Equilibrium selection in interdependent value auctions *

ELNAZ BAJOORI † DRIES VERMEULEN ‡

August 7, 2018

Abstract

In second-price auctions with interdependent values, bidders do not necessarily have dominant strategies. Moreover, such auctions may have many equilibria. In order to rule out the less intuitive equilibria, we define the notion of distributional strictly perfect equilibrium (DSPE) for Bayesian games with infinite type and action spaces. This equilibrium is robust against arbitrary small perturbations of strategies. We apply DSPE to a class of symmetric second-price auctions with interdependent values and show that the efficient equilibrium defined by Milgrom [27] is a DSPE, while a class of less intuitive, inefficient, equilibria introduced by Birulin [12] is not.

JEL Codes. C72, D44.

Keywords. Equilibrium selection, Strictly perfect equilibrium, Interdependent value auctions, Bayesian games with infinite type and action spaces.

---

*The authors would like to thank Philip J. Reny and János Flesch for their very helpful and constructive comments.

†Corresponding author, e.bajoori@bath.ac.uk. University of Bath, Dept. of Economics, Bath, BA2 7AY, United Kingdom.

‡d.vermeulen@maastrichtuniversity.nl. Maastricht University, School of Business and Economics, Dept. of Quantitative Economics, Maastricht, The Netherlands., P.O. Box 616, 6200 MD Maastricht, The Netherlands.
1 Introduction

In private value second-price auctions each bidder has a dominant strategy in which he bids his own valuation of the object. However, in second-price auctions with interdependent values bidders may not have dominant strategies while there may exist multiple equilibria in undominated strategies. For example Birulin (2003) [12] shows that any auction with an efficient ex-post equilibrium has a continuum of inefficient undominated ex-post equilibria. Therefore, a selection tool is needed in such Bayesian games to rule out the less intuitive equilibria. We develop a concept of trembling hand perfect equilibrium for Bayesian games with infinite type and action spaces, called distributional strictly perfect equilibrium (DSPE). In this equilibrium concept, players’ strategies are robust against any slight perturbations of their opponents’ strategies.

In finite normal form games, DSPE is equivalent to strictly perfect equilibrium by Okada (1981) [29], 1 as both definitions require robustness against arbitrary slight perturbations of strategies. It is known that strictly perfect equilibrium may not exist (see example 1.5.5 in van Damme [33]), therefore DSPE may not exist even in finite games. However, we can prove existence of a weaker notion of DSPE, which is called distributional perfect equilibrium, in Bayesian games for which the type space of each player is a separable metric space, the action space of each player is a compact metric space, and player types are drawn from a prior probability measure on the product of the type spaces that is absolutely continuous with respect to the product measure of its marginal probabilities. 2 3

We apply this notion of perfection to second price auctions with interdependent values in order to select among equilibria. In our auction model, there is one continuous and efficient equilibrium introduced by Milgrom (1981) [27] and a set of inefficient ex-post equilibria introduced by Birulin [12]. All of these equilibria are undominated and there is no dominant strategy. We prove that the only efficient equilibrium is DSPE and the rest are not. In other words, our results show

1The proof is available upon request.
2This property is called absolute continuity of information in Milgrom and Weber [28].
3For further studies on the existence of equilibrium in Bayesian games, we refer to Athey [3], Mcadams [23], Reny [31], Jackson et al. [17],[18], Mertens [26], Meirowitz [24], Milgrom and Weber [28], Mallozzi et al. [25], Kim and Yannelis [19], Zandt and Vives [35], and Balder [7].
that the efficient and continuous equilibrium is robust against arbitrary slight perturbation of strategies, while other discontinuous and inefficient equilibria are not. These discontinuous equilibria contain either overbidding or underbidding when compared with the efficient strategy profile by Milgrom [27]. We show that if each bidder believes that with a very small probability her opponent can bid according to the efficient bidding strategy, then overbidding or underbidding is not a best response.

Bajoori et al. (2016) [6], define the notion of Perfect BNE and apply this notion to an example in interdependent value auctions. In a Perfect BNE the strategies are robust against some slight perturbations of strategies. This definition is weaker than the notion of DSPE, but it is not strong enough to exclude all the discontinuous and inefficient equilibria in a more general model of an interdependent value auction.

To the best of our knowledge there is not much research done on the theory of equilibrium selection in interdependent value auctions. However, there are some relatively recent studies on the equilibrium selection in common value auctions such as Parreiras (2006) [30], Abraham et al. (2012) [2], Cheng and Tan (2008) [13], Larson (2009) [21], and Liu (2014) [22]. In the most relevant study by Liu [22], they use noisy bids rather than strategy perturbations. In their model, bidders need to believe that any bid is possible for their opponent even if this strategy does not have full support. They show that the equilibria introduced by Moilgrom [27] are robust against such noisy bids, while the Birulin [12] equilibria are note. Our selection tool is a refinement of Bayesian Nash equilibrium that uses strategy perturbations and is successful in excluding discontinuous and inefficient equilibria in interdependent value auctions.

To illustrate our results, we discuss the following example.

**Example 1.1** Consider a 2-bidder second-price auction with interdependent values \( v_1 = t_1 + \frac{1}{2}t_2 \) and \( v_2 = \frac{1}{2}t_1 + t_2 \). Types \( t_1, t_2 \) for bidder 1 and 2 are drawn independently from \([0, 1]\) according to the uniform distribution. Each bidder simultaneously submits a bid from the set \([0, \frac{3}{2}]\). The efficient BNE introduced by Milgrom [27] is the strategy pair \((b_1, b_2)\) in which \( b_i = \frac{3}{2}t_i \) for every bidder \( i = 1, 2 \). We can show that this BNE is robust against any slight perturbations...
of the strategies (Proposition 4.2). On the other hand, the class of inefficient ex post equilibria introduced by Birulin [12] are as follows. For fixed \(s_1, s_2 \in [0, 1]\) with \(s_1 < s_2\), define

\[
\hat{b}_1(t_1) = \begin{cases} 
  s_2 + \frac{1}{2}t_1 & \text{if } t_1 \in [s_1, s_2] \\
  \frac{3}{2}t_1 & \text{otherwise,}
\end{cases} \\
\hat{b}_2(t_2) = \begin{cases} 
  s_1 + \frac{1}{2}t_2 & \text{if } t_2 \in [s_1, s_2] \\
  \frac{3}{2}t_2 & \text{otherwise.}
\end{cases}
\]

An example of such an equilibrium is depicted in Figure 1. In each such an equilibrium, when the realizations of both types are within the set \([s_1, s_2]\), bidder 1 wins the auction and pays the amount \(\hat{b}_2(t_2) = s_1 + \frac{1}{2}t_2\), which is his minimum ex post valuation. Thus, in this case bidder 1 overbids, while bidder 2 underbids. So, since such equilibria lead to inefficient outcomes, we would like to be able to rule them out. Proposition 4.4 shows that the equilibria \((\hat{b}_1, \hat{b}_2)\) are not DSPEs. Intuitively, to exclude such equilibria it is enough to make the bidders believe that with a tiny probability their opponents bid according to the efficient bidding strategy. Then, the overbidding (underbidding) is not a best response against underbidding (overbidding).

![Figure 1](image_url)

The paper is structured as follows. First, we provide some preliminary notions in Section 2. In section 3 we introduce the concept of distributional strictly perfection. In section 4, first we present our interdependent value auction model. Then, we have two subsections; one on selected equilibrium and the other on discarding equilibria.
2 Preliminaries

A Bayesian game $\Gamma$ is defined as follows. There are $n$ players. For each player $i$, the set $T_i$ of types is a complete separable metric space, and the set $A_i$ of actions is a compact metric space. Let $d_{T_i}$ and $d_{A_i}$ denote the respective metrics, and let $\mathcal{T}_i$ and $\mathcal{A}_i$ denote the induced Borel $\sigma$-fields on $T_i$ and $A_i$ respectively. Moreover, let $T = \times_{i=1}^n T_i$ and $A = \times_{i=1}^n A_i$.

Let $\mu$ be a probability measure on the product $\sigma$-field $\mathcal{T} = \otimes_{i=1}^n \mathcal{T}_i$ with marginal probability $\mu_i$ on $\mathcal{T}_i$ for every player $i$. We assume that each $\mu_i$ is a completely mixed probability measure.

Player $i$’s payoff function $\pi_i : T \times A \to \mathbb{R}$ is assumed to be bounded and jointly measurable.

We define two classes of strategies: behavioral strategies and distributional strategies. We also define Bayesian Nash Equilibrium (BNE).

A pure strategy for player $i$ in $\Gamma$ is a measurable function $p_i : T_i \to A_i$. A behavioral strategy for player $i$ is a function $\beta_i : T_i \times A_i \to [0, 1]$ such that

1. the section function $\beta_i(t_i, \cdot) : A_i \to [0, 1]$ is a probability measure for every $t_i \in T_i$, and

2. the section function $\beta_i(\cdot, B) : T_i \to [0, 1]$ is measurable for every $B \in A_i$.

When player $i$ plays according to the behavioral strategy $\beta_i$, for each type $t_i \in T_i$, he chooses his action according to the probability measure $\beta_i(t_i, \cdot)$.

A behavioral strategy $\beta_i$ for player $i$ is called pure if there is a pure strategy $p_i$ for player $i$ such that $\beta_i(t_i, \cdot) = D_{p_i(t_i)}(\cdot)$ for every type $t_i \in T_i$, where $D$ denotes Dirac measure. For simplicity, we denote this pure behavioral strategy by $p_i$.

