# Economics and Computation

ECON 425/563 and CPSC 455/555

Professor Dirk Bergemann and Professor Joan Feigenbaum

Lecture III

In case of any questions and/or remarks on these lecture notes, please contact Oliver Bunn at oliver.bunn(at)yale.edu.

# 1 Bayesian Games

In the last lecture we have introduced the notion of a Bayesian game:

**Definition 1** *A* ***Bayesian game*** $\Gamma_B$ *is given by*

$$\Gamma_B = \{\mathcal{I}, \{A_i\}_{i\in\mathcal{I}}, \{T_i\}_{i\in\mathcal{I}}, p(\cdot), \{u_i(a,t)\}_{i\in\mathcal{I}}\}. \tag{1}$$

Remember the tuple-notation for a system of elements for every player in $\mathcal{I}$:

- The set $A := A_1 \times ... \times A_I$ denotes all possible tuples of actions. A typical element $a \in A$ contains an action $a_i$ for each player $i \in \mathcal{I}$.
- The set $A := A_1 \times ... \times A_I$ denotes all possible tuples of types. A typical element $t \in T$ contains an action $t_i$ for each player $i \in \mathcal{I}$.

Once again, the elements constituting a Bayesian game can be verbalized as follows:

- There is a set $\mathcal{I}$ of players that are involved in the game.
- Each player $i \in \mathcal{I}$ has a set of actions available that she can possibly take.
- Capturing the essence of the Bayesianity of the definition, each player $i \in \mathcal{I}$ has an element of private information about herself, the type-set $T_i$. Only player $i$ knows for sure which element $t_i \in T_i$ actually prevails.
- In contrast to player $i \in \mathcal{I}$, all other players can only draw conclusions about the type $t_i \in T_i$ for player $i$ according to the common prior probability distribution $p$.
- Finally, each player has a utility function $u$. Observe that, at this level of generality, the utility function depends on the pair $(a, t)$, i.e. not only all players' actions - as captured by the vector $a$ - influence the resulting utilities of the game, but also all types of the players - as captured by the vector $t$ - play a role. Put differently, the actual realization of the type of another player can potentially affect my utility.

Taken the structure of any Bayesian game $\Gamma_B$ as given, the players formulate strategies, i.e. for any possible realization of their private type $t_i \in T_i$ they determine which action or which probability distribution over actions they will choose in the game.

**Definition 2**

*A* ***pure strategy*** $s_i$ ***for player*** $i \in \mathcal{I}$ ***in the Bayesian game*** $\Gamma_B$ *is given by*

$$s_i : T_i \to A_i. \tag{2}$$

*A* ***mixed strategy*** $\sigma_i$ ***for player*** $i \in \mathcal{I}$ ***in the Bayesian game*** $\Gamma_B$ *is given by*

$$\sigma_i : T_i \to \Delta(A_i). \tag{3}$$

**Remark 1** *To denote a system of strategies, one for each player, we will simply write every players' strategy in a vector (as we have proceeded for the sets $A$ and $T$.)*

### 1.1 Bayesian Nash Equilibrium (BNE)

Building on the notion of a (pure/mixed) strategy as defined above, it is possible to think about equilibria of Bayesian games. We will use the notion of a Nash equilibrium from the part on complete information games and generalize this notion to Bayesian games, i.e. situations in which private or incomplete information plays a role.

**Remark 2** *As already mentioned in the previous lectures, we will adhere to the system of Expected-Utility maximization.*

**Assumption 1** *The common prior will be assumed to satisfy*

$$p(t) > 0 \qquad \forall t \in T.$$

*That is, every possible tuple of types among the players has some positive probability.*

**Remark 3** *Assumption 1 is solely made in order to avoid unnecessary complications that are related to conditioning on private information of players.*

**Definition 3 (Ex-Ante Pure-Strategy BNE)**
*A strategy profile $s^* = (s_1^*, ..., s_I^*)$ is called a* ***pure-strategy ex-ante BNE*** *iff for all players $i \in \mathcal{I}$*

$$\sum_{t \in T} p(t) u_i \left( (s_i^*(t_i), s_{-i}^*(t_{-i}), t) \right) \geqslant \sum_{t \in T} p(t) u_i \left( (s_i'(t_i), s_{-i}^*(t_{-i}), t) \right) \qquad \forall s_i'. \tag{4}$$

So, this definition looks at the strategies that constitute a BNE for each player separately. For each player $i$, all opponents' actions are taken as fixed ($s_{-i}^*(t_{-i})$ remains unchanged). Now, one compares the actions that are prescribed by the equilibrium-strategy $s_i^*(t_i)$ to any other possible choice of actions $s_i'(t_i)$. The concept of an ex-ante BNE refers to the fact that the expectation is taken over all players' types $t \in T$. So, one can argue that player $i$ has not yet learned her type and also evaluates all of her possible types according to the common prior. She formulates a contingent plan $s_i^*(t_i)$ that depends on the explicit realization of her type that she will learn later on.

**Definition 4 (Ex-Ante Mixed-Strategy BNE)**
*A strategy profile $\sigma^* = (\sigma_1^*, ..., \sigma_I^*)$ is called a* ***mixed-strategy ex-ante BNE*** *iff for all players $i \in \mathcal{I}$*

$$\sum_{t \in T} p(t) u_i \left( (\sigma_i^*(t_i), \sigma_{-i}^*(t_{-i})), t \right) \geqslant \sum_{t \in T} p(t) u_i \left( (s_i'(t_i), \sigma_{-i}^*(t_{-i})), t \right) \qquad \forall s_i'. \tag{5}$$

The difference between (4) and (5) is simply the fact that pure strategies have been substituted by mixed strategies. Importantly, observe that one can restrict attention to pure strategies if one checks for alternative strategies for player $i$. This relates to the fact - emphasized in Lecture I - that a player needs to be indifferent to all pure strategies that are involved in a mixed-strategy best response. So, if a player is already weakly better off with a particular mixed strategy than with any possible pure strategy, one cannot assemble any mixed strategy from the pure ones that will give the player a higher payoff than the original mixed strategy.

