' type is $t$. Let $m = (m_i, m_{-i}(t))$. Note that $m_i$ can be chosen such that there is no $t'$ for which $m_i(t') = m_i$ and, hence, $m \in M(m(\cdot))$. By (2) above, every pair of types is separated by at least two senders. Therefore, $T_m = \{t\}$ and $\mu_m(t) = 1$ by unprejudiced beliefs. But then, it must be that $e(m) = e(t)(m)$ which together with the above strict inequality, implies a profitable deviation for $(i, t)$ in the incomplete information game, a contradiction.

**(4) Existence and uniqueness:** B1 and B2 ensure that the non-distorted outcome exists, is unique, and all types are separated by at least two senders. By steps (1), (2), and (3), this can be the only OSFI outcome.

We prove that under assumptions A0-A4, any non-distorted outcome in which each pair of types is separated by at least two senders, is an OSFI outcome. This will complete the proof of our theorem. It is enough to specify the receiver's beliefs and actions after the out-of-equilibrium signal profiles that can arise after unilateral deviations. For signal profiles $m \in M(m(\cdot))$, where sender $i$ is identified as a deviator and $m_{-i} = m_{-i}(t)$ for some $t$, we have that $T_m = \{t\}$ by B2. To comply with OSFI, we set $\mu_m(t) = 1$ and $e(m) = e(t)(m)$. The receiver's action deters the deviation of $(i, t)$ from $m_i(t)$ to $m_i$ because $(m(t), e(t)(\cdot))$ is the subgame perfect equilibrium of the complete information game corresponding to $t$.

Consider now any out-of-equilibrium signal profile $m \notin M(m(\cdot))$ that could be a result of a unilateral deviation. (OSFI does not restrict beliefs after such $m$.) This $m$ can be written as $m = (m_i(t'), m_{-i}(t))$ for some $i, t', t$. Furthermore, it can be that there exists $j \neq i$ such that $m$ can also be written as $m = (m_j(t), m_{-j}(t'))$, in which case, w.l.o.g., we assume that $t < t'$. We set $\mu_m(t) = 1$ and $e(m) = e(t)(m)$. The same argument as above tells that $(i, t)$ has no incentive to deviate from $m_i(t)$ to $m_i(t')$. We also need to consider the incentives of $(j, t')$ if $j$ exists.

By A0, A1, and A3, $e(t)(m) < e(t')(m)$ (because we assumed that $t < t'$). By A2, $u_j(t', m, e(t)(m)) < u_j(t', m, e(t')(m))$. In addition, because $(m(t'), e(t')(\cdot))$ is the subgame perfect equilibrium of the complete information game corresponding to $t'$, $u_j(t', m, e(t')(m)) \leq u_j(t', m(t'), e(t')(m(t')))$. Hence, $u_j(t', m, e(t)(m)) < u_j(t', m(t'), e(t')(m(t')))$, meaning that $(j, t')$ also has no incentive to deviate from $m_j(t')$ to $m_j(t)$.

## 4 Discussion

In this section, we elaborate on some of our assumptions and contrast (O)SFI with other related solution concepts. Thus, in Section 4.1, we give conditions that imply B1 and B2. These conditions also allow us to state Theorem 1 only in terms of SFI, i.e., without the requirement of open-mindedness. In Section 4.2, we compare our results to those in Cho and Sobel (1990) for single-sender signaling games. In doing so, we provide definitions of D1 and FI, which are tailored to our multi-sender setting. Here, we also connect SFI to the notion of strategic stability that has been introduced by Kohlberg and Mertens (1986). In Section 4.3, we explain with the help of an example that B2 is unnecessary for the results if we take arbitrary finite approximations of the infinite signaling games. Finally, in Section 4.4, we relate SFI to other notions of FI that can be found in the literature.

### 4.1 Assumptions Implying B1 and B2, and Dispensing with Open-mindedness

Here we give assumptions that guarantee that, in the game, there exists a unique non-distorted outcome and that all types are separated by *all* the senders. Thus, they imply B1 and B2. As a side product, in this class of games, we can also dispense with the requirement of open-mindedness; meaning, Theorem 1 holds for SFI outcomes.

Let the signal spaces be compact intervals of the form $M_i = [0, c]$. Let $c_{-i} = (c, \ldots, c) \in M_{-i}$. The assumptions beyond A0-A4 are as follows:

**A5** 1. $u_i(t, m, e(t)(m))$ is concave and differentiable in $m_i$ for all $i, m_{-i}, t$,
2. For each $t$, $\sum_{i \in S} r_i u_i(t, m, e(t)(m))$ is diagonally strictly concave in $m \in M$ for some fixed $(r_1, \ldots, r_{|S|}) > 0$.[11]

**A6** $\exists i, m_i \in M_i : u_i(T-1, m_i, c_{-i}, e(0)(m_i, c_{-i})) > u_i(T-1, c, c_{-i}, e(T)(c, c_{-i}))$.

**A7** 1. $D_{m_i} u_i(t, m, e(t)(m))|_{m_i=0} > 0$, $D_{m_i} u_i(t, m, e(t)(m)))|_{m_i=c} < 0$, for all $i, t, m_{-i}$,
2. $D_{m_i} u_i(t, m, e(t)(m)))$ is weakly increasing in $m_{-i}$ for all $i, m_i, t$,

[11] The definition is due to Rosen (1965): $\sum_{i \in S} r_i u_i(t, m, e(t)(m))$ is diagonally strictly concave in $m \in M$ for some fixed $r = (r_1, \ldots, r_{|S|}) > 0$ if for $g(m, r) = (r_i D_{m_i} u_i(t, m, e(t)(m)))_{i \in S}$ as a vector, we have that for every pair of vectors $\hat{m}, \check{m} \in M : (\hat{m} - \check{m}) \cdot g(\check{m}, r) + (\check{m} - \hat{m}) \cdot g(\hat{m}, r) > 0$, where $\cdot$ denotes the inner product. One can easily verify that the example in Section 2.3 satisfies this condition by choosing $r_i = \rho$ for all $i \in S$ and for any $\rho > 0$.

3. $D_{m_i} u_i(t, m, e(t)(m)))$ is strictly increasing in $t$ for all $i, m$.[12]

A5 and A6 can be considered the multi-sender counterparts of A5 and A6 in Cho and Sobel (1990). (See the next section for comparison with Cho and Sobel (1990).) In short, A5 implies B1, and A6 ensures that there will be no pooling on the top signal because it is too costly. A7 is the only assumption that has no counterpart in Cho and Sobel (1990), and it implies B2. A7(1) ensures that any maximization of the senders is interior. A7(2) and A7(3) are additional supermodularity conditions implying that the equilibria of the complete information games are increasing in $t$ and, together with A7(1), are strictly increasing.