Following Milgrom and Weber [28], a distributional strategy for player $i$ in a Bayesian game is a probability measure $\gamma_i$ on $T_i \times A_i$ such that the marginal distribution of $\gamma_i$ on $T_i$ equals the probability measure $\mu_i$. That is, $\gamma_i(U \times A_i) = \mu_i(U)$ for every $U \in \mathcal{T}_i$. Given a behavioral strategy $\beta_i$, the induced distributional

---

4A probability measure is completely mixed if it assigns strictly positive weight to every nonempty open set.
strategy $\gamma_i$ is uniquely determined by

$$\gamma_i(U \times B) = \int_U \beta_i(t_i, B) \mu_i(dt_i)$$

for every rectangle $U \times B \in T_i \otimes A_i$. This induces a many-to-one mapping from behavioral strategies to distributional strategies that preserves the players’ ex ante expected payoffs.

The vector $\beta = (\beta_1, \beta_2, \ldots, \beta_n)$, where $\beta_i$ is player $i$’s behavioral strategy, is called a behavioral strategy profile. For every player $i$ we write $\beta_{-i} = (\beta_j)_{i \neq j=1}^n$ to denote the behavioral strategy profile for his opponents. The definitions are similar for distributional strategies.

Let $\Delta_i$ be the set of all distributional strategies for player $i$ and $\Delta = \times_{i=1}^n \Delta_i$. Let $\Pi_i : \Delta \to \mathbb{R}$ be the expected payoff of player $i$ playing distributional strategy $\gamma_i$ against a distributional strategy profile $\gamma_{-i}$. So,

$$\Pi_i(\gamma_i, \gamma_{-i}) = \int \pi_i(t, a) \gamma_i(da_i \mid t_i) \gamma_{-i}(da_{-i} \mid t_{-i}) \mu(dt),$$

where $t = (t_1, \ldots, t_n) \in T$, $a = (a_1, \ldots, a_n) \in A$, and $\gamma_i(da_i \mid t_i)$ is a version of conditional probability on $A_i$ given $t_i$. By integrals with respect to $\gamma_{-i}(da_{-i} \mid t_{-i})$ we mean the iterated integrals with respect to $\gamma_j(da_j \mid t_j)$ for all $j \neq i$.

As the measure $\mu$ is absolutely continuous with respect to $\hat{\mu}$, by the Radon-Nikodym theorem there is a measurable function $f : T \to \mathbb{R}$ such that for every measurable set $G \in T$ we have $\mu(G) = \int_G f d\hat{\mu}$. Therefore, the expected payoff $\Pi_i$ can be expressed in an easier form as

$$\Pi_i(\gamma_i, \gamma_{-i}) = \int \pi_i(t, a) f(t) d\gamma_i d\gamma_{-i}.$$

Let $\beta_i$ be a behavioral strategy corresponding to the distributional strategy $\gamma_i$, for every player $i$. Then, by Theorem 10.2.1 in Dudley [14]5 we conclude that

$$\Pi_i(\gamma_i, \gamma_{-i}) = \int \pi_i(t, a) f(t) \beta_i(t_i, da_i) \mu_i(dt_i) \beta_{-i}(t_{-i}, da_{-i}) \mu_{-i}(dt_{-i}).$$

5Dudley’s notation and terminology in [14] differs from ours. Part (II) of Theorem 10.2.1. in Dudley [14] with our terminology can be expressed as follows: For every measurable function $g : T_i \times A_i \to \mathbb{R}$ we have

$$\int_{T_i \times A_i} g d\gamma_i = \int_{T_i} \int_{A_i} g \beta_i(t_i, da_i) \mu_i(dt_i).$$
We say that a behavioral strategy $\beta_i$ for every player $i$ against a behavioral strategy profile $\beta_{-i}$. Thus,

$$\Pi_i^t(\beta_i, \beta_{-i} \mid t_i) = \int \pi_i(t, a) f(t) \beta_i(t_i, da_i) \beta_{-i}(t_{-i}, da_{-i}) \mu_{-i}(dt_{-i}).$$

Furthermore, notice that

$$\Pi_i(\gamma_1, \ldots, \gamma_n) = \int \Pi_i^t(\beta_i, \beta_{-i} \mid t_i) \mu_i(dt_i).$$

In the special case where player $i$ chooses a pure behavioral strategy $p_i$, his expected payoff is

$$\Pi_i^t(p_i, \beta_{-i} \mid t_i) = \int \pi_i(t, (p_i(t_i), a_{-i})) f(t) \beta_{-i}(t_{-i}, da_{-i}) \mu_{-i}(dt_{-i}).$$

One can also define $\Pi_i^t(\sigma_i, \beta_{-i} \mid t_i)$ and $\Pi_i^t(a_i, \beta_{-i} \mid t_i)$ respectively for every probability measure $\sigma_i$ on $A_i$ and every action $a_i \in A_i$.

A probability measure $\sigma_i$ on $A_i$ is called a best response of player $i$ for type $t_i \in T_i$ against a behavioral strategy profile $\tau_{-i}$, if for every probability measure $\sigma'_i$ on $A_i$ we have

$$\Pi_i^t(\sigma_i, \tau_{-i} \mid t_i) \geq \Pi_i^t(\sigma'_i, \tau_{-i} \mid t_i).$$

The set of such best responses is denoted by $BR_i(t_i, \tau)$. A behavioral strategy $\beta_i$ is called a best response of player $i$ against the behavioral strategy profile $\tau_{-i}$, if $\beta_i(t_i, \cdot) \in BR_i(t_i, \tau)$ for every $t_i \in T_i$. A behavioral strategy profile $\beta = (\beta_1, \ldots, \beta_n)$ is a Bayesian Nash Equilibrium (BNE) if $\beta_i$ is a best response to $\beta_{-i}$ for every player $i$.

A distributional strategy $\gamma_i$ of player $i$ is a best response against a distributional strategy profile $\eta_{-i}$, if $\Pi_i(\gamma_i, \eta_{-i}) \geq \Pi_i(\tilde{\gamma}_i, \eta_{-i})$, for every distributional strategy $\tilde{\gamma}_i$ of player $i$. Let $BR_i : \Delta \rightarrow \Delta_i$ denote the best response correspondence for player $i$ on the set of distributional strategies. Moreover, define $BR : \Delta \rightarrow \Delta$, for every $\gamma \in \Delta$, as $BR(\gamma) = \times_{i=1}^n BR_i(\gamma)$.

We say that a behavioral strategy $\beta_i$ (weakly) dominates behavioral strategy $\hat{\beta}_i$ if, for their respective induced distributional strategies $\gamma_i$ and $\hat{\gamma}_i$, it holds that

$$\Pi_i(\gamma_i, \gamma_{-i}) \geq \Pi_i(\hat{\gamma}_i, \gamma_{-i}).$$
for all $\gamma_{-i} \in \Delta_{-i}$, and with a strict inequality for at least one $\gamma_{-i} \in \Delta_{-i}$. A behavioral strategy $\beta_i$ is called (weakly) dominant if it dominates every other behavioral strategy of player $i$. A behavioral strategy $\beta_i$ is undominated if there is no other behavioral strategy that dominates $\beta_i$.

3 Distributional strictly perfect equilibrium

In this section, we define perfect equilibrium for Bayesian games in terms of behavioral strategies. Since only the induced distribution on opponents’ actions matter for computation of expected payoffs, we use distributional strategies as a tool to verify whether the behavioral strategy profile satisfies the conditions of trembling hand perfect equilibrium.

Let $C_b(T_i \times A_i)$ be the set of all continuous and bounded functions with domain $T_i \times A_i$. The weak topology on $\Delta_i$ is the coarsest topology for which all the maps $\gamma_i \mapsto \int h \ d\gamma_i$, $h \in C_b(T_i \times A_i)$ are continuous. Since $T_i \times A_i$ is a separable metric space, the weak topology on $\Delta_i$ is separable and metrizable with the weak (Prohorov) metric, denoted by $\rho_w$.

**Definition 3.1** A behavioral strategy profile $\beta = (\beta_1, \ldots, \beta_n)$ is called *distributional strictly perfect equilibrium* (DSPE) if, given the induced profile $\gamma = (\gamma_1, \ldots, \gamma_n)$ of distributional strategies, for every completely mixed strategy profile $\nu = (\nu_1, \ldots, \nu_n)$ and every sequence $(\varepsilon_k^i)_{k=1}^{\infty}$ of real numbers converging to zero, the sequence $(\gamma_k^i)_{k=1}^{\infty}$ defined by $\gamma_k^i = (1 - \varepsilon_k^i) \cdot \gamma_i + \varepsilon_k^i \cdot \nu_i$ for every $i$ is such that

$$\lim_{k \to \infty} \rho_w(\gamma_k^i, BR_i(\gamma^k)) = 0,$$

in which $\gamma^k = (\gamma_1^k, \ldots, \gamma_n^k)$.

A behavioral strategy profile is a distributional strictly perfect BNE if it is both distributionally strictly perfect and a BNE.