**Definition 5 (Interim Pure-Strategy BNE)**

*A strategy profile $s^* = (s_1^*, ..., s_I^*)$ is called a **pure-strategy interim BNE** iff for all players $i \in \mathcal{I}$ and all types $t_i \in T_i$*

$$\sum_{t_{-i} \in T_{-i}} p(t_{-i}|t_i) u_i \left( (s_i^*(t_i), s_{-i}^*(t_{-i})), t \right) \geqslant \sum_{t_{-i} \in T_{-i}} p(t_{-i}|t_i) u_i \left( (a_i, s_{-i}^*(t_{-i})), t \right) \qquad \forall a_i \in A_i. \quad (6)$$

Compare (6) to (4). In contrast to the notion of an ex-ante BNE, the interim-notion builds on the fact that player $i$ already knows her type, her piece of private information. So, player $i$ can condition on $t_i$ in the common prior and needs only to consider the expectation with respect to the other players' types (over $T_{-i}$). Furthermore, one does not have to specify a different strategy $s_i'$ about what type $t_i$ of player $i$ will do, but one can simply take another action $a_i$ on the right-hand side of the optimality-condition.

**Definition 6 (Interim Mixed-Strategy BNE)**

*A strategy profile $\sigma^* = (\sigma_1^*, ..., \sigma_I^*)$ is called a **mixed-strategy interim BNE** iff for all players $i \in \mathcal{I}$ and all types $t_i \in T_i$*

$$\sum_{t_{-i} \in T_{-i}} p(t_{-i}|t_i) u_i \left( (\sigma_i^*(t_i), \sigma_{-i}^*(t_{-i})), t \right) \geqslant \sum_{t_{-i} \in T_{-i}} p(t_{-i}|t_i) u_i \left( (a_i, \sigma_{-i}^*(t_{-i})), t \right) \qquad \forall a_i \in A_i. \quad (7)$$

Comparing (7) to (6), the same remarks apply as in the previous shift from (4) to (5). Due to the characteristic property of the interim-notion, one does not have to worry about alternative strategies anymore, but the fact that player $i$'s type $t_i$ is fixed in the optimality-condition allows to restrict attention to alternative actions $a_i$.

**Remark 4** *If the type-space $T_i$ for each player $i \in \mathcal{I}$ is continuous, the sums in all the above definitions have to be replaced by integrals over the corresponding sets. The essence of the definitions remains the same.*

As the following proposition shows, the notions of ex-ante- and interim-BNE are in fact equivalent.

**Proposition 1**

1. *Every ex-ante-BNE (in pure or mixed strategies) is also an interim-BNE (in the corresponding strategies).*
2. *Every interim-BNE (in pure or mixed strategies) is also an ex-ante-BNE (in the corresponding strategies).*

The proof of this proposition will be included in these lecture notes, once the first problem has been submitted by all people who take the course for credit.
An important consequence of the proposition is the fact that one does not have to care about the distinction between ex-ante- and interim-BNE. A solution according to one notion will also be a solution according to the other notion. In practise, it is often much more convenient to work with the interim-notion. In this case, one simply fixes one player's type and optimizes pointwise for this fixed type to obtain a player's best-response strategy. Repeating this exercise for all other players of the game yields an interim BNE.

To conclude the definitions relating to BNE, a final notion - ex-post BNE - is introduced.

**Definition 7 (Ex-Post Pure-Strategy BNE)**
*A strategy profile $s^* = (s_1^*, ..., s_I^*)$ is called a **pure-strategy ex-post BNE** iff for all players $i \in \mathcal{I}$ and all types $t = (t_i, t_{-i}) \in T$*

$$u_i\left((s_i^*(t_i), s_{-i}^*(t_{-i}), t)\right) \geqslant u_i\left((a_i, s_{-i}^*(t_{-i}), t)\right) \qquad \forall a_i \in A_i. \tag{8}$$

The notion of an ex-post BNE corresponds to the situation in which all players' types are known by all players. Hence, there is no private information anymore. For player $i$, there is no need to take expectations anymore, but she simply compares the utility from the BNE to any other utility that may arise if she chooses another action $a_i$.

**Definition 8 (Ex-Post Mixed-Strategy BNE)**
*A strategy profile $\sigma^* = (\sigma_1^*, ..., \sigma_I^*)$ is called a **mixed-strategy ex-post BNE** iff for all players $i \in \mathcal{I}$ and all types $t = (t_i, t_{-i}) \in T$*

$$u_i\left((\sigma_i^*(t_i), \sigma_{-i}^*(t_{-i}), t)\right) \geqslant u_i\left((a_i, \sigma_{-i}^*(t_{-i}), t)\right) \qquad \forall a_i \in A_i. \tag{9}$$

This definition generalizes in the same manner as previously outlined from its pure-strategy analogue.

The following proposition shows that the notion of ex-post BNE implies the notion of interim-BNE (and hence the notion of ex-ante BNE due to the equivalence from Proposition 1).

