Four Essays in Auction Theory and Contest Theory

Inaugural-Dissertation
zur Erlangung des Grades eines Doktors
der Wirtschafts- und Gesellschaftswissenschaften
durch die
Rechts- und Staatswissenschaftliche Fakultät
der Rheinischen Friedrich-Wilhelms-Universität
Bonn

vorgelegt von
Dmitriy Knyazev
aus Ryazan

Bonn, 2018

Dekan:	Prof. Dr. Daniel Zimmer
Erstreferent:	Prof. Dr. Benny Moldovanu
Zweitreferent:	Prof. Dr. Tymon Tatur

Tag der mündlichen Prüfung: 06.07.2018

Acknowledgments

First of all, I would like to thank Benny Moldovanu and Tymon Tatur for many helpful discussions and continuous support. Communication with them and their encouragement have helped a lot to improve this thesis. I am also very grateful to Stephan Lauermann and Dirk Bergemann for their comments and support especially at the late stage of writing this thesis. I benefited also from the discussions with Matthias Kräkel, Daniel Krähmer, Sven Rady, Larry Samuelson, Urs Schweizer, and Deszö Szalay.

I thank my fellow students and colleagues for many helpful comments, proofreading, and the nice time we have spent together, particularly Daniel Becker, Nina Bobkova, Holger Herbst, Andreas Kleiner, Benjamin Schickner. I am grateful for the collaboration with Bo Chen and Xiandeng Jiang. Bo Chen is also a great friend who has helped me a lot with nonacademic issues during my time in Bonn. Special thanks for the administrative support go to Silke Kinzig, Pamela Mertens, Britta Altenburg and Olaf Brodersen from the Bonn Graduate School of Economics. I acknowledge financial support from the German Research Foundation (DFG).

Last but not least, I want to thank my parents—Lidia Knyazeva and Andrey Riss— and my lovely wife Ekaterina, who have always supported me during the work on this dissertation.

Contents

Introduction

This thesis consists of four essays that belong to the literature on the theory of auctions and the theory of contests. Chapter 1 and Chapter 2 are devoted to the research on auctions and study the role of symmetry in auction design. In Chapter 1, I show how the designer can favor one of the bidders by choosing auction rules even if the chosen auction has to be symmetric. In Chapter 2, the symmetric revenue-maximizing auction is completely characterized. Chapter 3 and Chapter 4 contribute to the contest theory. In Chapter 3, the role of head starts in search contests is investigated. Chapter 4 finds optimal prize structures of elimination contests for a general form of the designer's objective. Chapter 2 is based on joint work with Bo Chen, and Chapter 3 is based on joint work with Bo Chen and Xiandeng Jiang.

Chapter 1 contributes to the mechanism design literature and considers the problem of favoritism in auctions from the mechanism design perspective. The auction designer has one favorite among bidders and maximizes his utility by choosing an auction format conditional on the favorite's value. Alternatively, one could think that one of the bidders can choose the auction format he would like to participate in depending on his value. To prevent favoritism, several restrictions are imposed on the designer in my model. Many real-life auctions WTO and EU require that procurement auctions conducted under their regulation have to be non-discriminatory that is symmetric. I show that this restriction is not sufficient to prevent discrimination. Namely, even if the designer is restricted to using anonymous and dominant strategy incentive compatible auctions, she can transfer all potential revenue to her favorite and guarantee him the interim utility at least equal to his value for any allocation rule. The equivalence of anonymity with respect to bids and anonymity with respect to true values is also established in this case. This form of favoritism is easily detectable. To prevent this obvious form of favoritism I add the restriction of non-positive trasnfers. Altogether, anonymity, dominant strategy incentive compatibility and non-positive transfers restrictions do not allow the designer to perform intra-auction favoritism, that is there is no particular value favored in equilibrium compared to others. However, intra-auction favoritism is still possible, where the deisgner chooses different auctions for different values of her favorite. Thus, the auction choice still depends on the favorite's value. The designer chooses a second-price auction with pooling, where she commits to not distinguishing values in pooling regions and using lotteries to determine a winner. To fully prevent favoritism, the deterministic auctions restriction is added. Altogether, these restrictions allow implementing only a specific class of second-price auctions with a generalized reserve price. For each bidder, this reserve price depends on opponents' bids. The designer chooses the standard second-price auction from this class and no favoritism is possible.