The above equilibrium concept may not exist even in finite games. DSPE is equivalent to strictly perfect equilibrium in finite normal form games and we

---

6See Aliprantis and Border [1]. Let $\mu$ and $\nu$ be two probability measures on a $\sigma$-field $\Sigma$. Weak (Prohorov) metric is defined for $\mu$ and $\nu$ by $\rho_w(\mu, \nu) = \inf \{\varepsilon > 0 \mid \forall B \in \Sigma : \mu(B) < \nu(B^c) + \varepsilon \text{ and } \nu(B) < \mu(B^c) + \varepsilon\}$.

7The proof is available upon request.
see that in example 1.5.5 in van Damme [33] there is no strictly perfect equilibrium. However, we can define a weaker version of DSPE that is called distributional perfect equilibrium and prove that for every Bayesian game, the set of distributional BNEs is not empty under two additional assumptions: First, the measure \( \mu \) to be absolutely continuous with respect to \( \hat{\mu} = \times_{i=1}^{n} \mu_i \). Second, payoffs to be continuous in actions for every type, that is, \( \pi_i(t, \cdot) : A \rightarrow \mathbb{R} \) is continuous for every \( t \in T \).

A behavioral strategy profile \( \beta = (\beta_1, \ldots, \beta_n) \) is called distributional perfect if for the induced distributional strategy profile \( \gamma = (\gamma_1, \ldots, \gamma_n) \) there exists a sequence \( (\gamma^k_{i})_{k=1}^{\infty} = (\gamma^k_1, \ldots, \gamma^k_n)_{k=1}^{\infty} \) of completely mixed distributional strategy profiles such that for every player \( i \) it holds that

\[
(i) \lim_{k \rightarrow \infty} \rho_w(\gamma^k_i, \gamma_i) = 0,
(ii) \lim_{k \rightarrow \infty} \rho_w(\gamma^k_i, BR_i(\gamma^k)) = 0.
\]

A behavioral strategy profile is a distributional perfect BNE if it is both distributional perfect and a BNE.

We can show that the set of distributional perfect BNEs is nonempty. The proof can be found in the appendix and it is based on the existence proof in Simon and Stinchcombe [?] and Milgrom and Weber [28].

# 4 Interdependent value auctions

In this section we study symmetric second-price auctions with interdependent values for two bidders and apply strict distributional perfection in this context.

Let \( \Lambda \) be the following sealed-bid second-price auction for a single indivisible object. There are 2 bidders. Prior to bidding, each bidder \( i \) receives a private signal \( t_i \in T_i = [\underline{t}, \overline{t}] \), which is called the type of bidder \( i \). Signals are drawn independently according to a distribution \( \mu_i \). Bidder \( i \)'s valuation of the object

---

8Let \( \mu \) and \( \nu \) be two measures on a \( \sigma \)-field \( \Sigma \). The measure \( \mu \) is absolutely continuous with respect to \( \nu \) if for every \( G \in \Sigma \), \( \nu(G) = 0 \) implies \( \mu(G) = 0 \).

9In Milgrom and Weber [28] this property is called Absolutely Continuous Information. Absolutely continuous information is a rather weak assumption. It is satisfied when the type spaces are finite, or when each player’s type is drawn independently from his type space.
may depend on both types and is denoted by $v_i(t_1, t_2)$. Each bidder $i$, after observing his own type, submits a bid from a set $A_i$ independently of his opponent, where $A_i$ is sufficiently large in the sense that it contains the range of $v_i$. Given the bids submitted by the bidders, the highest bid wins the auction for the price equal to the second highest bid. Suppose the following assumptions in this auction:

1. $v_1(t_1, t_2) = v_2(t_2, t_1)$.
2. $v_i$ is continuously differentiable.
3. $\frac{\partial v_i}{\partial t_i}(t_1, t_2) > \frac{\partial v_i}{\partial t_j}(t'_1, t'_2) \geq 0$ for all $t_1, t_2, t'_1, t'_2$.

Condition [3], without the non-negativity requirement, is a strong form of the single crossing condition (SCC). It is for example satisfied by linear valuation functions $v_1(t_1, t_2) = a \cdot t_1 + b \cdot t_2$ for non-negative numbers $a, b$ with $a > b$.

Condition [3] has several important consequences, of which we mention three. First, it implies that $v_i$ is strictly increasing in $t_i$ and increasing in $t_j$ for $j \neq i$. Second, in combination with condition [1] it ensures that the bidder with the higher type has the higher ex post valuation. Third, the pair of strategies $(b_1, b_2)$ in which $b_i = v_i(t_i, t_i)$ for every bidder $i$, is an efficient equilibrium.

More information on SCC can be found in for example Krishna [20].

In the auction $\Lambda$ the set of available bids for bidder $i$ is an interval $A_i = [a, \bar{a}]$, where $v_i(t, \bar{t}) = a$ and $v_i(\bar{t}, \bar{t}) = \bar{a}$, for $i = 1, 2$. In this auction bidders do not have any dominant strategy (disregarding the classic case of private values for a moment), and there are many equilibria.

In the next two subsections we show how DSPE can be used to refine among these equilibria. We show that the efficient equilibrium introduced in Milgrom [27] is a DSPE. In contrast with this, we show that within the class of inefficient equilibria introduced in Birulin [12] none of them are DSPE.

In the event of a tie for the highest bid, the winner is chosen according to a probability distribution which may depend on the identity of the highest bidder and his type as well. In our setting, a tie will only occur with probability 0, even given the type of one bidder, which makes the specification of this tie-breaking rule irrelevant for the expected profits.

Thus, common value auctions do not satisfy SSC.

An outcome in the auction is called efficient if the winner is the one with the highest ex post valuation.
4.1 Equilibrium refinement I: equilibrium selection

Milgrom [27] introduced the following class of strategy pairs. Let $\phi : [t, \bar{t}] \rightarrow [\underline{t}, \bar{t}]$ be an increasing bijection. Define the strategy pair $(b_1, b_2)$ by $b_1(t_1) = v_1(t_1, \phi(t_1))$ and $b_2(t_2) = v_2(\phi^{-1}(t_2), t_2)$. In Proposition 4.1 we show that within this class of continuous strategy pairs only one focal strategy pair is a BNE. In Proposition 4.2 we show that this focal BNE is a DSPE. We also argue that the BNE is not in dominant strategies (unless we are considering the special case of pure private values). Thus, dominance cannot be used as a means to refine among equilibria in this context.

**Proposition 4.1** Let $(b_1, b_2)$ be the strategy pair defined above. Then, $(b_1, b_2)$ is a BNE if and only if $\phi(t) = t$.

**Proof. A.** Suppose that $\phi(t) = t$ for every $t \in [t, \bar{t}]$. We show that $(b_1, b_2)$ is a BNE. Define $w(t) = v_1(t, t)$ for every $t \in [\underline{t}, \bar{t}]$. Clearly $w$ is an increasing bijection from $[\underline{t}, \bar{t}]$ to $[\underline{a}, \bar{a}]$. Suppose that bidder 1 uses bid function $b_1(t_1) = v_1(t_1, t_1)$. The expected payoff of bidder 2 having type $t_2$ and bidding $p_2$ is:

$$\Pi_2^b(b_1, p_2 \mid t_2) = \int_\underline{t}^{w^{-1}(p_2)} (v_2(t_1, t_2) - v_1(t_1, t_1)) \mu_1(dt_1).$$

As $v_2$ is strictly increasing in the second argument, and moreover

$$v_2(t_1, t_2) - v_1(t_1, t_1) = v_2(t_1, t_2) - v_2(t_1, t_1),$$

it is clear that $v_2(t_1, t_2) - v_1(t_1, t_1) \geq 0$ precisely when $t_1 \leq t_2$. So, the maximum of $\Pi_2^b(b_1, p_2 \mid t_2)$ is obtained by choosing $p_2$ such that $w^{-1}(p_2) = t_2$. Hence, the optimal bidding strategy is

$$b_2(t_2) = w(t_2) = v_1(t_2, t_2) = v_2(t_2, t_2).$$

**B.** Conversely, suppose that $\phi(t) \neq t$ for some $t \in [\underline{t}, \bar{t}]$. We show that $(b_1, b_2)$ is not a BNE.

If $\phi(t) > t$. Then

$$v_2(t, \phi(t)) > v_2(\phi(t), t).$$
by assumption [3] in our auction model. So, \( v_2(t, \phi(t)) \neq v_2(\phi(t), t) \). A similar argument applies for the case \( \phi(t) < t \).

Thus, there is a type \( \hat{t}_2 \) with \( v_2(\hat{t}_2, \phi(\hat{t}_2)) \neq v_2(\phi(\hat{t}_2), \hat{t}_2) \). Write \( t_2^* = \phi(\hat{t}_2) \).

We first argue that \( t_2^* \neq \phi^{-1}(t_2^*) \). Suppose that \( t_2^* = \phi^{-1}(t_2^*) \). Then \( \phi(t_2^*) = t_2^* = \phi(\hat{t}_2) \), so that \( \hat{t}_2 = t_2^* = \phi(\hat{t}_2) \). Then \( v_2(\hat{t}_2, \phi(\hat{t}_2)) = v_2(\phi(\hat{t}_2), \hat{t}_2) \). This contradicts our assumption.