**Proposition 2** *Every ex-post-BNE (in pure or mixed strategies) is also an interim-BNE (in the corresponding strategies).*

**Proof of Proposition 3** *So, consider a mixed-strategy ex-post BNE, i.e. for all players $i \in \mathcal{I}$ and all tuples of types $t \in T$, the inequality (9) holds.*[1] *All of these inequalities will remain valid if both sides are multiplied by the positive number (remember Assumption 1) $p(t_{-i}|p_{t_i})$. Hence:*

$$
\begin{aligned}
& \forall i \in \mathcal{I}, \forall t \in T: \qquad u_i\left((\sigma_i^*(t_i), \sigma_{-i}^*(t_{-i}), t)\right) \geqslant u_i\left((a_i, \sigma_{-i}^*(t_{-i}), t)\right) \quad \forall a_i \in A_i \\
\Leftrightarrow\ & \forall i \in \mathcal{I}, \forall t \in T: \\
& p(t_{-i}|t_i) u_i\left((\sigma_i^*(t_i), \sigma_{-i}^*(t_{-i}), t)\right) \geqslant p(t_{-i}|t_i) u_i\left((a_i, \sigma_{-i}^*(t_{-i}), t)\right) \quad \forall a_i \in A_i \\
\Rightarrow\ & \forall i \in \mathcal{I}, \forall t_i \in T_i: \\
& \sum_{t_{-i} \in T_{-i}} p(t_{-i}|t_i) u_i\left((\sigma_i^*(t_i), \sigma_{-i}^*(t_{-i}), t)\right) \geqslant \sum_{t_{-i} \in T_{-i}} p(t_{-i}|t_i) u_i\left((a_i, \sigma_{-i}^*(t_{-i}), t)\right) \quad \forall a_i \in A_i.
\end{aligned}
$$

*The last implication is simply due to the fact that the inequality-relation is preserved if one takes the sum over $T_{-i}$ if all terms of the sum satisfy the analogous inequality. Furthermore, observe that the last line is exactly the definition of a mixed-strategy interim-BNE, concluding the proof.*

**Remark 5** *The converse implication of Proposition 2 is not true. Not every interim-BNE is also an ex-post BNE. This can be seen from the simple example of a Bayesian game from Lecture II:*

| | | ***Type $t_2^1$ of Player 2*** | | | ***Type $t_2^2$ of Player 2*** | |
|---|---|---|---|---|---|---|
| | | *Left* | *Right* | | *Left* | *Right* |
| ***Player 1*** | *Up* | *1,1* | *0,0* | *Up* | *1,0* | *0,1* |
| | *Down* | *0,1* | *1,0* | *Down* | *0,0* | *1,1* |

*Without explicitly mentioning it, we have found an interim BNE for every possible common prior that may occur in this game as follows:*

$$
s_1^* = \begin{cases} U & \text{if } \alpha > \frac{1}{2}, \\ \lambda U + (1-\lambda) D \text{ for any } \lambda \in [0,1] & \text{if } \alpha = \frac{1}{2}, \\ D & \text{if } \alpha < \frac{1}{2}. \end{cases}
$$

[1] Once established for mixed strategies, the result follows trivially for pure strategies, because any pure strategy is a degenerate mixed strategy, where all probability is concentrated on one element.

$$s_2^*(t_2) = \begin{cases} L & \textit{if } t_2 = t_2^1, \\ R & \textit{if } t_2 = t_2^2. \end{cases}$$

*So, assume that $\alpha = \frac{1}{4}$. In this case, player 1 chooses "Down". But for player 2 being of type $t_2^1$ (implying that player 2 will definitely choose "Left"), player 1 will not play a best response. Thus, we have found a violation of the definition of an ex-post BNE for an equilibrium that clearly satisfies the definition of an interim BNE.*

## 1.2 First-Price Sealed-Bid Auction

Imagine the following situation:
An auction involving 2 bidders is conducted for a specific object, say a painting. Each bidder has a specific valuation for the painting that is her private information. The bidders are asked to submit their bids simultaneously in a sealed envelope. These envelopes will be opened by the auctioneer who awards the painting to the person who has submitted the highest bid. This bidder has to pay exactly her bid for the object.

In the following, this fairly general story will be translated very precisely into the general setting of Bayesian games as outlined above. Hereby, the language of auctions involving bidders and bids will be used in place of the game-theoretic language of players and strategies:

- The set of players $\mathcal{I}$ is given by the set of bidders $(1, 2)$.
- The set of possible actions $A_i$ for each player $i \in \{1, 2\}$ is given by $\mathbb{R}_+$. An action for player $i$ is a specific bid - a non-negative number - that a player/bidder submits.
- The set of possible types $T_i$ for each player $i \in \{1, 2\}$ is given by the set $[0, 1]$. Hereby, the interval $[0, 1]$ is a modeling-choice that makes some of the following computations a little bit easier. A type for player $i$ is a particular number between 0 and 1.
- The common prior $p$ on the set of possible types $T_1 \times T_2$ is given by

  $$t_i \sim iid \quad U[0, 1].$$

  That is, both bidders' valuations are distributed uniformly on $[0, 1]$ and they are independent from each other.
- In order to properly define the bidders' utility functions $u_1, u_2$, one needs a proper assessment of the bidders' bidding strategies. Hence, we will first define those:

For $i \in \{1,2\}$, bidder $i$'s bidding strategy $b_i$ - we have called these elements $s_i$ before - is given as a mapping from types to actions, i.e.