**Corollary 1** *Under assumptions A0-A7, there exists a unique SFI outcome which is the unique non-distorted outcome.*

***Proof*** See the proof in Section A in the Appendix. □

### 4.2 Comparison to Cho and Sobel (1990)

#### 4.2.1 Distortion versus Non-distortion

The main difference between our setup and that of Cho and Sobel (1990) is that we have multiple senders, although their types are perfectly correlated. Both papers establish that in a class of monotonic games, under certain restrictions on out-of-equilibrium beliefs that stem from strategic stability, all types separate in equilibrium. However, separation of types occurs in a very different way in the two setups. We show that separating signals must be non-distorted, whereas in the single sender case, separating signals are subject to distortion due to simple equilibrium forces. In what follows, we connect the two setups and explain the main differences.

To do so, let us consider a simplified multi-sender setup in which the receiver's payoff does not depend on the signals and the payoff of each sender only depends on his own signal, the action of the receiver, and on the sender's type, but not on the sender's identity. Namely, let $v(t, m, a) = v(t, a)$, $e(t) = \text{argmax}_{a \in A} v(t, a)$, and for all $i \in S$, $M_i = [0, c]$ and $u_i(t, m, a) = u(t, m_i, a)$. We replace A5-A7 with:

**A5'** $u(t, m_i, a)$ is strictly concave and differentiable in $m_i$ for all $t, a$.

**A6'** $\exists m_i \in M_i : u(T-1, m_i, e(0)) > u(T-1, c, e(T))$.

**A7'** 1. $\frac{\partial u(t,m_i,a)}{\partial m_i}|_{m_i=0} > 0$, $\frac{\partial u(t,m_i,a)}{\partial m_i}|_{m_i=c} < 0$ for all $t, a$,

2. $\frac{\partial u(t,m_i,a)}{\partial m_i}$ is strictly increasing in $t$ and $a$.

Note that diagonal strict concavity (A5(2)) simplifies to a *strict* concavity of $u$ (A5').

Suppose now that the senders can collude and act together as if they were a single sender with the objective of maximizing the sum of their payoffs. Then, we are back to the single-sender setup of Cho and Sobel (1990) with multidimensional signals. We

[12] While A5, A6, and A7 are stated in terms of $u_i(t, m, e(t)(m))$, they directly translate to the primitives $u_i(t, m, a)$ if $u_i$ is additively separable in $m_i$ and $a$ (for example, as in Kohlberg (2003)) or if $v$ does not depend on $m$ (for example, similar to Spence (1973)).

will compare the unique SFI outcome of the multi-sender game with the symmetric D1 equilibrium outcome of this single-sender game. To find the latter, it is instructive first to characterize the D1 equilibrium outcome of the game where sender $i$ is the single sender with a one-dimensional signal. According to Propositions 4.4 and 4.5 in Cho and Sobel (1990), the D1 equilibrium outcome is unique and is in separating signals that are the solution to the following inductively defined programs $(Q(t))_{t\in\bar{T}}$.

Let $u^{*d}(0) = \max_{m_i\in M_i} u(0, m_i, e(0))$. Suppose $u^{*d}(0), \ldots, u^{*d}(t-1)$ have been defined. Then, $u^{*d}(t)$ is the value of the program $Q(t)$:

$$\max_{m_i\in M_i} u(t, m_i, e(t))$$

$$\text{subject to } u^{*d}(s) \geq u(s, m_i, e(t)) \text{ for } s = 0, \ldots, t-1,$$

and its solution is denoted by $m_i^d(t)$ and, of course, it is the same for all $i \in S$. The solution of this program is known as the Riley outcome. A5' and A7'(1) guarantee that each type $t$ of sender $i$ also has a unique and interior non-distorted signal $m_i(t) \in (0, c)$ that maximizes the unconstrained objective in $Q(t)$ with value $u_i^*(t) \geq u_i^{*d}(t)$. A7', together with A3, ensures that $m_i(t)$ is strictly increasing in $t$. A5' and A6', together with A4, also imply that $m_i(t) \leq m_i^d(t) < c$ and, hence, $m_i^d(t)$ is also strictly increasing in $t$. Unlike Cho and Sobel (1990), the additional assumption A7' also ensures that the constraints in $Q(t)$ can only bind for $t-1$ and, hence, the programs can be simplified.[13] It is simple to construct examples when $m_i(t) < m_i^d(t)$ holds for some $t \geq 1$. Let us assume that this is the case for type $\tau$.

Consider now the single-sender game with multidimensional signals and the payoff function $U(t, m, a) = \sum_{i\in S} u(t, m_i, a)$. Note that in any separating equilibrium $(\bar{m}(\cdot), e(\cdot), \mu)$, $e(\bar{m}(t)) = e(t)$ for all $t$. Therefore, in what follows, we focus on the (possible) separating equilibrium signals. Thus, on the one hand, it is easy to see that $((m_i^d(t))_{i\in S})_{t\in\bar{T}}$ is still (a part of) a D1 equilibrium outcome in this single-sender game (just change $u$ to $U$ and $m_i$ to $m$ in the programs $(Q(t))_{t\in\bar{T}}$ and recall that we are in a symmetric setup), while $((m_i(t))_{i\in S})_{t\in\bar{T}}$ is not even a PBE outcome because the single sender of type $\tau-1$ would deviate from $m(\tau-1)$ to $m(\tau)$ (or else the solution to $Q(\tau)$ would be $m_i^d(\tau) = m_i(\tau)$ for all $i$). On the other hand, $((m_i^d(t))_{i\in S})_{t\in\bar{T}}$ is a PBE outcome of the multi-sender game, but clearly not an SFI outcome. After a deviation

[13] We verify these statements starting with $m_i(t) \leq m_i^d(t)$. By A6', there is a feasible solution to the program $Q(t)$ to the right of $m_i(t)$. By A5', it is enough to consider the smallest such solution as a candidate for $m_i^d(t)$. Denote this signal by $m_i'$. Suppose that at $m_i'$, the incentive constraint of type $\tau' < t$ is binding. Furthermore, suppose that there exists another candidate solution that is to the left of $m_i(t)$. By A5', it is enough to consider the largest such solution. Denote this signal by $m_i''$ and suppose that at $m_i''$, the incentive constraint of type $\tau'' < t$ is binding. $\tau'$ weakly prefers $m_i'$ to $m_i''$ and, hence, by A4, $t$ strictly prefers $m_i'$ to $m_i''$. It follows that $m_i(t) \leq m_i^d(t)$ must hold and, given that $m_i(t)$ is strictly increasing in $t$, $m_i^d(t)$ also strictly increases. Additionally, it must be that $\tau' = t-1$. Suppose to the contrary that the constraint in $Q(t)$ is binding for $\tau' < t-1$. Since $\tau'$ weakly prefers $m_i^d(\tau')$ to $m_i^d(t-1)$, he also weakly prefers $m_i^d(t)$ to $m_i^d(t-1)$. But then, by A4, $t-1$ strictly prefers $m_i^d(t)$ to $m_i^d(t-1)$, which is a contradiction.

of sender $i$ of type $\tau$ from $m_i^d(\tau)$ to a non-equilibrium signal $m_i$ sufficiently close to $m_i(\tau)$, the receiver identifies sender $i$ as a deviator and, by unprejudiced beliefs, she still believes, after observing $m_{-i}^d(\tau)$, that the senders' type is $\tau$ and, hence, keeps her action $e(\tau)$. Therefore, such a deviation by sender $i$ of type $\tau$ is profitable. In fact, we know from Corollary 1 that the only SFI outcome is in the non-distorted signals $((m_i(t))_{i\in S})_{t\in\bar{T}}$.