Now suppose that bidder 1 uses bid function \( b_1(t_1) = v_1(t_1, \phi(t_1)) \). The expected payoff of bidder 2 having type \( t_2^* \) and bidding \( p_2 \) is:

\[
\Pi_2^b(b_1, p_2 \mid t_2^*) = \int_{t_2^*}^{w^{-1}(p_2)} (v_2(t_1, t_2^*) - v_1(t_1, \phi(t_1))) \, \mu_1(dt_1).
\]

Define \( f(t_1) = v_2(t_1, t_2^*) - v_1(t_1, \phi(t_1)) \). Since

\[
f(t_1) = v_2(t_1, t_2^*) - v_1(t_1, \phi(t_1)) = v_2(t_1, t_2^*) - v_2(\phi(t_1), t_1)
\]

it follows that

\[
f'(t_1) = \frac{\partial v_2}{\partial t_1} - \frac{\partial v_2}{\partial t_2} \cdot \phi'(t_1) = \frac{\partial v_2}{\partial t_1} - \frac{\partial v_2}{\partial t_2} - \frac{\partial v_2}{\partial t_1} \cdot \phi'(t_1).
\]

So, by assumption [3], \( f \) is strictly decreasing. Further,

\[
f(\phi^{-1}(t_2^*)) = v_2(\phi^{-1}(t_2^*), t_2^*) - v_2(\phi(\phi^{-1}(t_2^*)), \phi^{-1}(t_2^*))
\]

\[
= v_2(\phi^{-1}(t_2^*), t_2^*) - v_2(\phi^{-1}(t_2^*), \phi^{-1}(t_2^*))
\]

Now recall that \( t_2^* \neq \phi^{-1}(t_2^*) \). If \( t_2^* > \phi^{-1}(t_2^*) \). Then \( f(\phi^{-1}(t_2^*)) > 0 \). As the function \( f \) is strictly decreasing, we can conclude that \( \Pi_2^b(b_1, p_2 \mid t_2^*) \) is maximized at a type \( \tilde{t}_1 > \phi^{-1}(t_2^*) \). If \( t_2^* < \phi^{-1}(t_2^*) \), then \( \Pi_2^b(b_1, p_2 \mid t_2^*) \) is maximized at a type \( \tilde{t}_1 < \phi^{-1}(t_2^*) \). In both cases, \( b_2(t_2^*) \neq v_2(\phi^{-1}(t_2^*), t_2^*) \).

Thus, this class of strategy pairs contains only one BNE, namely the pair for which \( b_i = v_i(t_i, t_i) \) for every \( i = 1, 2 \). We show that the BNE is also distributionally perfect.

**Proposition 4.2** The strategy pair \((b_1, b_2)\), where \( b_i = v_i(t_i, t_i) \) for \( i = 1, 2 \) is a DSPE. Further, \( b_i \) is undominated for every bidder \( i \).
Proof. We prove that the BNE \((b_1, b_2)\) is a DSPE. Let \(\gamma_i\) be the distributional strategy induced by \(b_i\). For every \(k \in \mathbb{N}\) define \(\gamma^k_i = (1 - \varepsilon_k) \cdot \gamma_1 + \varepsilon_k \cdot \nu_1\), where \(\varepsilon_k \in (0, 1), \varepsilon_k \to 0\) as \(k \to \infty\), and \(\nu_1\) is any completely mixed distributional strategy on the product space \(T_1 \times A_1\).\(^{13}\)

It suffices to show that for each \(k\) there is a strategy \(\eta^k_2 \in BR_2(\gamma^k_1)\) such that

\[
\lim_{k \to \infty} \rho^w(\gamma_2, \eta^k_2) = 0.
\]

The proof is in two steps. Let \(t_2\) be a type of bidder 2.

A. We show that \(p_2 = b_2\) is the unique maximizer of \(F(p_2) = \Pi_2^k(b_1, p_2 \mid t_2)\) and \(F(p_2)\) is continuous. Recall that

\[
F(p_2) = \Pi_2^k(b_1, p_2 \mid t_2) = \int_{t}^{w^{-1}(p_2)} (v_2(t_1, t_2) - v_1(t_1, t_1)) \mu_1(dt_1),
\]

where \(w\) denotes the strictly increasing function \(w(t) = v_1(t, t)\). Since the integrand

\[
v_2(t_1, t_2) - v_1(t_1, t_1) = v_2(t_1, t_2) - v_2(t_1, t_1)
\]

is continuous and strictly decreasing in \(t_1\), and \(w^{-1}(p_2)\) is continuous and strictly increasing in \(p_2\), it is clear that \(F(p_2)\) is continuous. Further, since the integrand is zero at \(t_1 = t_2\), it is also clear that \(F(p_2)\) has a unique maximum location at value \(p_2\) with \(w^{-1}(p_2) = t_2\). Thus,

\[
p_2 = w(t_2) = v_1(t_2, t_2) = v_2(t_2, t_2) = b_2(t_2)
\]

is the unique maximizer of \(F(p_2)\).

B. We prove that for each \(k\) there is a strategy \(\eta^k_2 \in BR_2(\gamma^k_1)\) such that

\[
\lim_{k \to \infty} \rho^w(\gamma_2, \eta^k_2) = 0.
\]

Let \(p_2\) be a pure behavioral strategy for bidder 2 and \(\sigma_2\) be the distributional strategy induced by \(p_2\). Then, the expected payoff of bidder 2 choosing \(\sigma_2\) against

\(^{13}\)That is, \(\eta_1\) is a completely mixed measure on \(T_1 \times A_1\) and \(\eta_1(S_1 \times A_1) = \mu_1(S_1)\) for every \(S_1 \in T_1\).
\( \gamma_1^k \) is:
\[
\Pi_2(\gamma_1^k, \sigma_2) = \int \int_{p_2 > a_1} (v_2(t_1, t_2) - a_1) \, d\gamma_1^k \, d\sigma_2
\]
\[
= \int_{T_2} \int_{p_2 > a_1} (v_2(t_1, t_2) - a_1) \, d\gamma_1^k \, \mu_2(dt_2)
\]
\[
= \int_{T_2} \Pi_2^b(\beta_1^k, p_2 \mid t_2) \, \mu_2(dt_2),
\]

where \( \beta_1^k \) is a behavioral strategy corresponding to \( \gamma_1^k \) for every \( k \in \mathbb{N} \). Let \( b_2^k \) be any maximizer of \( \Pi_2^b(\beta_1^k, p_2 \mid t_2) \) for every \( k \), where
\[
\Pi_2^b(\beta_1^k, p_2 \mid t_2) = (1 - \varepsilon_k) \int_{p_2 > a_1} (v_2(t_1, t_2) - a_1) \, d\gamma_1
\]
\[
+ \varepsilon_k \int_{p_2 > a_1} (v_2(t_1, t_2) - a_1) \, d\eta_1
\]
\[
= (1 - \varepsilon_k) \int_{p_2 > a_1} (v_2(t_1, t_2) - v_1(t_1, \phi(t_1))) \, \mu_1(dt_1)
\]
\[
+ \varepsilon_k \int_{p_2 > a_1} (v_2(t_1, t_2) - a_1) \, d\eta_1
\]
\[
= (1 - \varepsilon_k) \Pi_2^b(b_1, p_2 \mid t_2) + \varepsilon_k \int_{p_2 > a_1} (v_2(t_1, t_2) - a_1) \, d\eta_1.
\]

Note that the maximizer \( b_2^k \) exists, because \( \Pi_2^b(\beta_1^k, p_2 \mid t_2) \) is a continuous function over the compact set \( T_2 \). We know that \( p_2 = b_2 \) is the unique maximizer of \( \Pi_2^b(b_1, p_2 \mid t_2) \). Thus, according to Lemma 6.1, we have that \( b_2^k \) converges to \( b_2 \) when \( k \to \infty \). Now, let \( \eta_2^k \) be the induced distributional strategy by \( b_2^k \). As \( b_2^k \in BR_2(t_2, \beta_1^k) \), then we have \( \eta_2^k \in BR_2(\gamma_2^k) \). Furthermore, since \( b_2^k(t_2) \) converges to \( b_2(t_2) \) for every \( t_2 \in T_2 \), we conclude \( \rho(v_i, \eta_2^k) \to 0 \) when \( k \to \infty \).

With a similar argument for bidder 2, we can conclude that \( (b_1, b_2) \) is a DSPE.

Finally, as bidding \( b_2 \) is the unique best response for bidder 2 against \( b_1 \), we conclude that \( b_2 \) is undominated. A similar argument holds for \( b_1 \).

In the next proposition we prove that this strategy pair is not in dominant strategies, unless we are in the pure private values case.

**Proposition 4.3** Suppose that the valuation function \( v_i \) is not private values. Then the strategy \( b_i = v_i(t_i, t_i) \) is not dominant.
Proof. Assume that $v_1$ is not private values. We prove that $b_1 = v_1(t_1, t_1)$ is not a dominant strategy. In particular, we construct a bid function $d_2$ for bidder 2 and an open set $I_1$ of types for bidder 1, such that there is a bid $a \in A_1$ for bidder 1 that yields a strictly higher payoff against $d_2$ than $b_1(t_1) = v_1(t_1, t_1)$ for any type $t_1 \in I_1$. This then is a sufficient argument.

Since $v_1$ is not private values, by condition [3] there is an $s_1$ in $T_1$ and an open set $B_2$ in $T_2$ such that the map $v_1(s_1, \cdot) : T_2 \to A_2$ is strictly increasing on $B_2$.