In his seminal paper, Myerson (1981) finds the optimal auction that maximizes the revenue of the designer. In general, for heterogeneous bidders, his construction is asymmetric. Deb and Pai (2017) show that there exists and a symmetric auction and a Bayesian equilibrium of a new constructed auction that implements the same expected outcome. However, their implementation can lead to auctions with multiple equilibria where there is no reason to prefer one equilibrium over another. In Chapter 2, multiplicity of equilibria is eliminated by considering strategy-proof auctions. The optimal strategy-proof symmetric auction is obtained. It turns out to be a second-price auction with a generalized reserve price defined in Chapter 1. Hence, for each bidder the optimal reserve price depends on what other bidders bid in the auction.

Chapter 3 studies the effects of head starts in innovation contests. The model in this chapter is similar to Taylor (1995). Chapter 3 studies continuous time version of Taylor (1995) and introduces heterogeneity in the form of head start for one of contestants. A two firm winner-takes-all contest in which each firm decides when to stop a privately observed search for innovations (with recall) is analyzed. The firm with a superior innovation at the outset has a head start. The firm with the most successful innovation at a common deadline wins. It is shown that a large head start guarantees a firm victory without incurring cost. However, a medium-sized head start ensures defeat for the firm if the deadline is sufficiently long. In the latter case, the competitor wins the entire rent of the contest. The head start firm may still increase its expected payoff by discarding its initial innovation in order to indicate a commitment to search. The effects of early stage information disclosure and cost advantages are studied, respectively.

Chapter 4 considers multi-stage elimination contests where agents' efforts at different stages generate some output for the organizer. Rosen (1986) studies the similar problem where the organizer wants to induce the same level of effort at all stages and maximize it. In Chapter 4, I find the optimal prize structure for a general class of the organizer's prferences over stages. Depending on these preferences, various prize structures can be optimal. If the output function depends much more on efforts applied at later stages than on those applied at the earlier ones, the optimal prize structure can be non-monotone, that is, at some stages prizes fall and the agents who are more successful may earn less. Necessary and sufficient conditions for the optimality of such structures are provided. I also show that for any increasing prize shape there exists an output function such that this prize shape is optimal. Further, I consider the case of limited liability, where the principal is not allowed to use negative prizes but can choose a contest success function (CSF). There is always an efficient equilibrium under which the principal is able to extract the full surplus from the agents and the corresponding optimal prize structure is always increasing. Moreover, under some plausible assumptions, the optimal CSF is necessary convex, which corresponds to the most frequently used prize schemes in practice.

Chapter 1

Favoritism in Auctions: A Mechanism Design Approach

1.1 Introduction

In this chapter, I approach the problem of favoritism in auctions from the mechanism design perspective. Thus, I study how the designer chooses the auction format to maximize her favorite's utility under different sets of restrictions on the implemented auction rules. In my model, favoritism does not arise due to some hidden actions, unfair manipulation with bids, fictitious bidders, cheating or other unfair actions; rather, favoritism is solely due to the design of the mechanism. There are two main reasons that make the problem interesting: first, there are some restrictions on the auctions formats, which the designer needs to meet while choosing an auction; and second, in addition to the knowledge of her favorite's identity, the designer has information about how much her favorite values the good. This information can be used in the auction design, whereby the designer can choose different auctions for different values of her favorite. The main questions are what auction the designer chooses under different sets of restrictions to make her favorite better depending on his value and what is a good set of restrictions to prevent different forms of favoritism. The first question can be paraphrased in terms of the situation where the bidder chooses an auction. Namely, what auction would be chosen, if one particular bidder could choose an auction he would like to participate in.