### 4.2.2 Forward Induction and Related Concepts

Cho and Sobel (1990) also consider a larger class of monotonic games in which only the following assumption holds:

M For all $t, t', m, a, a'$ : $u(t, m, a') > u(t, m, a)$ implies $u(t', m, a') > u(t', m, a)$.

It says that if the sender of type $t$ strictly prefers one action to another, given some signal $m$, then any other type $t'$ will also do so. They prove in their Proposition 3.1 that in this class of games, D1 is equivalent to FI. We show that the analogous result also holds in the multi-sender setup.

We first provide very weak definitions of D1 and FI for pure strategy equilibria in multi-sender games. Similar to the single-sender case, they are based on the comparison of the incentives of different types to deviate. We only consider the out-of-equilibrium signals in $M(m(\cdot))$ and only the incentives of the sender who was identified as the deviator. Although one could restrict the beliefs even further after out-of-equilibrium signals (see Govindan and Wilson (2009) for a stronger definition of FI; see also a more detailed discussion in Footnote 14), the adopted definitions are sufficient for our purposes. Additionally, in analogy to SFI, we also define simple D1 (SD1) and show the equivalence between SFI and SD1. It then follows that we can also use SD1 instead of SFI in Corollary 1. We note that with some qualifications, as explained in Cho and Kreps (1987) on page 214, this equivalence holds for the intuitive criterion as well.

We start by defining sets of receiver's actions similar to the $P$-sets in Cho and Sobel (1990). Fix some, possibly distorted, equilibrium $(m(\cdot), e(\cdot), \mu)$. We only define these sets for the out-of-equilibrium signal profiles $m \in M(m(\cdot))$. Given $m$, let sender $i$ be the identified deviator and, hence, $m_{-i} = m_{-i}(\tau)$ for some $\tau \in T_m$. Consider *any* $t \in \bar{T}$, not necessarily in $T_m$, and let:

$$P_i(t|m_i) = \{a \in A | u_i^*(t) < u_i(t, m_i, m_{-i}(t), a)\},$$

$$P_i^0(t|m_i) = \{a \in A | u_i^*(t) = u_i(t, m_i, m_{-i}(t), a)\}.$$

Because $t$ does not need to be in $T_m$, $m_{-i}(t)$ can differ from $m_{-i}$. Intuitively, we now allow for the possibility that in addition to sender $i$, senders $-i$ have also deviated from $m_{-i}(t)$ to $m_{-i}$. Therefore, when considering the incentives of $(i, t)$ to deviate, we compare his equilibrium payoff with the payoff he would expect from deviation *assuming* that the other senders had not deviated (because the senders act non-cooperatively).

To keep matters as simple as possible in this section, we assume that $P_i^0(t|m_i)$ is non-empty for all $t \in \bar{T}$ and $m \in M(m(\cdot))$. In this way, we avoid that in the following definitions of D1 and FI, beliefs assign zero probability to every type after some $m \in M(m(\cdot))$. Denote by $T(m'_{-i}) = \{t'|m_{-i}(t') = m'_{-i}\}$ the set of types of senders $-i$ who pool on $m'_{-i}$, which may or may not be equal to $m_{-i}$. Note that $T(m_{-i}) = T(m_{-i}(\tau)) = T_m$.

We say that in a D1 equilibrium, the receiver can attach positive probability to type $t \in \bar{T}$ after signal profile $m$ only if there is no $t' \in T(m_{-i}(t))$ such that

$$P_i(t|m_i) \cup P_i^0(t|m_i) \subseteq P_i(t'|m_i). \tag{1}$$

In words, suppose there is some $(i, t)$ who has weak incentives to deviate to $m_i$ (l.h.s. of (1)), and there is $t'$ for which sender $(i, t')$ can also induce the signal profile $(m_i, m_{-i}(t))$, i.e., $t' \in T(m_{-i}(t)))$. Now, if $(i, t')$ has strict incentives to deviate (r.h.s. of (1)), then the receiver should attach 0 probability to type $t$ in a D1 equilibrium. Note that one cannot compare the incentives of $(i, t)$ and $(i, t')$ if senders $-i$ of type $t$ and $t'$ do not pool, i.e., if $m_{-i}(t) \neq m_{-i}(t')$. More importantly, unlike SFI, type $t \notin T_m$ can get a positive probability under D1.[14]

To define a multi-sender version of FI (or equivalently, NWBR), we require that the receiver can attach positive probability to type $t \in \bar{T}$ after signal profile $m$ only if:

$$P_i^0(t|m_i) \nsubseteq \cup_{t' \in T(m_{-i}(t)) \setminus \{t\}} P_i(t'|m_i), \tag{2}$$

that is, only if $m_i$ is a weak best response for $(i, t)$ (as defined in Section 2.2) given that senders $-i$ send the equilibrium signal $m_{-i}(t)$. Note again that $m_{-i}(t)$ can differ from $m_{-i}$ because $t$ does not have to be in $T_m$ and so, FI can assign a positive probability to $t \notin T_m$.

Because of unprejudiced beliefs, SFI requires that we assign zero probability to all types outside $T_m$. Therefore, we only need to apply the test in (2) to $t \in T_m$. If we do so, we find that the receiver can only attach positive probability to types in $F_m$, exactly as required by the definition of SFI in Section 2.2. In a similar way, we can also define an SD1 equilibrium by restricting the test in (1) only to $t \in T_m$ while the types outside $T_m$ are automatically assigned zero probability. This is a powerful restriction as demonstrated by the example in Section 2.3. Additionally, Vida and Honryo (2021) provide an example when pooling of two types $t, t' \in T_m$ may only be justified by

[14] In principle, when observing $m$ (with $m_{-i} = m_{-i}(\tau)$ for some $\tau \in T_m$), one should also consider the incentives of senders $j \in S \setminus \{i\}$ to deviate to $m_j$. When $t \in T_m$, there is no need to do it because $m_j = m_j(\tau) = m_j(t)$ is $(j, t)$'s equilibrium message for all $j \in S \setminus \{i\}$. However, if $t \notin T_m$, there might be another type $t' \in T(m_{-j}(t))$ for which $(j, t')$ has a stronger incentive to deviate to $m_j = m_j(\tau)$ than $(j, t)$ has and, for that reason, the receiver should put 0 probability on $t$. Moreover, $(m_j, m_{-j}(t)) \notin M(m(\cdot))$ (because $m_j = m_j(\tau)$). Therefore, there might be another sender $k \neq j$ of type $t''$, possibly $(i, \tau)$ himself, who could reach the information set $(m_j, m_{-j}(t))$ by unilateral deviation. In this case, one should also compare the incentives to deviate across different types of different players, namely, those of $(j, t)$ and $(k, t'')$. The definition of FI in Govindan and Wilson (2009) also requires such comparisons. Finally, the above considerations are not relevant in the case of SFI because by unprejudiced beliefs, types $t \notin T_m$ are already assigned 0 probability.

putting a positive probability on a third type $t'' \notin T_m$ who separates from $t, t'$, and this belief is not ruled out by FI. However, it is automatically excluded by SFI.