Suppose that $B_2 \cap [t, s_1)$ is not empty. Then we may assume that $t_2 < s_1$ for all $t_2 \in B_2$. Then, for every $t_2 \in B_2$ we have $v_1(s_1, t_2) < v_1(s_1, s_1)$. Define

\[ y(t_1) = \int_t^{s_1} v_1(t_1, t_2) \mu_2(dt_2). \]

Since $v_1(s_1, t_2) \leq v_1(s_1, s_1)$ for every $t_2 < s_1$, and $v_1(s_1, t_2) < v_1(s_1, s_1)$ for every $t_2 \in B_2$, it follows that

\[ y(s_1) = \int_t^{s_1} v_1(s_1, t_2) \mu_2(dt_2) < v_1(s_1, s_1) = b_1(s_1). \]

So, $y(s_1) < b_1(s_1)$. Since $t < s_1$ and $b_1$ is strictly increasing and continuous, there is $u_1 < s_1$ with $b_1(u_1) > y(s_1)$.

Take $I_1 = (u_1, s_1)$. Then, since $y$ and $b_1$ are strictly increasing functions, for every $t_1 \in I_1$ we have $y(t_1) < y(s_1) < b_1(u_1) < b_1(t_1)$.

Take the (constant) bid function $d_2(t_2) = b_1(u_1)$ for every $t_2 \in T_2$ for bidder 2. Take $t_1 \in I_1$ and any $t_2 \in T_2$. Since bidder 1 bids according to $b_1(t_1)$, and $d_2(t_2) = b_1(u_1) < b_1(t_1)$, bidder 1 wins the auction. Consequently, his expected payoff against $d_2$ is

\[ \int_{T_2} (v_1(t_1, t_2) - b_1(u_1)) \mu_2(dt_2) = y(t_1) - b_1(u_1) < 0. \]

A similar argument applies to the case where $B_2 \cap (s_1, t]$ is not empty. This concludes the proof.
4.2 Equilibrium refinement II: discarding equilibria

Birulin in [12] on page 678 introduces the following class of discontinuous asymmetric strategy profiles. \(^{14}\) For given \(s_1, s_2 \in [t, \bar{t}]\) with \(s_1 < s_2\), define the strategy pair \((\hat{b}_1, \hat{b}_2)\) by

\[
\hat{b}_1(t_1) = \begin{cases} 
  v_1(s_2, t_1) & \text{if } t_1 \in [s_1, s_2] \\
  v_1(t_1, t_1) & \text{otherwise,}
\end{cases}
\]

\[
\hat{b}_2(t_2) = \begin{cases} 
  v_2(t_2, s_1) & \text{if } t_2 \in [s_1, s_2] \\
  v_2(t_2, t_2) & \text{otherwise.}
\end{cases}
\]

Birulin [12] shows that, under SCC, the strategy pair \((\hat{b}_1, \hat{b}_2)\) is an ex post equilibrium that allocates the object inefficiently. Moreover, this equilibrium may be in undominated strategies. In the equilibrium above, bidder 1 overbids and bidder 2 underbids for types between \(s_1\) and \(s_2\). In the proposition below we prove that strict distributional perfection rules out these equilibria. \(^{15}\) The intuition behind this result is that if there is a small chance that one bidder bids \(b_i = v_i(t_i, t_i)\) for \(t_i \in [s_1, s_2]\) instead of \(\hat{b}_i\), then \(\hat{b}_j\) will no longer be a best response for bidder \(j\), where \(j \neq i\). In other words, underbidding (overbidding) would not be a best response against overbidding (underbidding) if even with a tiny probability one bidder bids according to the efficient bidding strategy \(b_i = v_i(t_i, t_i)\). As an extreme case, we see that the ex post equilibrium \(^{16}\) in which one bidder always bids \(\pi\) and the other bidder always bids \(\bar{a}\) is not a DSPE.

**Proposition 4.4** The equilibrium \((\hat{b}_1, \hat{b}_2)\) is not a DSPE.

**Proof.** A. First we show that if bidder 1 bids according to a behavioral completely mixed strategy that converges to \(\hat{b}_1\), then bidding \(\hat{b}_2\) for bidder 2 is not optimal on a set of types with strictly positive measure. Assume bidder 1 bids according to \(\beta^k_1\) strategy:

\[
\beta^k_1(t_1) = (1 - 2\delta_k)\hat{b}_1(t_1) + \delta_k b_1(t_1) + \delta_k \eta_1(t_1),
\]

\(^{14}\)Our presentation of the bid functions is slightly different from Birulin [12]. He presents the bid functions in terms of the opponents' valuations. We do this for each bidder in terms of his own valuations.

\(^{15}\)In finite games, an equilibrium is perfect if and only if it is in undominated strategies. This is in general not true in games with infinite action spaces, even if undominatedness is replaced by limit undominatedness. Bajoori et. al in [5] give a detailed analysis on this issue. They prove that, in normal form games under appropriate assumptions, every weak perfect equilibrium is in limit undominated strategies, but not vice versa. Example 6 in Bajoori et. al [5] provides the relevant counter-example.

\(^{16}\)This kind of equilibrium is often called a wolf and sheep equilibrium.
where \( b_1(t_1) = v_1(t_1, t_1) \), \( \eta_1 \) is any completely mixed strategy for bidder 1, and \( \delta_k \to 0 \) when \( k \to \infty \). Clearly, \( \beta_1^k \) converges to \( \hat{b}_1 \) when \( k \to \infty \) pointwisely. We show that bidding \( \tilde{b}_2(t_2) = v_2(s_1, s_2) \) is strictly better than bidding \( \tilde{b}_2 \) for \( t_2 \in (s_1, s_2) \) for bidder 2. Note that \( \tilde{b}_2 = v_2(t_2, s_1) \) for \( t_2 \in (s_1, s_2) \). Define \( w(t) = v_1(t, t) \). Bidder 2’s expected payoff by bidding \( \tilde{b}_2 = v_2(t_2, s_1) \) against \( \beta_1^k \) for every \( t_2 \in (s_1, s_2) \) is

\[
\Pi_2^k(\beta_1^k, \tilde{b}_2 | t_2) = (1 - 2\delta_k) \int_{s_1}^{s_1} (v_2 - v_1(t_1, t_1)) d\mu(t_1) + \delta_k \int_{s_1}^{w^{-1}(\tilde{b}_2)} (v_2 - v_1(t_1, t_1)) d\mu(t_1) + \delta_k \int_{\tilde{b}_2 > a_1} (v_2 - a_1) d\eta_1(t_1).
\]

Moreover, bidder 2’s expected payoff by bidding \( \tilde{b}_2 = v_2(s_1, s_2) \) against \( \beta_1^k \) for every \( t_2 \in (s_1, s_2) \) is

\[
\Pi_2^k(\beta_1^k, \tilde{b}_2 | t_2) = (1 - 2\delta_k) \int_{s_1}^{s_1} (v_2 - v_1(t_1, t_1)) d\mu(t_1) + \delta_k \int_{s_1}^{w^{-1}(\tilde{b}_2)} (v_2 - v_1(t_1, t_1)) d\mu(t_1) + \delta_k \int_{\tilde{b}_2 > a_1} (v_2 - a_1) d\eta_1(t_1).
\]

By condition [3] in our auction model, we have \( \tilde{b}_2(t_2) = v_2(t_2, s_1) < v_2(s_1, s_2) = \tilde{b}_2(t_2) \). Because the function \( w \) is a strictly increasing bijection, we can have \( w^{-1}(\tilde{b}_2) < w^{-1}(\tilde{b}_2) \). Hence, \( \Pi_2^k(\beta_1^k, \tilde{b}_2 | t_2) < \Pi_2^k(\beta_1^k, \tilde{b}_2 | t_2) \). Therefore, \( \tilde{b}_2(t_2) \notin BR_2(t_2, \beta_1^k) \) for \( t_2 \in (s_1, s_2) \) and every \( k \).

B. To show that \( (\tilde{b}_1, \tilde{b}_2) \) is not a DSPE, we show that for the induced distributional strategy profile \( (\gamma_1^k, \gamma_2^k) \)

\[
\exists \nu = (\nu_1, \nu_2) \text{ and } \exists (\varepsilon_k)_{k=1}^{\infty} = (\varepsilon_1^k, \varepsilon_2^k)_{k=1}^{\infty}
\]

in which \( \nu_1, \nu_2 \) are completely mixed distributional strategies and \( \varepsilon_1^k, \varepsilon_k^2 \to 0 \) when \( k \to \infty \) such that for \( \gamma_1^k = (1 - \varepsilon_k^1) \hat{\gamma}_1 + \varepsilon_k^1 \nu_1 \) and \( \gamma_2^k = (1 - \varepsilon_k^2) \hat{\gamma}_2 + \varepsilon_k^2 \nu_2 \) we have

\[
p^w(\gamma_2^k, BR_2(\gamma_1^k)) > \varepsilon.
\]
for an $\varepsilon > 0$. Let $\varepsilon_k^2 = 0$ for every $k$, then we have $\gamma_k^k = \hat{\gamma}_2$ for every $k$. Therefore, it is enough to prove that

$$\rho^w(\hat{\gamma}_2, BR_2(\gamma_1^k)) > \varepsilon.$$ 

for an $\varepsilon > 0$.