There are many real-life auctions where the problem of favoritism is relevant. For example, consider a situation where the principal intends to sell some good using an auction. If she is not sufficiently informed about the market and potential buyers, she could hire an expert to design the auction format to achieve goals such as revenue maximization, efficiency maximization, etc. However, the designer's incentives can differ from those of the principal. As a result, the auction format chosen by the designer can substantially differ from that preferred by the principal. In this chapter, this conflict of interest arises in a situation where the designer has a favorite among potential buyers. One possible reason for this would be a bribe from this particular buyer or any other form of collusion. Subsequently, the designer's objective could be maximization of this particular buyer's utility and the principal could not achieve her goal in the auction outcome. If the principal's objective is revenue maximization and the designer chooses an auction format where the good directly goes to her favorite, this outcome is a disaster for the principal in terms of collected revenue, which is equal zero. Thus, the principal would like to limit the freedom of an auction format choice given to the designer to prevent favoritism. Another situation is a government auction; for example, a government procurement auction, where one of the participating companies may be partially or fully owned by a government. In this case, the government could prefer to choose the auction format that favors this company.

However, one essential requirement for the rules of a procurement is that they guarantee fair competition. Institutions like the European Commission and the WTO set procurement guidelines that should ensure the absence of positive and negative discrimination. In particular, "equal treatment, non-discrimination, mutual recognition, proportionality and transparency" (European Comission, 2014) are required. "Each Party shall seek to avoid introducing or continuing discriminatory measures that distort open procurement" (WTO, 2011). Nevertheless, statistics show that discrimination in procurement is present. According to an estimate (PricewaterhouseCoopers and Ecorys, 2013), the costs of corruption in public procurement in eight EU countries ranged from €1.4 billion to €2.2 billion in 2010. More than half of foreign bribery cases occurring involved obtaining a public procurement contract (OECD, 2014). 10-30% of the investment in a publicly-funded construction project may be lost through mismanagement and corruption (CoST, 2012). The question is why the implemented legal restrictions cannot prevent discrimination and favoritism and how legal restrictions should be changed.

It is obvious that if there are no restrictions imposed on the designer, then the designer can simply allocate the good to her favorite and not charge him anything. This is an example of a situation, which I call perfect favoritism. Namely, perfect favoritism is possible if the designer can guarantee her favorite the ex-post utility higher than his value in any equilibrium of the auction. Hence, some restrictions are needed to prevent this. Probably the most natural attempt to avoid such obvious favoritism is to impose an *anonymity* restriction to eliminate direct discrimination by identity of the bidder. Anonymity means that the allocation and transfer rules should only depend on the submitted bids, rather than the identities of bidders. However, it emerges that anonymity alone is not a particularly useful restriction for several reasons. First, as shown by Deb and Pai (2017), given some asymmetric auction the designer is often able to construct an anonymous auction, which has an equilibrium such that it provides the same expected outcome as the original auction. Thus, if we assume that the designer can choose an equilibrium, then anonymity restriction alone is not a binding constraint at all. One further reason is the first main result of this chapter (Theorem 1.2), showing that if there is some anonymous and *dominant strategy incentive compatible* (DIC) auction that generates revenue R, then there exists another anonymous and DIC auction that has the same allocation rule and where the whole revenue R is transferred to the favorite. For example, the designer can implement the allocation rule of a second-price auction and transfer all collected revenue to his favorite. Hence, the favorite either obtains the good for free or obtains the revenue weakly higher than his value. Therefore, the designer can implement perfect favoritism in an anonymous and dominant strategy incentive compatible auction. This result is stronger than the result of Deb and Pai (2017) in the sense that it does not use the fact that the designer chooses a particular equilibrium. It should be also emphasized that if the auction is DIC, then standard anonymity restriction with respect to bids implies "true" anonymity with respect to values in the corresponding direct auction (Theorem 1.1).