With these definitions, we now establish the equivalence between (simple) D1 and (simple) FI.

***Remark 1*** Suppose assumption M holds for all $i$ with $u$ replaced by $u_i$. Then, D1 is equivalent to FI and SD1 is equivalent to SFI.

***Proof Remark 1:*** We prove the equivalence between D1 and FI. The argument is similar for SD1 and SFI. Suppose some type $t$ is eliminated by FI after the out-of-equilibrium signal profile $m$, where $i$ is identified as the deviator. Pick any $a \in P_i^0(t|m_i)$. Because $t$ is eliminated by FI, there is $t'$ such that $a \in P_i(t'|m_i)$. Suppose now there is some $b \in P_i(t|m_i)$. By the definition of sets $P_i$'s, $(i, t)$ strictly prefers $b$ to $a$ at $m' = (m_i, m_{-i}(t))$. By assumption M, $(i, t')$ also strictly prefers $b$ to $a$ at $m'$ and, therefore, $b \in P_i(t'|m_i)$. It follows that $P_i^0(t|m_i) \subseteq \cup_{t' \in T(m_{-i}(t)) \setminus \{t\}} P_i(t'|m_i)$ implies $P_i(t|m_i) \cup P_i^0(t|m_i) \subseteq \cup_{t' \in T(m_{-i}(t)) \setminus \{t\}} P_i(t'|m_i)$. We want to show that $t$ is also eliminated by D1. Suppose not. It can only happen if there are some $t', t'' \in T(m_{-i}(t)) \setminus \{t\}$ and $a', a'' \in P^0(t|m_i) \cup P(t|m_i)$ such that $a' \in P(t'|m_i)$, $a'' \in P(t''|m_i)$, $a' \notin P(t''|m_i)$, and $a'' \notin P(t'|m_i)$. But then $u_i(t', m', a') > u_i(t', m', a'')$ and $u_i(t'', m', a'') > u_i(t'', m', a')$, which contradicts assumption M.

### 4.2.3 Connection with Stability

Cho and Sobel in their Proposition 3.2 also prove that, in finite games, FI (and D1) is equivalent to strategic stability à la Kohlberg and Mertens (1986), given assumption M. We give the formal definition of strategically stable sets and outcomes in Section B in the Appendix. FI, however, is not equivalent to strategic stability in the multi-sender setup even if all of our assumptions are satisfied and even if we adopt the stronger definition of FI of Govindan and Wilson (2009). The example in Section 2.3 gives an FI equilibrium outcome with distorted price signals that is not an SFI outcome. This outcome also does not survive iterative forward induction and, hence, is unstable according to Proposition 6B in Kohlberg and Mertens (1986).[15] However, we now show that a result in the spirit of Proposition 3.2 of Cho and Sobel (1990) holds for SFI in the multi-sender games.

First, we prove that strategic stability implies (O)SFI in generic finite games. It is well known that stable outcomes generically exist, possibly only in mixed strategies. Therefore, once we generalize the definition of SFI to mixed strategies in the obvious way, we also have the existence of (O)SFI outcomes in generic finite games. This

[15] We argued in Section 2.3 that it is a weak best response for firm 2 of type 0 to send signal 2.9 given that firm 1 of type 0 sends signal 2. To show this, we set $e(2, 2.9) = 0.595$. This action is sequentially rational for the entrant if $\mu_{(2,2.9)}(0) = 0.595$ and $\mu_{(2,2.9)}(1) = 0.405$. However, it is never a weak best response for firm 1 of type 1 to send signal 2. Therefore, in the first iteration of FI, we can restrict the belief and action of the entrant to $\mu_{(2,2.9)}(0) = e(2, 2.9) = 1$. With this restriction, it is now never a weak best response for firm 2 of type 0 to send signal 2.9. Then, in the second iteration of FI, the entrant's belief after the signal profiles $(m_1, 2.9)$ must be concentrated on $t = 1$, i.e., $\mu_{(m_1,2.9)}(1) = 1$. But then, firm 1 of type 1 will deviate from the distorted signal profile (2.8, 2.9) by choosing its best response signal $m_1 = (2.9 + 2 + 1)/2 = 2.95$.

result does not even require the assumption that types are perfectly correlated. The proof builds on Theorem 1 in Vida and Honryo (2021) and Proposition 6B of Kohlberg and Mertens (1986). The former shows that in generic arbitrary signaling games, any stable outcome can be maintained as a PBE outcome with unprejudiced beliefs, while the latter shows that stable outcomes survive FI. However, it is not immediate that a stable set contains an equilibrium having both properties at the same time.

**Proposition 1** *In generic finite signaling games, stable outcomes are (O)SFI outcomes.*

***Proof*** See the proof in Section B in the Appendix. □

Second, we obtain in the following proposition that the non-distorted outcomes of infinite game are the limits of stable outcomes of finite games that approximate the infinite game. Because SFI outcomes are non-distorted in monotonic infinite games according to Corollary 1, it follows that SFI also implies stability, albeit in the finitistic sense, in our class of games.

**Proposition 2** *Under assumptions A0-A3 and B2, any non-distorted outcome of the infinite game is the limit of stable outcomes of the approximating finite games.*[16]

***Proof*** See the proof in Section D in the Appendix. □

### 4.3 A (Non-)Example à la Spence, Dispensing with B2, SFI versus OSFI

Consider a two-sender, two-type version of the job market signaling model of Spence (1973), where education does not increase the marginal products. For concreteness, let $u_i(t, m, a) = a - m_i^2(2-t)$ and $M_i = [0, 1]$ for $i = 1, 2$, $v(t, m, a) = -(t-a)^2$, $A = [0, 2]$, $\bar{T} = \{0, 1\}$, and $\pi(0) = \pi(1) = 1/2$. The corresponding subgame perfect equilibrium education levels of the complete-information games are 0 for both types and both senders, i.e., $N(0) = \{((0, 0), 0)\}$ and $N(1) = \{((0, 0), 1)\}$. Because $m(0) = m(1)$, but $e(0) \neq e(1)$, non-distorted outcome does not exist in the incomplete-information game and, according to Theorem 1, neither does an OSFI outcome. This game violates assumption B2 and neither A7(1) nor A7(3) are satisfied at $m_i = 0$. Yet, there exist SFI outcomes that are necessarily distorted. In these SFI outcomes, one of the senders pools on the non-distorted 0 education level and the other sender separates according to the Riley outcome (we refer to this as the one-sided Riley outcome).[17] In what follows, we offer a solution for this type of games.