Let $\tilde{\beta}_2^k$ be such that $\tilde{\beta}_2^k(t_2, \cdot) \in BR_2(t_2, \beta_1^k)$ for every $t_2 \in T_2$ and every $k$. In part A we show that $\hat{b}_2(t_2) \notin BR_2(t_2, \beta_1^k)$ for $t_2 \in (s_1, s_2)$. As the function $\Pi_t^2(\beta_1^k, p_2 \mid t_2)$ is continuous in $p_2$, then any point in a neighborhood of $\hat{b}_2(t_2)$ is not a best response against $\beta_1^k$ either. Therefore, there are $\varepsilon > 0$ and $\lambda > 0$ such that if for every $t_2 \in (s_1, s_2)$ we define

$$B_{t_2} = \left(\hat{b}_2(t_2) - \lambda, \hat{b}_2(t_2) + \lambda\right),$$

then we have

1. $s_1 + \varepsilon < s_2 - \varepsilon$,

2. $\mu_2(s_1 + \varepsilon, s_2 - \varepsilon) > \varepsilon$,

3. $\tilde{\beta}_2^k(t_2, (B_{t_2})^\varepsilon) = 0$.

Now, let $B = \{(t_2, a_2) \mid t_2 \in (s_1, s_2 - \varepsilon), a_2 \in B_{t_2}\}$. Clearly, $B$ is measurable. Also, let $\tilde{\gamma}_2^k$ and $\gamma_1^k$ be the distributional strategies induced by $\tilde{\beta}_2^k$, and $\beta_1^k$ respectively. One can write

$$\gamma_1^k = (1 - \varepsilon_k^1)\hat{\gamma}_1 + \varepsilon_k^1 \nu_1$$

in which $\varepsilon_k^1 = 2\delta_k$ and $\nu_1$ is the completely mixed distributional strategy induced by behavioral strategies $b_1$ and $\eta_1$. Moreover, we have $\tilde{\gamma}_2^k \in BR_2(\gamma_1^k)$. As $\hat{b}_2(t_2)(B_{t_2}) = 1$ and $\tilde{\beta}_2^k(t_2, (B_{t_2})^\varepsilon) = 0$, we have

$$\tilde{\gamma}_2^k(B) = \int_{s_1 + \varepsilon}^{s_2 - \varepsilon} \hat{b}_2(t_2)(B_{t_2})\mu_2(dt_2) = \mu_2(s_1 + \varepsilon, s_2 - \varepsilon)$$

and

$$\tilde{\gamma}_2^k(B^\varepsilon) = \int_{s_1}^{s_2} \tilde{\beta}_2^k(t_2, (B_{t_2})^\varepsilon)\mu_2(dt_2) = 0.$$}

\(^{17}\text{Notice that } \hat{b}_2(t_2)\text{ is interpreted as a behavioral strategy that assigns probability } 1 \text{ on } \hat{b}_2(t_2)\text{ for every } t_2 \in T_2.\)
Consequently,
\[ \hat{\gamma}_2(B) > \hat{\gamma}_2^k(B^c) + \varepsilon, \]
then \( \rho^w(\hat{\gamma}_2, \hat{\gamma}_2^k) > \varepsilon \) for every \( k \). Hence, we have \( \rho^w(\hat{\gamma}_2, BR_2(\gamma_1^k)) > \varepsilon \) for every \( k \). The proof is complete. 

5 Conclusion

We introduce the concept of distributional strictly perfect equilibrium (DSPE) as a refinement of Bayesian Nash equilibrium. We apply this concept in symmetric second-price auctions with interdependent values for two bidders to select the efficient BNE, and to rule out the less intuitive equilibria. In particular, we show that the efficient BNE discussed in Milgrom [27] is a DSPE. We also show that DSPE rules out all the discontinuous, inefficient, equilibria introduced by Birulín [12]. These equilibria feature overbidding and underbidding. Intuitively, we show that underbidding (overbidding) cannot be a best response against overbidding (underbidding), if one bidder bids according to the efficient equilibrium with a small probability.

6 Appendix

Lemma 6.1 Let \( f \) and \( g \) be two continuous functions on an interval \([a, b]\). For every \( k \in \mathbb{N} \), define
\[
f^k(x) = (1 - \varepsilon_k) f(x) + \varepsilon_k g(x)
\]
in which \( \varepsilon_k \to 0 \) as \( k \to \infty \). Let \( \gamma^k \) be a maximizer of \( f^k \) for every \( k \). Suppose that function \( f \) has a unique maximizer \( x = y \), then \( \gamma^k \to y \) when \( k \to \infty \).

Proof. We have \( f^k \) converges to \( f \) uniformly. For every \( k \), \( \gamma^k \) is any maximizer of \( f^k \), which exists since \([a, b]\) is compact. Consider any convergent subsequence \( \{\gamma^{k_t}\}_{t=1}^\infty \) of the sequence \( \{\gamma^k\}_{k=1}^\infty \). Assume \( \gamma^{k_t} \) converges to \( y' \in [a, b] \) when \( k_t \to \infty \). Then we have \( f^{k_t}(\gamma^{k_t}) \geq f^{k_t}(x) \) for every \( x \in [a, b] \). Now, by uniform convergence, we have \( f^{k_t}(\gamma^{k_t}) \to f(y') \) and \( f^{k_t}(x) \to f(x) \), when \( k_t \to \infty \).
Therefore, we have $f(y') \geq f(x)$ for every $x \in [a, b]$. Then, by uniqueness of maximizer we have $y' = y$. Thus, any convergent subsequence has the same limit $y$. Hence, we conclude that $\{y^k\}_{k=1}^{\infty}$ is convergent and converges to $y$ when $k \to \infty$.

**Existence proof of weak distributional perfect BNE**

To prove that the set of weak distributional perfect BNEs are non-empty, we first show the following lemmas.

**Lemma 6.2** Consider a Bayesian game $\Gamma$. Let $\beta = (\beta_1, \ldots, \beta_n)$ be a behavioral strategy profile and $\gamma = (\gamma_1, \ldots, \gamma_n)$ be the induced distributional strategy profile. If $\gamma_i \in BR_i(\gamma)$, then $\beta_i(t_i, \cdot) \in BR_i(t_i, \beta)$ for $\mu_i$-a.e. $t_i \in T_i$.

**Proof.** Suppose for every player $i$ that $\gamma_i \in BR_i(\gamma)$. Then, we have

$$
\Pi_i(\gamma_i, \gamma_{-i}) \geq \Pi_i(\tilde{\gamma}_i, \gamma_{-i}),
$$

for every player $i$’s distributional strategy $\tilde{\gamma}_i$. Therefore, for every player $i$’s behavioral strategy $\tilde{\beta}_i$ we have

$$
\int_{T_i} \Pi_i^b(\tilde{\beta}_i, \beta_{-i} | t_i) \mu_i(dt_i) \geq \int_{T_i} \Pi_i^b(\tilde{\beta}_i, \beta_{-i} | t_i) \mu_i(dt_i).
$$

(1)

Theorem 9.1 in [34] implies that for every player $i$ there is a behavioral strategy $\zeta_i$ such that $\zeta_i(t_i, \cdot) \in BR_i(t_i, \beta)$ for every $t_i \in T_i$, which means having type $t_i \in T_i$

$$
\Pi_i^b(\zeta_i, \beta_{-i} | t_i) \geq \Pi_i^b(a_i, \beta_{-i} | t_i),
$$

for every $a_i \in A_i$. To prove that $\beta_i(t_i, \cdot)$ for $\mu_i$-a.e. $t_i \in T_i$ is a best response for every player $i$ against $\beta_{-i}$, we show that $\Pi_i^b(\zeta_i, \beta_{-i} | t_i) = \Pi_i^b(\beta_i, \beta_{-i} | t_i)$ for $\mu_i$-a.e. $t_i \in T_i$. Suppose the opposite. First, let

$$
S_i = \{ t_i \in T_i | \Pi_i^b(\zeta_i, \beta_{-i} | t_i) < \Pi_i^b(\beta_i, \beta_{-i} | t_i) \}.
$$

To prove that $\beta_i(t_i, \cdot)$ for $\mu_i$-a.e. $t_i \in T_i$ is a best response for every player $i$ against $\beta_{-i}$, we show that $\Pi_i^b(\zeta_i, \beta_{-i} | t_i) = \Pi_i^b(\beta_i, \beta_{-i} | t_i)$ for $\mu_i$-a.e. $t_i \in T_i$. Suppose the opposite. First, let

$$
S_i = \{ t_i \in T_i | \Pi_i^b(\zeta_i, \beta_{-i} | t_i) < \Pi_i^b(\beta_i, \beta_{-i} | t_i) \}.
$$

18Suppose $(T, M)$ is a measurable space, $X$ is a topological space with the Borel $\sigma$-algebra $B$, and $F$ is a correspondence from $T$ to $X$. Denote $GrF = \{(t, x) : x \in F(t)\}$. Suppose $u : GrF \to R$ is $M \otimes B$ measurable and $u(t, \cdot)$ is upper semicontinuous on $F(t)$ for $t \in T$, and let $\nu(t) = \sup\{u(x, t) : x \in F(t)\}$ for $t \in T$. Theorem 9.1 in [34] states that if $F$ is compact-valued and measurable, $X$ is separable metric, and $u$ is the limit of a decreasing sequence of Carathéodory maps, then $\nu$ is measurable and there exits $f : T \to X$ measurable such that $u(\cdot, f(\cdot)) = \nu$. 
Clearly, $S_i$ is $\mu_i$-measurable. Suppose $\mu_i(S_i) > 0$. Notice that given $t_i \in T_i$, for $\beta_i(t_i, \cdot)$-a.e. $a_i \in A_i$ we have

$$\Pi_i^b(\beta_i, \beta_{-i} | t_i) = \Pi_i^b(a_i, \beta_{-i} | t_i).$$

Then, given $t_i \in S_i$, for $\beta_i(t_i, \cdot)$-a.e. $a_i \in A_i$ we conclude that

$$\Pi_i^b(\zeta_i, \beta_{-i} | t_i) < \Pi_i^b(a_i, \beta_{-i} | t_i),$$

which is a contradiction with optimality of $\zeta_i$ for every type. Second, let

$$S_i = \{ t_i \in T_i | \Pi_i^b(\zeta_i, \beta_{-i} | t_i) > \Pi_i^b(\beta_i, \beta_{-i} | t_i) \}$$

and suppose $\mu_i(S_i) > 0$. Therefore, we have

$$\int_{S_i} \Pi_i^b(\zeta_i, \beta_{-i} | t_i) \mu_i(dt_i) > \int_{S_i} \Pi_i^b(\beta_i, \beta_{-i} | t_i) \mu_i(dt_i).$$