$$b_i : [0,1] \to \mathbb{R}_+.$$

- According to the distinctive principle of a First-Price auction (highest player wins and pays her own bid), the utility-function $u_i((b_i, b_j), (t_i, t_j))$, $i, j \in \{1,2\}, i \neq j$, is defined as follows. Explicitly, player $i$'s utility is described as

$$u_i((b_i, b_j), (t_i, t_j)) = \begin{cases} t_i - b_i & \text{if } b_i > b_j \\ \frac{1}{2}(t_i - b_i) & \text{if } b_i = b_j \\ 0 & \text{if } b_i < b_j \end{cases}. \tag{10}$$

  In the following, the terms "utility" and "payoff" will be employed exchangeably. Observe that we do not put the players actions $a$ as an argument of the utility-function, but the players' bids. This is due to the construction of the agents' strategies as mappings from types to actions, i.e. together with the types the bidding function denotes the action chosen by the player. The first case of the utility-function corresponds to player $i$ winning the auction (by submitting the highest bid ($b_i > b_j$)). In this case, $i$ receives her valuation for the good ("the pleasure of owning the item from now on"), but has to pay the price $b_i$. The third case corresponds to the player $i$ losing the auction. In this case, she does not receive her valuation for the good, but she does not have to pay anything either. The second case refers to the situation where the bidders are tied with their bids (both submit equal bids). In this situation, one has to decide about a tie-breaking rule. The most common and also most intuitive rule is to split the benefit of getting the good (the quantity $t_i - b_i$) between the two players.

The utility-function above has been specified at the ex-post level, i.e. for known types $t_i, t_j$ of both players. But in order to solve the First-Price auction, we need to derive a system of best-response correspondences that either satisfy the definition of a Bayesian Nash equilibrium at the ex-ante- or the interim-level. This is due to the fact that the ex-post-notion is not equivalent to either the ex-ante- or the interim-notion.
Comparing the ex-ante- and the interim-notion, it is more convenient - as discussed above - to work with the interim notion. So, the question arises of how to obtain the interim-payoff function from the ex-post-payoff?
This transition is performed by taking the expectation of the ex-post payoff-function with respect to all the elements of the other player, i.e. $t_j$ and $b_j$. So, player $i$ is interested in maximizing the following expression

$$\max_{b_i} \mathbb{E}_{t_j, b_j^*} \left[ u_i \left( (b_i, b_j), (t_i, t_j) \right) | t_i, b_i \right]. \tag{11}$$

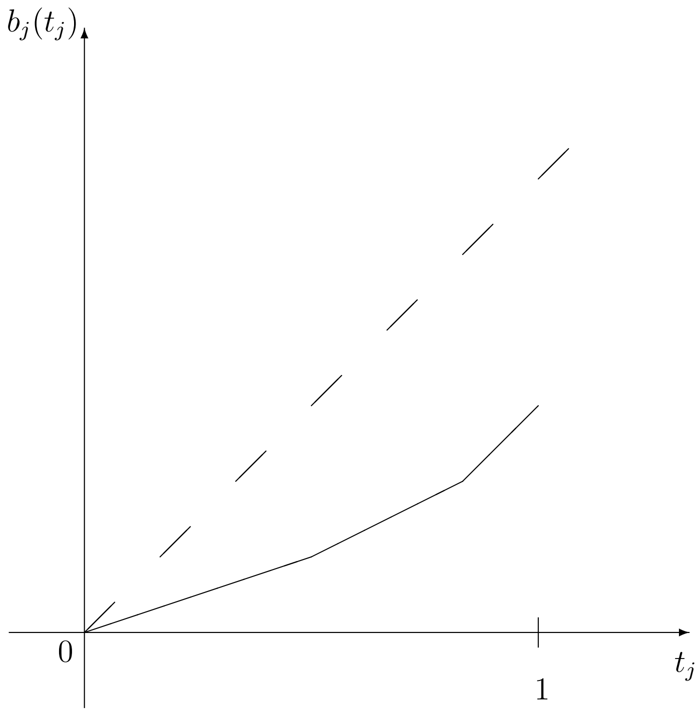
So, player $i$ averages over those elements that she does not have any control over. She does not know the other player's type and, hence, she has no idea about the value $b_j^*$. This will be the value of the bid that player $j$ submits which is a function of player $j$'s private information $t_j$. Since we are interested in a BNE, i.e. in mutual best responses, player $i$ already takes the optimality of player $j$'s bidding behavior into account. According to the interim-notion of a BNE, player $i$ already knows about her type $t_i$. So, she can condition on this element of private information. Furthermore, player $i$ will derive the specific bid $b_i$ as a function of her private information $t_i$. This will also be privately known to her.
Summarizing, player $i$'s objective is to choose the bid $b_i$ optimally as to maximize the expression (11). Now, (11) can be rewritten as

$$
\begin{aligned}
& (t_i - b_i) \cdot \mathbb{P}\left[b_i > b_j^*(t_j)\right] + \frac{1}{2}(t_i - b_i) \cdot \mathbb{P}\left[b_i = b_j^*(t_j)\right] + 0 \cdot \mathbb{P}\left[b_i < b_j(t_j)\right] \qquad (12) \\
= \; & (t_i - b_i) \cdot \mathbb{P}\left[b_i > b_j^*(t_j)\right] + \frac{1}{2}(t_i - b_i) \cdot \mathbb{P}\left[b_i = b_j^*(t_j)\right]. \qquad (13)
\end{aligned}
$$

The first step of the preceding transformation is done according to the following rule on conditional expectations and/or probabilities:

$$\mathbb{E}[X] = \mathbb{E}\left[X|Y > 0\right]\mathbb{P}[\{Y > 0\}] + \mathbb{E}\left[X|Y = 0\right]\mathbb{P}[\{Y = 0\}] + \mathbb{E}\left[X|Y < 0\right]\mathbb{P}[\{Y < 0\}].$$

Explicitly, set $X$ to be our payoff-function and $Y$ to be the difference $b_i - b_j$.
In the next step of the analysis, we aim at "inverting" the opponent's bidding function, i.e. we would like to infer the type from the bid. This will allow us to further simplify the interim-payoff function because player $i$ has a probability-assessment about the other bidder's type (but, importantly, not about the other bidder's bid; hence the "inversion"-procedure). In general, any bidder's bidding function will be given by a function that may look similar to the following graph.