[16] The approximation is done using the finitistic approach that has been introduced by Simon and Stinchcombe (1989) and is also used in Simon and Stinchcombe (1995). A precise description of the approximation can be found in Section C in the Appendix. See also the next section for an example.

[17] Steps (1), (2), and (3) in the proof of Theorem 1 go through with a slight modification. The Spence model satisfies assumptions A0-A4. However, the signal spaces are not open intervals, which was exploited in steps (1) and (2) of the proof, but now they are compact. Nevertheless, it can still be shown that types will separate for both senders. In particular, they cannot pool on the highest signal because it is too costly and they cannot pool on lower signals because of the argument in step (1). If one of the senders pools, then the other must separate according to the Riley outcome by forward induction. (Unprejudiced beliefs have no bite here because once all types of one sender pool, it is as if there was a single sender.) This outcome can

Based on this game, we now introduce, somewhat informally, the finitistic approach and explain how it allows us to circumvent the problem of non-existence of both non-distorted outcome and OSFI. Let us consider a sequence of games, indexed by $n$, that are identical to the original game, except that the senders' signal spaces $M_i^n$ are finite grids on $M_i$ with grid size $1/n$ and with $0 \in M_i^n$. Let $M^n = \Pi_{i \in S} M_i^n$. We now take the limit of OSFI outcomes as the solution concept. A distribution $\lambda \in \Delta(\bar{T} \times M \times A)$ is an OSFI* outcome if there is a sequence of $\lambda_n \in \Delta(\bar{T} \times M^n \times A)$ converging to $\lambda$ such that $\lambda_n$ is an OSFI outcome of the game with signal spaces $M^n$.

In the example, $\lambda_n$ is $(0, (0, 0), 0)$ with probability 1/2 and $(1, (1/n, 1/n), 1)$ with probability 1/2. The proof that $\lambda_n$ is an OSFI outcome, is analogous to step (4) in the proof of Theorem 1. The limit of this sequence $\lambda$, which is $(0, (0, 0), 0)$ and $(1, (0, 0), 1)$ with equal probabilities, is an OSFI* outcome. Distribution $\lambda$ is not a feasible outcome in the infinite game. However, because the types separate in the finite games all along the approximating sequence and only pool at the limit, we can be arbitrarily close to $\lambda$ and so, effectively, there exists a non-distorted "outcome".

The definition of OSFI* as given above does not guarantee uniqueness. As in step (1) in the proof of Theorem 1, OSFI* ensures that all types separate. However, it can be that the types are only separated by one sender. Namely, step (2) in the proof breaks down, and the one-sided Riley outcome survives SFI and OSFI* but is rejected by OSFI.[18]

To summarize, if the conditions of Corollary 1 hold, then there is no need to refer to open-mindedness because SFI selects the unique non-distorted outcome. If some of the conditions of Corollary 1 do not hold, but conditions A0-A4 hold, one can employ OSFI to eliminate equilibria that are distorted. However, a non-distorted outcome may not exist in which case one can resort to the finitistic approach and use OSFI*. We hasten to say that the monotonicity conditions in Theorem 1 and Corollary 1 are by no means necessary in order to apply (O)SFI in games. However, whenever the monotonicity conditions hold, working with OSFI necessarily requires the existence of a non-distorted equilibrium that satisfies B2.

### 4.4 Related Notions of Forward Induction

In the definition of SFI, we only apply unprejudiced beliefs and forward induction to the out-of-equilibrium signal profiles in $M(m(\cdot))$. It was already sufficient to produce sharp predictions in monotonic multi-sender signaling games. In general *finite* multi-sender signaling games, however, one may want to strengthen the definition in several

itself be ruled out by showing that the deviation of low type to a slightly positive education level is a weak best response only for the low type and, hence, by open-mindedness (which puts a small weight $\varepsilon$ on the high type), the low type obtains a fixed increase in his wage. One can also easily construct games in which there is an OSFI outcome and other SFI outcomes. Suppose there is a third dummy sender with 2 signals and constant utility. Depending on the signal of this dummy sender, the payoffs of the other senders are either as described in this section or $u_i(t, m, a) = m_i a - (2 - t)m_i^2$, allowing education to increase the marginal products of the senders. In the latter case, there is an outcome satisfying B2, with $N(0) = \{((0, 0), 0)\}$ and $N(1) = \{((1/2, 1/2), 1)\}$, which is also an OSFI outcome. In the former case, there are only SFI outcomes as described above.

[18] To ensure uniqueness of the OSFI* "outcome", we additionally need to require that *the same* $\bar{\varepsilon}$ is used for all $n$ when checking whether $\lambda_n$ is an OSFI outcome. Then, similar to step (2) in the proof of Theorem 1, we obtain that all types will be separated by (at least) two senders.

obvious ways: 1) apply unprejudiced beliefs and forward induction to all out-of-equilibrium signal profiles that can be reached by unilateral deviations; 2) extend the definition to the mixed strategies; and 3) restrict the receiver's action space only to those actions that are sequentially rational for some consistent belief system. (The last point means that SFI is applied to sequential equilibria instead of PBE.) With these modifications, SFI implies the notions of FI in Kohlberg and Mertens (1986), Cho (1987), and Govindan and Wilson (2009); SFI restricts the support of beliefs to the types for which there is a sender who can reach the given information set with a unilateral deviation, which the other notions do not do. However, SFI is implied by stability, as shown in Proposition 1.

Man (2012) and Zheng (2017) refine the definition of FI in Govindan and Wilson (2009) by considering the possibility of iterative application of FI. We have seen in the example of Section 2.3 that iterative FI can imply SFI (see Footnote 15). However, in general, the iterative application of FI does not necessarily yield unprejudiced beliefs. Nevertheless, it follows from Proposition 1 that for generic games, one can always require unprejudiced beliefs at the very end of the iteration of FI without losing existence. Finally, there are also environments where SFI and FI coincide. For example, when the senders' types are drawn independently, unprejudiced beliefs are already implied by sequential equilibrium (see Proposition 1 in Vida and Honryo (2021)).

There are also other approaches to define the notion of forward induction. Thus, Battigalli and Siniscalchi (2002) derive the intuitive criterion for single-sender signaling games using an epistemic model. van Damme (1989) proposes to understand the reasoning of forward induction using outside options in games. For more on forward induction, see the surveys by Van Damme (2002), Fudenberg and Tirole (1991), Hillas and Kohlberg (2002), and Kreps and Sobel (1994).