Since $\zeta_i(t_i, \cdot)$ is a best response for every type $t_i \in T_i$ against $\beta_{-i}$, for every $a_i \in A_i$ we have

$$\Pi_i^b(\zeta_i, \beta_{-i} | t_i) \geq \Pi_i^b(a_i, \beta_{-i} | t_i),$$

thus

$$\int_{T_i \setminus S_i} \Pi_i^b(\zeta_i, \beta_{-i} | t_i) \mu_i(dt_i) \geq \int_{T_i \setminus S_i} \Pi_i^b(\beta_i, \beta_{-i} | t_i) \mu_i(dt_i).$$

Consequently,

$$\int_{T_i} \Pi_i^b(\zeta_i, \beta_{-i} | t_i) \mu_i(dt_i) > \int_{T_i} \Pi_i^b(\beta_i, \beta_{-i} | t_i) \mu_i(dt_i).$$

This is a contradiction with (1). Overall, we showed that if there is a subset of $T_i$ on which the expected payoff of player $i$, playing according to $\zeta_i$ is different from playing according to $\beta_i$, the measure of that subset is zero. In other words, for every $i$ there exists a set $S_i \in \mathcal{T}$ with $\mu_i(S_i) = 0$ such that for every $t_i \in T_i \setminus S_i$ we have $\beta_i(t_i, \cdot) \in BR_i(t_i, \beta)$. This completes the proof.

**Lemma 6.3** Consider a Bayesian game $\Gamma$. For every player $i$ and every $\varepsilon > 0$, there is a set $E \in \mathcal{T}$ such that $\mu(E^c) < \varepsilon$ and the family $\{ \pi_i(t, \cdot) | t \in E \}$ is equicontinuous. \(^{19}\)

---

\(^{19}\)A family $\mathfrak{F}$ of real functions on the metric space $X$ is equicontinuous if to every $\varepsilon > 0$ corresponds a $\delta > 0$ such that for every $x, y \in X$ with $d(x, y) < \varepsilon$ we have $|f(x) - f(y)| < \delta$, for every $f \in \mathfrak{F}$. 
Proof. Let \( C(A) \) be the set of all continuous real functions on \( A \). For every player \( i \) consider the mapping \( t \rightarrow \pi_i(t, \cdot) \) which is a measurable mapping from \( T \) to \( C(A) \). Therefore, the probability measure \( \mu \) on \( T \) induces a probability measure on \( C(A) \), say \( \eta \). Since \( A \) is compact metric space, then \( C(A) \) is a compact separable metric space that implies that \( \eta \) is tight. That means, for every \( \varepsilon > 0 \) there is a compact set \( G \subset C(A) \) with \( \eta(G^c) < \varepsilon \). Let \( E \) be the inverse image of \( G \) under the above mapping. Then we have \( \mu(E^c) < \varepsilon \). Since \( G \) is compact, by Ascoli-Arzela theorem one concludes that \( G \) is equicontinuous. This completes the proof.

In the following lemma, we use the following metrics: For every \( t, s \in T \) and \( a, b \in A \), define

\[
d_T(t, s) = \left( \sum_{i=1}^{n} d^2_{T_i}(t_i, s_i) \right)^{\frac{1}{2}}, \quad d_A(a, b) = \left( \sum_{i=1}^{n} d^2_{A_i}(a_i, b_i) \right)^{\frac{1}{2}}.
\]

Moreover, let \( d_{T \times A}((t, a), (s, b)) = \left( d^2_T(t, s) + d^2_A(a, b) \right)^{\frac{1}{2}}. \)

Lemma 6.4 Consider a Bayesian game \( \Gamma \). For every player \( i \) and every \( \varepsilon > 0 \), there is a set \( K \in \mathcal{T} \) with \( \mu(K^c) < \varepsilon \) such that \( \pi_i \) is continuous on \( K \times A \) for every \( i \).

Proof. Take \( \varepsilon > 0 \) and \( a \in A \). First we prove that there is a set \( K \in \mathcal{T} \) with \( \mu(K^c) < \varepsilon \) such that the function \( \pi_i(\cdot, a) : K \rightarrow \mathbb{R} \) is continuous. As \( A \) is separable, it has a countable dense subset, say \( B \). Hence, there exists a sequence \( (a_k)_{k=1}^{\infty} \) in \( B \) such that \( a_k \neq a \), for every \( k \), and \( d_A(a_k, a) \rightarrow 0 \) as \( k \rightarrow \infty \). According to Lusin’s Theorem for the collection of the functions \( \{ \pi_i(\cdot, a_k) \mid k = 1, 2, \ldots \} \) there is a compact set \( E \in \mathcal{T} \) with \( \mu(E^c) < \frac{\varepsilon}{2} \) such that \( \pi_i(\cdot, a_k) \) is continuous on \( E \) for every \( k \in \mathbb{N} \). Moreover, by the lemma 6.3 we know that payoffs are equicontinuous, i.e. there is a set \( F \in \mathcal{T} \) with \( \mu(F^c) < \frac{\varepsilon}{2} \) such that the collection of the functions \( \{ \pi_i(t, \cdot) : A \rightarrow \mathbb{R} \mid t \in F \} \) is equicontinuous. Let \( K = E \cap F \). It is clear that \( \mu(K^c) < \varepsilon \). Moreover, because \( \pi_i(t, \cdot) \) is continuous in actions for every \( t \in K \), the sequence of the functions \( \{ \pi_i(\cdot, a_k) \}_{k=1}^{\infty} \) converges pointwise to \( \pi_i(\cdot, a) \) as \( k \rightarrow \infty \). It is easy to check that equicontinuity of payoffs and the fact that \( d_A(a_k, a) \rightarrow 0 \) as \( k \rightarrow \infty \), implies that the sequence of functions
(\pi_i(\cdot, a_k))^{\infty}_{k=1} uniformly converges to \pi_i(\cdot, a) on K as k \to \infty. Hence, \pi_i(\cdot, a) : K \to \mathbb{R} is continuous too.

Now, we prove that \pi_i is continuous at (t, a) \in K \times A. We know that \{\pi_i(t, \cdot) \mid t \in K\} is equicontinuous, then there is a \delta_1 > 0 such that if d_A(a, b) < \delta_1, we have |\pi_i(t, a) - \pi_i(t, b)| < \frac{\varepsilon}{2} for every t \in K. Also, we know \pi_i(\cdot, a) is continuous at t \in K, so there is a \delta_2 > 0 such that if d_T(t, s) < \delta_2, we have |\pi_i(t, a) - \pi_i(s, a)| < \frac{\varepsilon}{2}. Let \delta = \min\{\delta_1, \delta_2\}. Then, for every (s, b) \in K \times A with d_{T \times A}((t, a), (s, b)) < \delta we have d_A(a, b) < \delta_1 and d_T(s, t) < \delta_2. Consequently, we have

|\pi_i(t, a) - \pi_i(s, b)| < |\pi_i(t, a) - \pi_i(s, a)| + |\pi_i(s, a) - \pi_i(s, b)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.

This proves that \pi_i is continuous on K \times A.

Lemma 6.4 gives a slightly more general result than Lusin’s theorem. Indeed, Lusin’s theorem would guarantee the existence of a subset of T \times A such that \pi_i is continuous on that subset, while Lemma 6.4 would do that for a set of the form K \times A, where K \subseteq T.

The following lemma is the basic continuity result in Milgrom and Weber [28]. Here we provide a detailed proof of that.

**Lemma 6.5** Consider a Bayesian game \Gamma. For every player i, the function \Pi_i : \Delta \to \mathbb{R} is continuous with respect to the weak metric on \Delta.

**Proof.** To prove that \Pi_i is continuous at a distributional strategy profile \gamma, let (\gamma^k)^{\infty}_{k=1} be a sequence of distributional strategy profiles converging to \gamma with respect to the weak metric as k \to \infty. Now, by Lemma 6.4, for every i and every \delta > 0, there exist a continuous and bounded function v_\delta : T \times A \to \mathbb{R} and a set K \in \mathcal{T} such that \mu(K^c) < \delta, and v_\delta = \pi_i on K \times A.