I call intra-auction favoritism a situation where the designer can discriminate bidders within the auction (a bidder with a higher value obtains lower utility than a bidder with a lower value). To avoid intra-auction discrimination via transfers that results to perfect favoritism, I additionally impose the *non-positive transfers* restriction, which does not allow the designer to transfer collected revenue to her favorite. I analyze the case with two bidders and show that under these three restrictions the intra-auction favoritism is not possible and the favorite's preferred auction is a *second-price auction with pooling* (Proposition 1.1). This is the second important result of the chapter. Pooling means that the designer commits to not distinguishing among the bids in certain regions of the values domain and using a lottery to determine a winner. Pooling is always optimal when the favorite's and his opponent's values are sufficiently close. In this case, the winner is determined by a lottery and the payment is lower than in a second-price auction. Additionally, pooling may be used to reduce payments when the favorite wins. I also provide comparative statics results concerning how the choice of mechanism depends on the favorite's value (Proposition 1.3). Only the pooling region at the top changes its size, with all other things being equal. If the favorite's value is too low, then the top pooling region covers the whole set of possible values and the optimal mechanism emerges as a simple lottery.

Although intra-auction favoritism is not possible under anonymity, DIC and non-positive transfers[1], the designer makes the choice of the auction dependent on her favorite's value. Even if the

[1]This is true in the model with two bidders. If there are more than two bidders, then the intra-auction favoritism

chosen auction is fair (non-discriminatory), this is still a form of favoritism. I call this situation inter-auction favoritism. To illustrate the last point, consider a situation in which the designer can only choose among two auction formats: 1) a second-price auction and 2) a symmetric lottery. Both of these formats can be called fair. Indeed, in a second-price auction the bidder with the highest value wins the auction and has to pay the second highest bid. In a lottery, all bidders do not need even to make bids and thus they have the same probabilities of winning the good. However, bidders with different values could still prefer one of these formats to another. For example, if one of n bidders has a low value, he would certainly prefer a lottery rather than a second-price auction, since it gives him a chance to obtain the good for free with probability $1/n$. Meanwhile, a bidder with a high value could prefer a second-price auction rather than a lottery, since his chances of winning the good in the competition are high. Thus, although both described auction formats are fair, they are not equally valued by different bidders.

I show that by imposing one more restriction on the designer, it is possible to prevent any form of favoritism. Thus, I impose a *deterministic auctions* restriction, which does not allow the designer to use randomization to determine a winner if there is a unique highest bid. The third main result characterizes a class of auctions feasible under these four restrictions as *second-price auctions with a generalized reserve price* (Theorem 1.3). A generalized reserve price is different from the standard reserve price in the sense that it is unique for each bidder and depends on all bids of his opponents. However, it is constructed in a symmetric way to preserve anonymity restriction. Independent of the favorite's value, the auction maximizing the utility of the favorite in this class of auctions is a standard second-price auction without any reserve price (Proposition 1.4). Thus, this combination of four restrictions allows preventing any form of favoritism.

I also analyze what kind of favoritism is possible under different subsets of restrictions. I show that the above restrictions form a hierarchy with *non-positive transfers* at the top, *deterministic auctions* at the bottom and *anonymity+DIC* in the middle (Proposition 1.5). In other words, *non-positive transfers* always reduce the scope of favoritism. *Anonymity* helps if and only if *DIC* is imposed and vice versa. *Deterministic auctions* only matter in combination with *anonymity+DIC.*

This chapter is related to papers by Deb and Pai (2017) and Azrieli and Jain (2018). They show that for many mechanisms that are not anonymous, one can find a symmetric auction such that it has a Bayes-Nash equilibrium with the same expected revenue and bidder's utilities. Manelli and Vincent (2010) and Gershkov et al. (2013) show that in the independent private values model, there is equivalence of Bayesian and dominant strategy implementation in expected terms. This

can still be possible.

equivalence does not hold here due to the additional restrictions and in particular anonymity.

Collusion among buyers is studied in Graham and Marshall (1987) and Mailath and Zemsky (1991) for second-price auctions, as well as McAfee and McMillan (1992) for first-price auctions. Robinson (1985), Caillaud and Jehiel (1998), Che and Kim (2006), Marshall and Marx (2007) and Che and Kim (2009) compare possibilities of collusion among buyers or between a buyer and seller in different auction formats. In a setting with non-transferable payments, Condorelli (2012) and Chakravarty and Kaplan (2013) find the social welfare maximizing mechanism with a benevolent designer. They show that the optimal mechanism comprises contest and lottery regions depending on a distribution of values. In this chapter, the favorite's preferred auction under the restriction of non-positive transfers exhibits similar properties.