## 5 Extension

The notion of unprejudiced beliefs and, hence, the notion of SFI is not context specific and can be easily generalized to arbitrary extensive-form games with perfect recall. We could strengthen the definition of FI in Govindan and Wilson (2009) to SFI by additionally requiring that beliefs also satisfy the following property. Consider an information set that can be reached with $k$ or more deviations, but not with less, from the equilibrium path. One could require that beliefs only put positive probability on those nodes that can be reached by $k$ deviations. This restriction is essentially equivalent to unprejudiced beliefs in multi-sender signaling games and, hence, together with FI, it is equivalent to SFI.

To ensure the existence of such an equilibrium, one can consider the extension of Proposition 1 to generic finite extensive-form games with perfect recall. Such an extension is not trivial for arbitrary $k$, and we leave it for future research whether such an equilibrium exists.[19]

[19] However, for $k = 1$, one can apply Mertens's notion of stability (see Mertens (1989, 1991)) and prove a version of Proposition 1 along the lines described in Footnote 24 in Vida and Honryo (2021).

## 6 Conclusion

We have introduced a new and powerful solution concept that can be easily generalized and applied to any (even infinite) extensive-form games with perfect recall. In generic finite multi-sender games, our solution concept is implied by strategic stability (even when the senders' types are not perfectly correlated as the proof of Proposition 1 does not require it) and, hence, our solution generically exists. In Theorem 1, we have shown how powerful the equilibrium selection is in monotonic infinite multi-sender games and proven that the solution is non-distorted. Moreover, we have also shown that the selected equilibrium outcome is the limit of stable outcomes of approximating finite games.

## Appendix

## A Proof of Corollary 1

The result follows from Van Zandt and Vives (2007, Corollary 15), Rosen (1965, Theorems 1 and 2), Cho and Sobel (1990, Proposition 4.5), and our Theorem 1. A5(1) ensures that the complete information games have equilibria (Theorem 1 in Rosen, 1965) and A5(2) ensures that these equilibria are unique (Theorem 2 in Rosen, 1965). Thus, assumption B1 is satisfied. Let us denote these unique equilibrium outcomes by $(m(t), e(t)(m(t))) \in N(t)$ for every $t \in \bar{T}$. A7 ensures that we are in the smooth case of Van Zandt and Vives (2007) and their Corollary 15 implies that the equilibria of the complete information games are strictly increasing in $t$ ($m(t) < m(t')$ when $t < t'$). Hence, assumption B2 is also satisfied. (In fact, all senders separate in the non-distorted outcome.) By Theorem 1, this non-distorted outcome is an OSFI outcome and hence, also an SFI outcome. It remains to be shown that it is a unique SFI outcome. A6 implies that there cannot be pooling on the top signal, just as in Proposition 4.5 of Cho and Sobel (1990). According to the proof of step (1) in Theorem 1 (which only uses forward induction and unprejudiced beliefs), all types separate in any SFI outcome. We now prove that all signals must be non-distorted in any separating SFI which completes the proof of the corollary. We do this by contradiction, and we only rely on unprejudiced beliefs without any reference to open-mindedness or forward induction.

Consider a separating equilibrium in which signals are sent according to some $\bar{m}(\cdot)$. Suppose there is a type $t'$ for which $\bar{m}(t') \neq m(t')$ and consider the lowest such type. (Hence, all types $t$ below $t'$ play according to $\bar{m}(t) = m(t)$.) Let sender $i$ be the one who is not best responding at $\bar{m}(t')$ in the complete information game corresponding to $t'$. First we claim that there must be a type $t < t'$ such that $m_{-i}(t) = \bar{m}_{-i}(t')$. Suppose not. Clearly, we can find a deviation $m_i$ such that $m = (m_i, \bar{m}_{-i}(t')) \in M(\bar{m}(\cdot))$ and $u_i(t', m, e(t')(m)) > u_i(t', \bar{m}(t'), e(t')(\bar{m}(t')))$. The previous inequality assumes that the receiver still believes that the senders' type is $t'$. Because $m_{-i}(t) \neq \bar{m}_{-i}(t')$ for all $t < t'$, $T_m$ may only contain types above $t'$, in addition to $t'$ itself. By unprejudiced beliefs, the receiver's belief (weakly) increases to $\mu_m \geq_{stoch.\,dom.} \mu_{\bar{m}(t')}(t') = 1$. Therefore, the receiver's action $e(\mu)(m)$ must be (weakly) higher than $e(t')(m)$. By A2,

$u_i(t', m, e(\mu)(m)) > u_i(t', \bar{m}(t'), e(t')(\bar{m}(t')))$, meaning that $(i, t')$ has a profitable deviation, a contradiction. Thus, there exists $t < t'$ such that $m_{-i}(t) = \bar{m}_{-i}(t')$.

Note that $u_i(t, m(t), e(t)(m(t))) \geq u_i(t, \bar{m}(t'), e(t')(\bar{m}(t')))$ holds because we are in equilibrium and only $i$ separates types $t$ and $t'$. If $\bar{m}_i(t') < m_i(t)$, then by A4, $u_i(t', m(t), e(t)(m(t))) > u_i(t', \bar{m}(t'), e(t')(\bar{m}(t')))$, which implies that $(i, t')$ has a profitable deviation, a contradiction. Thus, $m_i(t) < \bar{m}_i(t')$ must hold.

Now, by A7, it must be that senders $-i$ of type $t'$ are also not best responding at $\bar{m}(t')$ in the complete information game corresponding to $t'$. (Therefore, in fact, none of the senders is best responding.) This is because the derivatives (which were all 0 for senders of type $t$ at $m(t)$ by A7(1)) become strictly positive for senders $-i$ of type $t'$ at $\bar{m}(t')$ by A7(2) and A7(3). Take any $j \in -i$. We can repeat the above argument that we did for sender $i$, but now for sender $j$ to establish that there exists $t'' < t'$ such that $\bar{m}_{-j}(t'') = m_{-j}(t')$ and $m_j(t'') < \bar{m}_j(t')$ hold.

We have established that $m_i(t) < \bar{m}_i(t') = m_i(t'')$ and $m_j(t'') < \bar{m}_j(t') = m_j(t)$. However, one of these inequalities cannot hold because the equilibria of complete information games are strictly increasing in types, and so we have arrived at a contradiction. Thus, all signals must be non-distorted, $\bar{m}(t) = m(t)$ for all $t$.

## B 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 $\Gamma$ of a finite multi-sender signaling game. Let $\sigma = (\sigma_1, \ldots, \sigma_{|S|})$, where $\sigma_i$ is a completely mixed-strategy of sender $i \in S$.[20] For $\delta > 0$, consider the set of all normal form games $\Gamma'$ that have the same strategy space as $\Gamma$ 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 $\Gamma'$, 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 $\Gamma$. A game in this set is called a $(\sigma, \delta)$ perturbation of $\Gamma$.