As \mu(K^c) < \delta, we have \gamma^k(K^c \times A) < \delta, for every k. Similarly, \gamma(K^c \times A) < \delta. Moreover, because v_\delta and \pi_i are bounded, there is an M > 0 such that |\pi_i - v_\delta| \leq M. Hence, for every k we have

\[ \int_{T \times A} |\pi_i - v_\delta| f(t) \, d\gamma^k = \int_{K^c \times A} |\pi_i - v_\delta| f(t) \, d\gamma^k \leq M \cdot \delta. \]
Similarly,
\[
\int_{T \times A} |\pi_i - v_\delta| f(t) \, d\gamma = \int_{K \times A} |\pi_i - v_\delta| f(t) \, d\gamma \leq M \cdot \delta.
\]

Furthermore, as \( f \) is \( \hat{\mu} \)-integrable, there is a sequence \((f_\ell)_{\ell=1}^\infty\) of bounded and continuous functions such that
\[
\int_T |f(t) - f_\ell(t)| \hat{\mu}(dt) \to 0, \quad \text{as } \ell \to \infty. \]
We have
\[
|\Pi_i(\gamma^k) - \Pi_i(\gamma)| = |\int \pi_i f(t) d\gamma^k - \int \pi_i f(t) d\gamma|
\leq \int |\pi_i - v_\delta| f(t) d\gamma^k + \int |f(t) - f_\ell(t)| v_\delta d\gamma^k
+ \int v_\delta f_\ell(t) d\gamma^k - \int v_\delta f_\ell(t) d\gamma
+ \int |f(t) - f_\ell(t)| v_\delta \, d\gamma + \int |\pi_i - v_\delta| f(t) \, d\gamma
\leq 2M\delta + \int |f(t) - f_\ell(t)| v_\delta \, d\gamma^k
+ \int v_\delta f_\ell \, d\gamma^k - \int v_\delta f_\ell \, d\gamma + \int |f(t) - f_\ell(t)| v_\delta \, d\gamma.
\]

Hence, \(|\Pi_i(\gamma^k) - \Pi_i(\gamma)|\) converges to zero when \( k \to \infty, \ell \to \infty, \) and \( \delta \to 0. \)
This proves that \( \Pi_i \) is continuous at \( \gamma. \)

We prove the existence of a weak distributional perfect BNE, with the help of
the following fixed point theorem from Glicksberg [16].

**Theorem 6.6** Let \( S \) be a nonempty compact and convex subset of a locally convex Hausdorff space. Let \( F : S \to S \) be an upper hemicontinuous correspondence with nonempty and convex values. Then, \( F \) has a fixed point.

**Theorem 6.7** The set of distributional perfect BNEs is nonempty.

First we prove that the set of distributional strategy profiles satisfying the conditions in the definition of distributional perfect equilibrium is not empty. For every player \( i \), let \( \nu_i \) be a completely mixed distributional strategy. Note that
such a \( \nu_i \) exists, because \( T_i \times A_i \) is separable. Define for every \( k \in \mathbb{N} \)
\[
\Delta_i(k) = \{ \gamma_i \in \Delta_i \mid \gamma_i(B_i) \geq \frac{1}{k} \cdot \nu_i(B_i), \forall B_i \in T_i \otimes A_i \}.
\]
Moreover, let \( \Delta(k) = \times_{i=1}^n \Delta_i(k) \) and \( BR_k^i : \Delta(k) \to \Delta_i(k) \) be the best response correspondence for player \( i \) restricted to \( \Delta_i(k) \)\(^{20}\) and define the correspondence \( BR_k : \Delta(k) \to \Delta(k) \) to be

\[
BR_k(\gamma) = \times_{i=1}^n BR_k^i(\gamma),
\]

for every \( \gamma \in \Delta(k) \). We verify the conditions of Theorem 6.6 for the correspondence \( BR_k \).

Since \( T \) is a complete and separable metric space, by Theorem 1.4 in [11] \( \mu \) is a tight measure.\(^{21}\) This fact together with the compactness of the set \( A \) implies that \( \Delta_i \) is a tight set of probability measures.\(^{22}\) Now, by Prohorov’s Theorem we conclude that \( \Delta_i \) is a compact metric space with respect to the weak metric. It is easy to see that \( \Delta_i(k) \) is a closed subset of \( \Delta_i \) with respect to the strong metric. Consequently, by Theorem V.3.13 in [15], it is closed with respect to the weak metric. Hence, \( \Delta_i(k) \) is compact with respect to the weak metric. Also, one can easily check that \( \Delta_i(k) \) is convex. Moreover, upper hemicontinuity of \( BR_k \) follows by Lemma 6.5. Now, we apply Theorem 6.6, which leads us to the existence of an equilibrium point \( \gamma^k \in \Delta(k) \), for every \( k \).

Now, define\(^{23}\)

\[
PBR_i(\gamma) = \{ (t_i, a_i) \in T_i \times A_i \mid (t_i, a_i) \in \text{supp}(\gamma_i) \text{ and } \gamma_i \in BR_i(\gamma) \}.
\]

It is clear that \( \gamma^k_i (PBR_i(\gamma^k)) \geq 1 - 1/k \upsilon_i(T_i \times A_i) = 1 - 1/k \). This implies that \( \rho^w(\gamma^k_i, BR_i(\gamma^k)) \leq 1/k \). Furthermore, as \( \Delta_i \) is compact with respect to the weak metric, without loss of generality we can assume that there is a distributional strategy profile \( \gamma \) such that \( \rho^w(\gamma^k, \gamma) \to 0 \) when \( k \to \infty \). Let \( \beta \) be a behavioral strategy profile corresponding to \( \gamma \). Obviously, \( \beta \) is a distributionally perfect.

---

\(^{20}\)Distributional strategy \( \gamma_i \) of player \( i \) is a best response restricted to \( \Delta_i(k) \) against a distributional strategy profile \( \eta_{-i} \in \Delta_{-i}(k) \), if \( \Pi_i(\gamma_i, \eta_{-i}) \geq \Pi_i(\tilde{\gamma}_i, \eta_{-i}) \), for every distributional strategy \( \tilde{\gamma}_i \in \Delta_i(k) \) of player \( i \).

\(^{21}\)Let \( X \) be a metric space with the Borel \( \sigma \)-field \( \Sigma \). A measure \( \mu \) on \( \Sigma \) is tight if for every \( G \in \Sigma, \mu(G) = \sup \{ \mu(K) \mid K \subset G, \text{and } K \text{ compact} \} \). Moreover, Theorem 1.4 in [11] states that if \( X \) is separable and complete, then each probability measure on \( \Sigma \) is tight.

\(^{22}\)A family of probability measures on a metric space is tight if for every \( \varepsilon > 0 \) there is a compact set \( K \) satisfying \( \mu(K) > 1 - \varepsilon \), for every \( \mu \) in the family.

\(^{23}\)Recall that \( \text{supp}(\mu) = \{ x \in X \mid \text{for every open set } G \text{ if } x \in G \text{ then } \mu(G) > 0 \} \).
We show that there is a BNE \( \hat{\beta} \) such that for \( \mu_i \)-a.e. \( t_i \in T_i \) we have \( \beta_i(t_i, \cdot) = \hat{\beta}_i(t_i, \cdot) \), for every player \( i \). Then, since \( \beta \) is distributionally perfect, trivially the BNE \( \hat{\beta} \) is also distributionally perfect. Hence, the set of distributional perfect BNEs is nonempty, as desired.

Since \( \beta = (\beta_1, \ldots, \beta_n) \) is distributionally perfect, by definition we know that for the induced distributional strategy profile \( \gamma = (\gamma_1, \ldots, \gamma_n) \) there exists a sequence \( (\gamma^k)_{k=1}^\infty = (\gamma^k_1, \ldots, \gamma^k_n)_{k=1}^\infty \) of completely mixed distributional strategy profiles such that for every player \( i \) we have \( \lim_{k \to \infty} \rho^w(\gamma^k_i, \gamma_i) = 0 \) and

\[
\lim_{k \to \infty} \rho^w(\gamma^k_i, BR_i(\gamma^k)) = 0.
\]

Therefore, for every \( i \), there exists a sequence \( (\hat{\gamma}^k_i)_{k=1}^\infty \) such that \( \hat{\gamma}^k_i \in BR_i(\gamma^k) \) for every \( k \in \mathbb{N} \) and \( \rho^w(\gamma^k_i, \hat{\gamma}^k_i) \to 0 \) as \( k \to \infty \). By the triangle inequality for \( \rho^w \) this implies for every \( i \) that \( \rho^w(\gamma_i, \hat{\gamma}_i^k) \to 0 \) as \( k \to \infty \). Now, by Lemma 6.5 we conclude the upper hemicontinuity of the correspondence \( BR_i \), for every player \( i \). Hence, \( \gamma_i \in BR_i(\gamma) \).

Now, by Lemma 6.2 we have \( \beta_i(t_i, \cdot) \in BR_i(t_i, \beta) \) for \( \mu_i \)-a.e. \( t_i \in T_i \). In other words, for every \( i \) there exists a set \( S_i \subseteq T_i \) with \( \mu_i(S_i) = 0 \) such that for every \( t_i \in T_i \setminus S_i \) we have \( \beta_i(t_i, \cdot) \in BR_i(t_i, \beta) \). Define \( \hat{\beta} = (\hat{\beta}_1, \ldots, \hat{\beta}_n) \) by letting \( \hat{\beta}_i(t_i, \cdot) = \beta_i(t_i, \cdot) \) for every \( t_i \in T_i \setminus S_i \) and selecting any \( \hat{\beta}_i(t_i, \cdot) \in BR_i(t_i, \beta) \) for every \( t_i \in S_i \). Note that the result of this selection is measurable. Clearly, \( \hat{\beta} \) is a BNE and satisfies the requirement of the definition.

\section*{References}