Extensive literature exists on the informed principal problem (see Myerson, 1983, Maskin and Tirole, 1990, 1992, Severinov, 2008, Mylovanov and Tröger, 2012, 2014 and Yilankaya, 1999). In such models, the design of a mechanism can reflect the information that the designer possesses. Thus, the choice of the mechanism can partially or fully reveal information to the agents. In this chapter, all main results are formulated for dominant strategy incentive compatible auctions. Since each bidder has a dominant strategy, he does not pay attention to the information revealed by the designer.

The remainder of this chapter is structured as follows. In the next section, I present the auction model used in the chapter. Then, I introduce the concept of favoritism. Subsiquently, I introduce the restrictions sequentially and discuss how they help (or ortherwise) to prevent different forms of favoritism. I conclude with a discussion of open issues. All major proofs are delegated to Appendix 1.A.

1.2 Auction Model

The designer has to conduct an auction to sell one indivisible good (object) to a set $N = \{1, ..., n\}$ of potential bidders. The bidders are characterized by independent private values v_i coming from continuously differentiable distributions F_i on $V_i = [\underline{v}_i, \overline{v}_i]$ with a positive density[2]. The designer has a favorite among the bidders and without loss of generality, I assume that it is the first bidder[3]. The designer knows the value of the favorite $v_1 = v^*$ and maximizes his interim utility[4].

[2] $\overline{v_i}$ could be equal to $+\infty$

[3] Otherwise, we can renumerate the bidders such that the favorite obtains a number 1.

[4] The assumption that the designer knows the favorite's value is quite natural. Since the designer wants to maximize the utility of the favorite, their incentives are completely aligned and the favorite would like to disclose the information about his value to the designer regardless.

The auction proceeds in the following steps:

1. The designer announces the rules of the auction.
2. Agents simultaneously decide whether they want to participate in the auction and if yes they make their bids.
3. The winner is determined according to the auction rules defined on step 1.

Each bidder i chooses a bid from a given set of admissible bids, $b_i \in \{\oslash\} \cup B_i$, where $B_i \subset \mathbb{R}_+$ and $b_i = \oslash$ mean that the bidder i does not participate in the bidding. By $\mathbf{B} = \times_{i=1}^{n} (\{\oslash\} \cup B_i)$ we denote the product set of admissible bid sets. $M \subset N$ is a set of bidders who participate in the bidding, namely $\forall i \in M : b_i \neq \oslash$. The number of participating bidders is $m = |M|$. I denote a vector of values $\mathbf{v} = (v_1, ..., v_n) \in \times_{i=1}^{n} V_i$ and vector of bids $\mathbf{b} = (b_1, ..., b_n) \in \mathbf{B}$. $N_{-i} = N \backslash \{i\}$, $\mathbf{v}_{-i}$, $\mathbf{b}_{-i}$ are used for the set of bidders without a bidder i. When the bids are submitted, an outcome of the auction has to be determined. Denote by a_j an allocation of the object where an agent j obtains the object. By a_0, I denote the allocation when the object remains unassigned. The set of possible allocations is $A = \{a_j\}_{j=0}^{n}$. An allocation is chosen according to an allocation rule $\mathbf{y} : \mathbf{B} \to [0,1]^n$, $\mathbf{y}(\mathbf{b}) = (y_1(\mathbf{b}), ..., y_n(\mathbf{b}))$, where $y_i(\mathbf{b}) := \Pr(a_i|\mathbf{b})$[5]. The allocation rule determines how often each allocation is chosen. Transfer rule $\mathbf{p} : A \times \mathbf{B} \to \mathbb{R}^n$, $\mathbf{p}(a, \mathbf{b}) = (p_1(a, \mathbf{b}), ..., p_n(a, \mathbf{b}))$, where $p_i(a, \mathbf{b})$ specifies how much agent i receives in the allocation a, given that a vector of bids $\mathbf{b}$ is submitted. Transfers $\mathbf{t} : \mathbf{B} \to \mathbb{R}^n$, $\mathbf{t}(\mathbf{b}) = (t_1(\mathbf{b}), ..., t_n(\mathbf{b}))$, where $t_i(\mathbf{b}) := \sum_{a \in A} p_i(a, \mathbf{b}) \Pr(a|\mathbf{b}) = \sum_{j=0}^{n} p_i(a_j, \mathbf{b}) \Pr(a_j|\mathbf{b}) = \sum_{j=0}^{n} p_i(a_j, \mathbf{b}) y_j(\mathbf{b})$ can be computed after the bids have been submitted, but before an allocation has been chosen.