It is an important feature of the concept of a BNE that we are aiming at determining a whole optimal bidding-function. In the following, a couple of properties and their intuition will be described that are considered desirable for the bidding function that represents an optimum in the context of a First-Price auction.

We want the bidding function to satisfy the following characteristics:

- At 0, we want the bidding function to be equal to zero. If the object is not worth anything to a bidder, then the bidder optimally does not give a positive bid.
- The dashed line in the graph represents the 45-degree line. This corresponds to the situation where the bid is exactly equal to the valuation of a bidder. If a bidder exactly places a bid on this line, then she obtains zero utility, which leaves her indifferent to not participating in the auction at all.
- The bidding-function will be located below (or at) the dashed line. This is due to the fact that players will make losses if they announce a bid above their valuation (above the 45-degree line in the graph) and are actually awarded the object. The

distance between the private type $t_i$ and the bid $b_i(t_i)$, i.e. the distance between the dashed line and the bidding curve, is sometimes referred to as **bid shedding**. This distance stands for the amount by which the bid is distorted away from the truth. In the context of the First-Price auction, positive bid shedding corresponds to the possibility of getting strictly positive utility in the case that the bidder actually receives the object.

- We want the bidding function to be strictly increasing. That is, if we compare two types of one bidder, then we want the higher type of the bidder to put in a strictly higher bid.[2]

Incorporating all the above reflections on bidding functions, we will guess a functional form of the opponent's bidding function that appears in the representation of the interim payoff function (12). The guess is that a BNE can be found in linear bidding strategies:

$$\begin{aligned} b_i(t_i) &= \alpha_i + \beta_i t_i, && (14) \\ b_j(t_j) &= \alpha_j + \beta_j t_j. && (15) \end{aligned}$$

This means that we are restricting attention to situations of linear bidding strategies. This does NOT mean that there are no potential other BNE involving non-linear bidding strategies. In the following, our task is to verify that the guess of linear bidding strategies is indeed a BNE and to determine the coefficients $\alpha_i, \alpha_j, \beta_i, \beta_j$.

But how do we use the guess to transform the above expression

$$(t_i - b_i) \cdot \mathbb{P}\left[b_i > b_j^*(t_j)\right] + \frac{1}{2}(t_i - b_i) \cdot \mathbb{P}\left[b_i = b_j^*(t_j)\right].$$

for the interim-expected-payoff of player $i$?

We will plug in the guess for player $j$'s bidding function in this expression. This is actually the only element that we can replace. We are not allowed to tamper with the bidding function $b_i$ because this function will solely be determined from optimality-considerations and we do not want to restrict the set over which we optimize. In the end, we need to check whether the optimality-condition that we will derive for $b_i$ actually is a linear function. So, we obtain

$$\begin{aligned} & (t_i - b_i) \cdot \mathbb{P}\left[b_i > b_j^*(t_j)\right] + \frac{1}{2}(t_i - b_i) \cdot \mathbb{P}\left[b_i = b_j^*(t_j)\right] \\ = \; & (t_i - b_i) \cdot \mathbb{P}\left[b_i > \alpha_j + \beta_j t_j\right] + \frac{1}{2}(t_i - b_i) \cdot \mathbb{P}\left[b_i = \alpha_j + \beta_j t_j\right] \\ = \; & (t_i - b_i) \cdot \mathbb{P}\left[b_i > \alpha_j + \beta_j t_j\right]. \end{aligned}$$

[2] The increasingness-property appears to be very reasonable. In contrast, the strictness is an assumption, but a very convenient one.

Observe that in the last step we have dropped the second summand. This is due to the fact that we restrict attention to strictly increasing bidding functions. In this case, the event that a bidding function (here, $b_i$) equals a specific value (here, $\alpha_j + \beta_j t_j$) is a $\mathbb{P}$-null event, i.e. it has probability zero.
Now, we have finally obtained an expression that only involves the other bidder's type and no longer his bidding function. In the next step, we will transform the above expression in order to be able to apply the distributional assumption (the common prior) for $t_j$:

$$\begin{aligned} & (t_i - b_i) \cdot \mathbb{P}\left[b_i > \alpha_j + \beta_j t_j\right] \\ = \quad & (t_i - b_i) \cdot \mathbb{P}\left[t_j < \frac{b_i - \alpha_j}{\beta_j}\right]. \end{aligned}$$

Given that $t_j \sim U[0,1]$, it follows that

$$\mathbb{P}\left[t_j < \frac{b_i - \alpha_j}{\beta_j}\right] = \frac{b_i - \alpha_j}{\beta_j}.$$

So, finally, player $i$'s objective function is given by

$$\max_{b_i} \left\{ (t_i - b_i) \frac{b_i - \alpha_j}{\beta_j} \right\}. \tag{16}$$