**Definition 4** A set of Nash equilibria of $\Gamma$ is *stable* if it is minimal with respect to the following property: $\mathcal{N}$ is a closed set of Nash equilibria of $\Gamma$ satisfying: for each $\epsilon > 0$, there is a $\delta > 0$ such that for any completely mixed $\sigma$, the $(\sigma, \delta)$ perturbations of $\Gamma$ have a Nash equilibrium $\epsilon$-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 that are never-weak-best responses (inferior), e.g., those that 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 identified as a deviator, are supported on $F_m$ if $F_m$ is not empty, and on $T_m$ if $F_m = \emptyset$. Additionally, open-mindedness can also be satisfied for some $\bar{\varepsilon}$ sufficiently small because the game is finite and the equilibria in the stable set are perfect.

[20] 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 strategies with behavioral ones.

## C The Finitistic Approach for Infinite Games

We extend the scope of the solution concepts defined for finite games to infinite games following the finitistic approach introduced by Simon and Stinchcombe (1989) and (1995).

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 $\bar{T}$, and fix the payoff functions of the senders $u = (u_1, \ldots, u_{|S|})$ and 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 the set of signals $M_i^n$ and the 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 $x^n \in \mathbb{R}^{\dim G^n}$ induced by $(u, v)$, where $\dim G^n = (|S| + 1)|\bar{T} \times M^n \times A^n|$, 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 5** Fix an infinite multi-sender signaling game form $G$ with $(u, v)$. We say that $\lambda \in \Delta(\bar{T} \times M \times A)$ 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 $(\lambda^n)_{n\in\mathbb{N}} : \lambda^n \in \Delta(\bar{T} \times M^n \times A^n)$ of the games $G^n$, with $(u^n, v^n)$, weakly converging (in the topology of weak convergence) to $\lambda$.

***Remark 2*** 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 (O)SFI* outcome always exists. We show in the supplementary material that a stable* outcome always exists.[21] It then immediately follows from Proposition 1 and Definition 5 that an (O)SFI* outcome also always exists.

## D Proof of Proposition 2

Consider a non-distorted outcome $\lambda$ of an infinite $G$ with $(u, v)$ as in section C. Because of A2, it is a Nash (in fact, a PBE) outcome, and one can easily see that it is also a Nash* (PBE*) outcome. Consider a pair $(m(\cdot), e(\cdot))$ for which $[m(\cdot), e(\cdot)] = \lambda$. We can pick a sequence of game forms $G^n \to G$ and a sequence of payoffs $y^n = (u^n, v^n) \to (u, v)$ such that in the finite games, $\lambda$ is still a Nash outcome, and in the corresponding complete information games, $m(t)$ is part of a strict subgame perfect equilibrium for all $t$ in a neighborhood of $y^n$ denoted by $B(y^n, \epsilon^n)$.[22] We show that $\lambda$ is a pure stable* outcome with the sequence of game forms $(G^n)_{n\in\mathbb{N}}$ together with the sequence of open

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

[22] In strict equilibria, deviators are always strictly worse off, and these equilibria remain strict in an open neighborhood of payoffs.

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 $\lambda$. For simplicity, we only perturb the senders' strategies and with the same $\delta$. The proof also goes through for the case when these $\delta$-s are different for different senders.

Fix an $\varepsilon > 0$. We design a $\bar{\delta}$ such that for all strategy perturbations $\sigma^*$ of the senders and all $\delta < \bar{\delta}$, we have an equilibrium $(\sigma', e)$ of the $(\sigma^*, \delta)$ perturbed game such that $\sigma'$ is $\varepsilon$ close to $m(\cdot)$. Fix a $\xi < \varepsilon$. We will set this $\xi$ 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. (perturbation) with probability $\delta$, the signal is chosen according to $\sigma_i^*$;
2. (restriction) with probability $(1-\delta)(1-\xi)$, the strategy $m_i(\cdot)$ is chosen.

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

There must be a mixed-strategy equilibrium in this game. Denote the senders' strategies in this equilibrium by $\sigma$. We show that in this mixed equilibrium, it must be that for all types $t$ of each sender $i$, we have $\sigma_i(m_i(t)|t) > 0$. Hence, $\sigma$ can be transformed in the obvious way into a $\sigma'$ that is part of an equilibrium of the perturbed game without restriction (corresponding to $\xi = 1$). This is because the restrictions are actually best responses. Moreover, the resulting $\sigma'$ will be $\varepsilon$-close, in fact, $\xi$-close to $m(\cdot)$. We proceed by contradiction and assume that for all $\xi < \epsilon$, there is sequence of $\delta$-s converging to 0 such that w.l.o.g. (because the game is finite) there is a $t$ with $\sigma_1(m_1(t)) = 0$. 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 $O_1 = \{m_1' \in M_1 | \forall t \in \bar{T}, m_1(t) \neq m_1'\}$ be the set of signals of sender 1 in the finite game that are never sent 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 $m = (m_1, m_{-1}(t))$ is sent by senders of type $t$ can be bounded from below by $\frac{1}{k}(1-\delta)\xi((1-\delta)(1-\xi))^{|S|-1}$ (independently of $\sigma^*$). The probability that the signal profile $m$ is sent by senders of type $t' \neq t$ can be bounded (independently of $\sigma^*$) from above by $((1-\delta)\xi + \delta)^2 \times 1^{|S|-2}$ because, by B2, $t$ is separated from $t'$ by at least two senders. Note that only $k$ depends on $n$. Hence, for each $k$, we can choose $\xi$ to be sufficiently small so that whenever $\delta < \xi$, the ratio

$$\frac{((1-\delta)\xi + \delta)^2}{\frac{1}{k}(1-\delta)\xi((1-\delta)(1-\xi))^{|S|-1}}$$

is arbitrarily close to 0. That is, the receiver puts weight arbitrarily close to 1 on the event that the signal profile arrived from senders of type $t$ and, hence, plays the action $e(\mu)(m)$, where $\mu_m(t) = 1$. But then, sender 1 of type $t$ gets strictly less, as opposed to sending the (complete-information) equilibrium signal $m_1(t)$, if $\xi$ is small enough, because the corresponding subgame perfect equilibrium was chosen to be strict (in a neighborhood) and so, we have a contradiction.

Second, suppose that $m_1 \notin O_1$, that is, there is a $t' \neq t$ such that $m_1 = m_1(t') \neq m_1(t)$. Let us write again $m = (m_1, m_{-1}(t))$ and note that $m$ is still off the equilibrium path of the original equilibrium by B2 (because $t$ and $t'$ are separated by at least two senders). Consider the set of types $T'$, with generic element $t''$, such that the signal profile $m$ can be reached by a unilateral deviation of some sender $i$ of type $t''$ from $m(t'')$ and $\sigma_i(m_i|t^{''}) > 0$ and $m_i \neq m_i(t'')$. Clearly, $t \in T'$ with $i = 1$ by assumption. Bounding the probabilities as before, one can show that the belief of the receiver will put probability arbitrary close to 1 on $T'$ if $\xi$ is chosen to be small and as $\delta$ approaches 0. Consider now the highest $\bar{t} \in T'$ and the corresponding player $i$. Now the receiver will either choose the action $e(\mu)(m)$, where $\mu_m(\bar{t}) = 1$ or, by A3, a strictly lower action. In both cases, we get a contradiction because sender $i$ of type $\bar{t}$ is strictly worse off by sending $m_i$ than $m_i(\bar{t})$. In the former case, it is as before because of the strictness of subgame perfect equilibrium of the corresponding complete information game. In the latter case, it is for the same reason but we must also invoke A2, namely, that senders prefer higher actions of the receiver.