Example 1.1. *The auction format is a simple lottery, where the winner and only the winner pays a fixed price γ independent of bids. Subsequently, the allocation rule is $\mathbf{y}(\mathbf{b}) = (1/n, ..., 1/n)$, bidder i pays $-\gamma$ if he obtains the object and 0 otherwise, namely $p_i(a_j, \mathbf{b}) = -\gamma$ if $i = j$ and $p_i(a_j, \mathbf{b}) = 0$ if $i \neq j$, and the transfers are $\mathbf{t}(\mathbf{b}) = (-\gamma/n, ..., -\gamma/n)$.*

The utility of an agent i who participates in the auction is

$$U_i(v_i|a) = v_i I\{a = a_i\} + p_i,$$

where $I : A \to \{0, 1\}$ is an indicator function equal to 1 if $a = a_i$ and 0 otherwise. The ex-post

[5] By $\Pr(a_i|\mathbf{b})$, I mean the probability that an allocation $a_i \in A$ is chosen conditional on a vector $\mathbf{b} \in \mathbf{B}$ is submitted.

utility of a bidder i given vector of bids $\mathbf{b}$ is as follows:

$$U_i(v_i|\mathbf{b}) = \sum\nolimits_{a\in A} U_i(v_i, a)\Pr(a|\mathbf{b}) = v_i y_i(\mathbf{b}) + t_i(\mathbf{b}).$$

For any vector of bidding strategies $\boldsymbol{\beta}(\mathbf{v}) = (\beta_1(v_1), ..., \beta_n(v_n))$ where $\beta_i : V_i \to \{\oslash\} \cup B_i$, we can define the interim utility of a bidder i as an expectation of his ex-post utility taken with respect to a vector of other bidders' values $\mathbf{v}_{-i}$, given that $\beta(\mathbf{v})$ is played. Thus,

$$U_i(v_i|\boldsymbol{\beta}) = v_i \mathbb{E}_{\mathbf{v}_{-i}} y_i(\boldsymbol{\beta}(\mathbf{v})) + \mathbb{E}_{\mathbf{v}_{-i}} t_i(\boldsymbol{\beta}(\mathbf{v})).$$

When it is clear which bidding strategy we consider, I simply use $U_i(v_i)$ rather than $U_i(v_i|\boldsymbol{\beta})$. Each bidder i participates in the auction, making a bid $b_i \neq \oslash$ if and only if the individual rationality constraint holds:

$$U_i(v_i|\boldsymbol{\beta}) \geq 0. \tag{1.1}$$

Definition 1.1 (feasible auction)**.**

A feasible auction $FA = (\mathbf{B}, \mathbf{y}, \mathbf{p})$ *is a collection of bid sets* $\mathbf{B}$*, an allocation rule* $\mathbf{y}$ *and a transfer rule* $\mathbf{p}$*, such that*

$$\begin{aligned} \forall i, \mathbf{b} \quad & 0 \leq y_i(\mathbf{b}) \leq 1, \\ \forall \mathbf{b} \quad & \sum\nolimits_i y_i(\mathbf{b}) \leq 1, \\ \forall i, a, \mathbf{b}_{-i} \quad & y_i(\mathbf{b}) = p_i(a, \mathbf{b}) = 0 \;\; if \;\; b_i = \oslash. \end{aligned}$$