So, player $i$'s objective is to maximize the product of her realized surplus $t_i - b_i$ and the probability of winning $\frac{b_i - \alpha_j}{\beta_j}$. Hereby the player faces the tradeoff to increase her probability of winning (by choosing a higher bid $b_i$) only at the cost of decreasing her surplus. The FOC for (16) is given by

$$\begin{aligned} & -\frac{b_i^* - \alpha_j}{\beta_j} + \frac{t_i - b_i^*}{\beta_j} \overset{!}{=} 0 \\ \Leftrightarrow \quad & 2b_i^* = \alpha_j + t_i \\ \Leftrightarrow \quad & b_i^* = \frac{t_i + \alpha_j}{2}. \end{aligned}$$

In summary, player $i$'s best-response bidding function $b_i^*$ is given by

$$b_i^*(t_i) = \frac{1}{2}\alpha_j + \frac{1}{2} t_i. \tag{17}$$

Observe two important features of this optimal response:

- Making use of the fact that the opponent's bidding strategy is linear, but NOT assuming anything about player $i$'s bidding strategy, one obtains linearity of $b_i^*$ in $t_i$. This is the property that we needed to check in order to make sure that there is a BNE in linear bidding strategies.

- As a consequence of the uniform distribution in the common prior of both players, the best responses do not depend on the slope $\beta_j$ opponent's bidding strategy.

Due to the symmetry of the underlying structure, player $j$'s best response can be obtained from (17) by exchanging $i$ and $j$:

$$b_j^*(t_j) = \frac{1}{2}\alpha_i + \frac{1}{2}t_j. \tag{18}$$

The last step that needs to be taken is the elimination of any element from the opponent's bidding strategy from the best-response-functions $b_i^*(t_i)$ and $b_j^*(t_j)$. This will be done by comparing (17) and (18) to the linearity-conditions (14) and (15) of the bidding strategies. So, comparing the best response

$$b_i^*(t_i) = \frac{1}{2}\alpha_j + \frac{1}{2}t_i$$

to the general functional form of $b_i^*$

$$b_i = \alpha_i + \beta_i t_i,$$

one obtains the following two conditions:

$$\frac{1}{2}\alpha_j = \alpha_i,$$
$$\frac{1}{2} = \beta_i.$$

Analogously, one obtains the following conditions for bidder $j$:

$$\frac{1}{2}\alpha_i = \alpha_j,$$
$$\frac{1}{2} = \beta_j.$$

All these four equations have to be satisfied simultaneously. Therefore, it follows that

$$\beta_i = \beta_j = \frac{1}{2}.$$

as well as

$$\alpha_i = \alpha_j = 0.$$

Summarizing, a BNE in linear bidding strategies is given by

$$b_i^*(t_i) = \frac{1}{2}t_i, \tag{19}$$
$$b_j^*(t_j) = \frac{1}{2}t_j. \tag{20}$$

**Remark 6** *If one does not choose the prior to be independent uniform distributions on* $[0,1]$*, the idea for the solution will be the same as above. The only condition that one needs to assume is strict monotonicity for the bidding function in the type of a bidder to infer types from bids. In the end, the optimality-condition for any bidder will be described by a differential equation.*

## 1.3 Second-Price Sealed-Bid Auction

The setting and the rules of this auction are - with one exception - identical to those of the First-Price auction. The exception is the amount that the winning bidder - who is still the one having submitted the highest bid - has to pay. She no longer has to pay her own bid, but the bid of the other bidder. That is, the utility-function for player $i$, where her opponent is denoted by $j$, is given by

$$u_i((b_i, b_j), (t_i, t_j)) = \begin{cases} t_i - b_j & \text{if } b_i > b_j, \\ \frac{1}{2}(t_i - b_j) & \text{if } b_i = b_j, \\ 0 & \text{if } b_i < b_j. \end{cases} \tag{21}$$

If the general setting is generalized to include more than one bidder, then the second-price auction refers to the situation in which the winning bidder (having submitted the highest bid) pays the second highest bid.
This auction-format is also called **Vickrey-auction**, after William Vickrey - another awardee of the Nobel-prize in Economics - who first brought forward this auction-format in 1962.

The search for Bayesian Nash equilibria is particularly simple in this auction-format, as it will be shown below. This derives from the fact that an equilibrium is not given in Bayesian strategies, i.e. those that take into account the opponent's actions and type-space, but in weakly dominant strategies. In other words, there will be a strategy for each player that is optimal for this particular player irrespective of the opponent's type. This equilibrium will be summarized in the following theorem:

**Theorem 1** *In a second-price auction, the best-response function for any bidder $i \in \{1, 2\}$ is given by*

$$b_i^*(t_i) = t_i,$$

*irrespective of the type of the other bidder.*

In short, in Second-Price auctions, it is optimal for both bidders to simply reveal the truth, i.e. to report their valuation as it really is. In comparison, bidders do not reveal the truth about their valuation by bidding in a First-Price auction (they only report half of their valuation).
Intuitively, in a Second-Price auction - as it is a simultaneous-move game as all the games that we have discussed so far - a player does not derive any direct utility from bidding higher, since this only affects the other player's utility, but not her own. Hence, it makes sense to obtain bidding-functions that are not distorted in any way.