Hence, there is a $\xi < \epsilon$ and a $\bar{\delta}$ such that for all $\delta < \bar{\delta}, \sigma_i(m_i(t)) > 0$ in the restricted and perturbed game, i.e., $m_i(t)$ is a best response for all $t$ and all $i$ and, hence, $\sigma$ can be transformed into a $\sigma'$ that is part of a Nash equilibrium of the perturbed game without restriction and is $\xi$-close to $m(\cdot)$. Q.E.D.

**Funding** Open access funding provided by Corvinus University of Budapest.

## Declarations

**Conflicts of Interest** The authors have no conflict of interest to report.



## References

Au, P.H., Kawai, K.: Competitive disclosure of correlated information. Econ. Theor. **72**, 767–799 (2021)
Bagwell, K., Ramey, G.: Oligopoly Limit Pricing. RAND Journal of Economics **22**, 155–172 (1991)
Baliga, S., Corchon, L.C., Sjostrom, T.: The Theory of Implemenetation When the Planner Is a Player. Journal of Economic Theory **77**, 15–33 (1997)
Banks, J., Sobel, J.: Equilibrium Selection in Signaling Games. Econometrica **55**, 647–661 (1987)
Battaglini, M.: Multiple Referrals and Multidimensional Cheap Talk. Econometrica **70**, 1379–1401 (2002)
Battigalli, P., Siniscalchi, M.: Strong Belief and Forward Induction Reasoning. Journal of Economic Theory **106**, 356–391 (2002)
Bester, H., Demuth, J.: Signalling Rivalry and Quality Uncertainty in a Duopoly. J. Ind. Compet. Trade **15**, 135–154 (2015)
Blume, L., Zame, W.: The Algebraic Geometry of Perfect and Sequential Equilibrium. Econometrica **62**, 783–794 (1994)
Chen, M., Serfes, K., Zacharias, E.: Prices as Signals of Product Quality in a Duopoly. Internat. J. Game Theory **52**, 1–31 (2023)
Cho, I.: A Refinement of Sequential Equilibrium. Econometrica **55**, 1367–1389 (1987)
Cho, I., Kreps, D.: Signaling Games and Stable Equilibria. Quart. J. Econ. **102**, 179–221 (1987)

Cho, I., Sobel, J.: Strategic Stability and Uniqueness in Signaling Games. Journal of Economic Theory **50**, 381–413 (1990)
Emons, W., Fluet, C.: Accuracy Versus Falsification Costs: The Optimal Amount of Evidence under Different Procedures. J Law Econ Organ **25**, 134–156 (2009)
Fudenberg, D., Tirole, J.: Game Theory. MIT Press, Cambridge, MA (1991)
Govindan, S., Wilson, R.: On Forward Induction. Econometrica **77**, 1–28 (2009)
Hartman-Glaser, B., Hébert, B.: The Insurance Is the Lemon: Failing to Index Contracts. Journal of Finance **75**, 463–506 (2020)
Hillas, J., Kohlberg, E.: Foundations of Strategic Equilibrium. *Handbook of Game Theory with Economic Applications*, Vol. III, ed. by R. Aumann and S. Hart. New York: Elsevier, 1597-1663 (2002)
Hong, F., Pang, Y.: Mitigation and Adaptation: An Informational Perspective. J. Econ. **141**, 57–92 (2024)
Kohlberg, E., Mertens, J.-F.: On the Strategic Stability of Equilibria. Econometrica **54**, 1003–1037 (1986)
Kohlberg, E.: Demonstration of Strength Requires Sacrifice. HBS working paper, (2003)
Kreps, D., Sobel, J.: Signalling. *Handbook of Game Theory with Economic Applications*, Vol. II, ed. by R. Aumann and S. Hart. New York: Elsevier, 849-867 (1994).
Laslier, J.-F., Van der Straeten, K.: Electoral competition under imperfect information. Econ. Theor. **24**, 419–446 (2004)
Man, P.: Forward Induction Equilibrium. Games Econom. Behav. **75**, 265–276 (2012)
Mertens, J.-F.: Stable Equilibria: A Reformulation. Part I. Definition and Basic Properties. Math. Oper. Res. **14**, 575–625 (1989)
Mertens, J.-F.: Stable equilibria: A Reformulation. Part II. Discussion of the Definition, and Further Results. Math. Oper. Res. **16**, 694–753 (1991)
Qiu, Y., Rao, R.C.: Can Merchants Benefit from Entry by (Amazon-Like) Platform if Multiagent Prices Signal Quality? Mark. Sci. **43**, 778–796 (2023)
Rosen, J.: Existence and Uniqueness of Equilibrium Points for Concave N-Person Games. Econometrica **33**, 520–534 (1965)
Schultz, C.: Polarization and Inefficient Policies. Rev. Econ. Stud. **63**, 331–343 (1996)
Schultz, C.: Limit Pricing When Incumbents Have Conflicting Interests. Int. J. Ind. Organ. **17**, 801–825 (1999)
Simon, L., Stinchcombe, M.: Extensive Form Games in Continuous Time: Pure Strategies. Econometrica **57**, 1171–1214 (1989)
Simon, L., Stinchcombe, M.: Equilibrium Refinement for Infinite Normal-Form Games. Econometrica **63**, 1421–1443 (1995)
Spence, M.: Job Market Signaling. Q. J. Econ. **87**, 355–374 (1973)
Vaccari, F.: Competition in Costly Talk. Journal of Economic Theory **213**, 105740 (2023)
Van Zandt, T., Vives, X.: Monotone Equilibria in Bayesian Games of Strategic Complementarities. Journal of Economic Theory **134**, 339–360 (2007)
van Damme, E.: Stable Equilibria and Forward Induction. Journal of Economic Theory **48**, 476–496 (1989)
van Damme, E.: Strategic Equilibrium. *Handbook of Game Theory with Economic Applications*, Vol. III, ed. by R. Aumann and S. Hart. New York: Elsevier, 1521-1596 (2002)
Vida, P., Honryo, T.: Strategic Stability of Equilibria in Multi-Sender Signaling Games. Games Econom. Behav. **127**, 102–112 (2021)
Zhang, Q.: Vagueness in Multidimensional Proposals. Games Econom. Behav. **121**, 307–328 (2020)
Zheng, B.: Strong Forward Induction. *The B.E. Journal of Theoretical Economics*, 17, 20160067 (2017).