**Remark 7** *The Second-Price auction and the First-Price auction share a common feature: In both formats, the object is allocated to the person with the highest valuation, i.e.*

*the allocation is* ***efficient*** *- it cannot be improved upon from a social-choice perspective. This is due to the fact that those bids will be winning the auction that come from the highest types, which can trivially be seen from the functional form of the optimal bidding functions. Interestingly, the First-Price auction is able to establish efficiency without the players reporting their true types.*

**Proof of Theorem 1** *Wlog, consider player* 1 *and fix his type at* $t_1$. *It will now be shown that player* 1*'s optimal bid is given by* $b^* := b_1(t_1) = t_1$. *In the following, player* 2*'s bid will be denoted by* $b_2$.[3] *Now, different cases will be examined and it will be shown that in every case player* 1 *cannot make herself strictly better off by changing her bid away from* $b^*$*:*

- *Assume that* $b_2 > b^*(= t_1)$*:*
  *In this case, the object is awarded to player* 2. *Player* 1 *gets utility* 0. *Changing her bid to any other* $b < b_2$ *does not change player* 1*'s utility at all. She still does not receive the object, hence her utility is still* 0. *But if player* 1 *chooses either* $b = b_2$ *or* $b > b_2$ *, then she will either receive utility* $\frac{1}{2}(t_1 - b_2)$ *(from the tie-breaking rule) or* $(t_1 - b_2)$ *(from winning the object). In both bases, this expression is negative due to the assumption that* $b_2 > b^* = t_1$, *so player* 1 *is strictly worse off than in the situation where she bids* $b^*$ *(and gets* 0 *utility).*

- *Assume that* $b_2 = b^*(= t_1)$*:*
  *In this case, the object is awarded to player* 1 *as well as player* 2. *Player* 1 *gets utility* $\frac{1}{2}(t_1 - b_2) = \frac{1}{2}(t_1 - b^*) = 0$. *Changing her bid to any other* $b < b_2$ *causes player* 1 *to lose the auction. Hence, the resulting utility for her will be* 0, *which leaves her indifferent. But if player* 1 *chooses* $b > b_2$ *, then she will always win the auction and receive utility* $(t_1 - b_2)$. *Because* $b_2 = t_1$, *this expression equals* 0, *leaving player* 1 *indifferent to the original choice of* $b^*$.

- *Assume that* $b_2 < b^*(= t_1)$*:*
  *In this case, the object is awarded to player* 1. *Player* 1 *gest utility* $t_1 - b_2$. *Changing her bid to any other* $b_2 < b$ *does not change player* 1*'s utility at all. She still receives the object and gets the same utility* $t_1 - b_2$. *But if player* 1 *chooses either* $b = b_2$ *or* $b < b_2$ *, then she will either receive utility* $\frac{1}{2}(t_1 - b_2)$ *(from the tie-breaking rule) or* 0 *(from not winning the auction). In both cases, her utility is strictly less than in the previous situation of bidding* $b^*$.

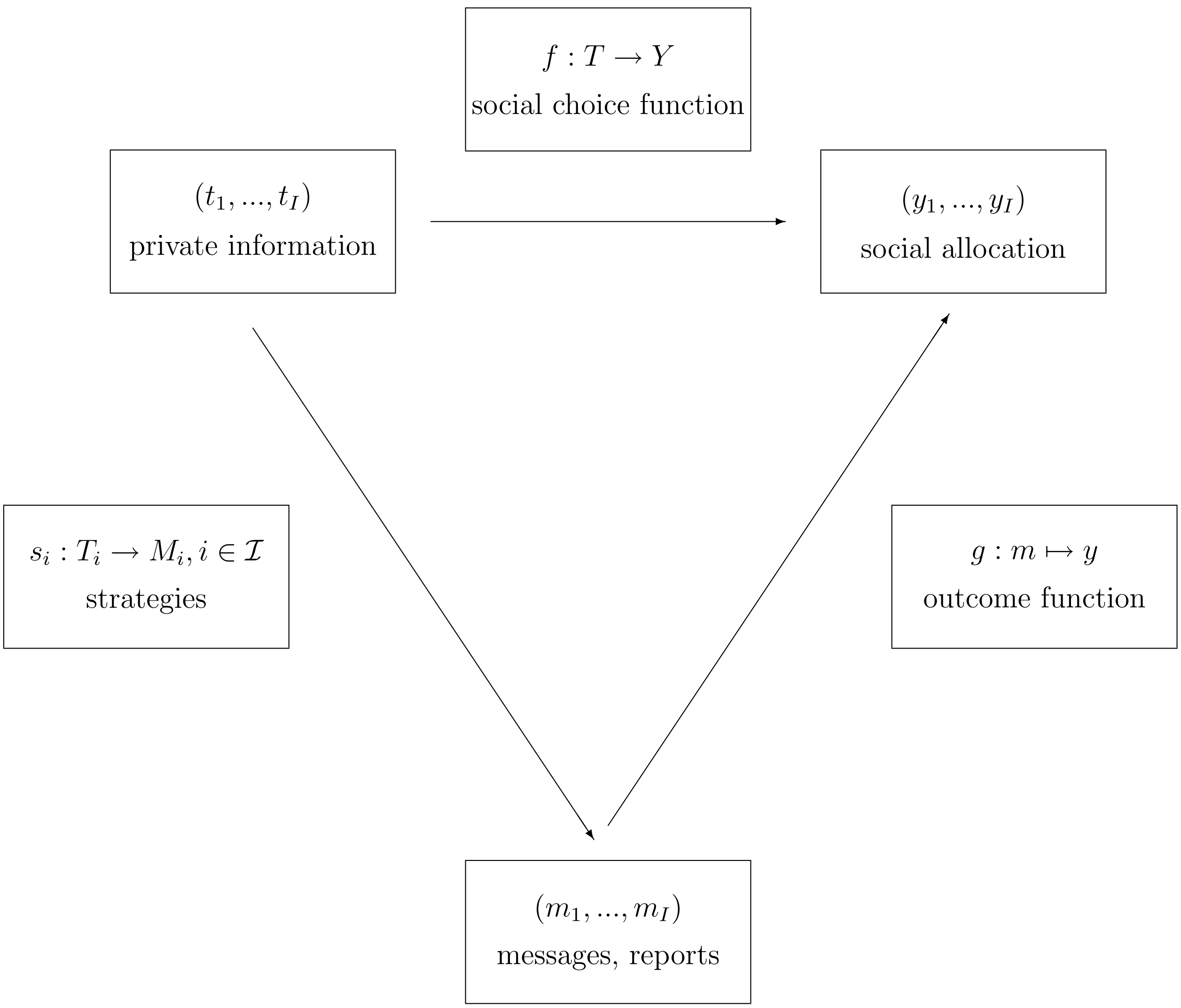
*So, for any possible constellation between* $b^*$ *and* $b_2$, *player* 1 *will not find it profitable to deviate from bidding* $b^* = b_1(t_1) = t_1$, *concluding the proof.*

[3] In slight abuse of notation, we do not consider $b_2$ as a function of player 2's type, but it will simply be an arbitrary number that player 2 may announce. This will actually yield the desired component that player 1's bidding strategy is optimal independent from player 2's type.

# 2 Mechanism Design

For the following discussion, we need to augment the current notation by the tuple of sets $M = M_1 \times ... \times M_I$. Each of the sets $M_i, i \in \mathcal{I}$, refers to messages that player $i$ can send. These elements can also be seen as reports or little pieces of information that player $i$ can unilaterally disclose. Again, $m \in M$ will refer to a specific tuple of messages, one for each player.

Now, consider the following diagram which captures the essence of **Mechanism Design**:




The logic behind this diagram can be described as follows:

- Essentially, the diagram aims at describing ways to map tuples of private information into allocations. So, given the types of the players from the set $\mathcal{I}$, who are predominantly called "agents" in this context, one seeks to decide about an allocation among the players.

- There are two ways to accomplish this mapping from types $(t_1, ..., t_I)$ to allocations $y_1, ..., y_I$:
  - There is a direct way via the social choice function $f$. This is referred to as a **direct mechanism**. There is an instance - known as the **social planner** - who knows the private information of all players. According to the types of the agents, the social planner decides about an allocation.
  - The concept of social planning sounds like a very artificial concept at first. The question naturally arises how the social planner can ever elicit the agents' private information. Will the agents not have an incentive to hide their private information in order to influence the allocation? As an example, consider the case of an auction and imagine that the auctioneer will simply ask all present bidders who wants to obtain the object.[4] In other words, the auctioneer tries to find the bidders' valuations by simply asking them. But every bidder knows that only the highest valuation will actually receive the object. So, it is completely natural for every bidder to shamelessly exaggerate on their valuations. Due to this effect, there is a second way to get from types to allocations:
  - Players first have certain strategies according to which they send messages. This is represented by the $s_i$-mapping, $i \in \mathcal{I}$, which originates from types and maps into sets of messages. Each player decides on his own (and, in his own interest) which message to send.
  - From the tuple of messages that originates from the players' strategies there is a mechanism $g$ that maps messages into social allocations. This mechanism is in fact a game that is to be designed, the reason for the term "Mechanism Design" per se. It takes the reports of the players as given and yields a particular social allocation as its result. It is a complete list of the rules of a game which uses the players messages as its input. The output will be an equilibrium of the game that will reflect a certain allocation among the players.
- Importantly, the game with all its rules and possible consequences is known to the agents at any stage in the diagram, i.e. in particular when they decide on their strategic action to report messages. Of course, it will be the specification of the game is decisive for each players' strategy in the first place. Knowing that their

[4] This question does not involve any money to be paid. It is really just the simple question: What is your valuation? Meaning: Do you want to receive the good?

messages will be plugged into the game strongly determines which kind of message the players want to send.

- So, as already argued in the beginning of this whole introduction, Mechanism Design reverts the logic of Game Theory. Whereas Game Theory takes the structure of the game as given and aims at determining the consequences of the game in the form of equilibria, Mechanism Design aims at creating a game such that a specific outcome function is implemented. Put differently, Mechanism Design has to reverse-engineer a game with specific equilibria in mind.

- The challenge of the task of Mechanism Design is the implementation of the game although nothing is a priori known about the specific types of the agents and the strategic considerations they make (their strategy-mapping).

- As it can be seen from Remark 7, there is certainly more than one way to achieve a certain social allocation:

  - The First-Price and the Second-Price auction in their general setup as described above correspond to certain choices of $g$.
  - Given the structure of the respective game, the players/bidders have to reflect on the optimal strategy (their optimal bid), corresponding to the arrow from types to messages. For the First-Price auction, we have concluded that it is optimal to bid half of one's valuation. Whereas in the Second-Price auction, the strategy-mapping is simply the identity-mapping. That is, for the Second-Price auction, the bidders find it optimal to simply report their true valuation and not distort their valuation in any way.
  - Both mechanisms or games allocate the object to the person with the highest valuation, i.e. the social choice function $f$ for both mechanisms is identical.

An important result is hidden in the context of a Second-Price auction. As argued above, the strategy-mapping of each of the agents is simply the identity-mapping, implying that truth-telling is optimal. Under fairly general circumstances, one can restrict the search for an optimal mechanism, one that implements a specific social choice function, to those mechanisms that take the agents' types (and NOT their messages) as an input, i.e. one can pre-suppose $s_i = id$. This is the so-called **Revelation Principle**. Put differently, one restricts attention to **direct mechanisms**. Hereby, directness refers to mapping types directly into allocations, and hence skipping the message-step.