Any feasible auction should completely ignore bidders who do not participate in the bidding. These bidders never receive the good or transfers. The solution concept is Bayes-Nash equilibrium (BNE). The profile of bidding strategies $\psi = \{\beta_i^*(v_i)\}_{i=1}^n$ constitutes a Bayes-Nash equilibrium of an auction if the interim utility from playing the equilibrium strategy is greater than any other strategy, i.e. for any v_i and for any $\beta_i(v_i)$:

$$\begin{aligned} & v_i \mathbb{E}_{\mathbf{v}_{-i}} y_i(\boldsymbol{\beta}^*(\mathbf{v})) + \mathbb{E}_{\mathbf{v}_{-i}} t_i(\boldsymbol{\beta}^*(\mathbf{v})) \geq \\ \geq \;\; & v_i \mathbb{E}_{\mathbf{v}_{-i}} y_i(\beta_1^*(v_1), ..., \beta_i(v_i), ..., \beta_n^*(v_n)) + \mathbb{E}_{\mathbf{v}_{-i}} t_i(\beta_1^*(v_1), ..., \beta_i(v_i), ..., \beta_n^*(v_n)). \end{aligned} \tag{1.2}$$

Definition 1.2 (no deficit)**.**

An equilibrium ψ *of a feasible auction* FA *is feasible if it does not run ex-post deficit:*

$$\psi : \sum\nolimits_{i=1}^n t_i(\boldsymbol{\beta}^*(\mathbf{v})) \leq 0. \tag{1.3}$$

In any feasible equilibrium, the sum of transfers to bidders is non-positive. However, without any further restrictions, a transfer to some particular bidder could be positive. It is important to emphasize here that non-positive transfers are only a restriction only on equilibrium outcome. Thus, it may not hold for any vector $\mathbf{v}$, but should hold for those vectors that appear in equilibrium ψ. Since the designer knows the value of the favorite, the auction can be such that it runs the deficit if the favorite makes a bid b_1 different from $\beta_1^*(v^*)$. However, this never happens in equilibrium ψ and hence it is sufficient that $\sum_i t_i(\boldsymbol{\beta}^*(\mathbf{v})) \leq 0$ only for $v_1 = v^*$ and for any $\mathbf{v}_{-1}$. This concludes the description of a model and now we continue with a concept of favoritism.

1.3 Favoritism

Denote by $\Psi(A)$ the set of all undominated feasible BNE of some auction A. I will now use notation $U_i(v_i, \psi)$ to denote the interim utility of a bidder i in a particular equilibrium $\psi \in \Psi(A)$.

Definition 1.3 (favorite's preferred equilibrium).

A favorite's preferred equilibrium (FPE) $\psi^*(A) : A \to \Psi(A)$ is the equilibrium that generates the highest interim utility for the favorite given his value v^* among all feasible undominated equilibria, namely for any $\psi \in \Psi(A)$:

$$U_1(v^*, \psi^*(A)) \geq U_1(v^*, \psi)$$

Definition 1.4 (favorite's preferred auction).

A favorite's preferred auction (FPA) is a feasible auction that maximizes the favorite's interim utility in FPE, namely,

$$FPA = \arg\max_{FA} U_1(v^*, \psi^*(FA)) \tag{1.4}$$

Since the choice of an auction may generally depend on the actual value of the favorite, it means that the favorite and all other bidders can be in different information sets when the auction starts. Hence, when an auction format is announced, other bidders can make an inference about a favorite's value. By manipulation with the auction format, the designer can exclude the participation of some potential bidders.

Taking into account the possibility of favoritism, some restrictions can be imposed on auctions proposed by the designer. I denote by $C = \{c_i\}_{i=1}^K$ the set of restrictions on $(\mathbf{y}, \mathbf{p}, \mathbf{t})$. Thus, the designer is not completely free in the choice of an auction. I introduce the following two definitions to take this into account.

Definition 1.5 (auction feasible under restrictions).

A feasible auction under set of restrictions C (later $FA(C)$) is a feasible auction $FA = (\mathbf{B}, \mathbf{y}, \mathbf{p})$ such that $(\mathbf{B}, \mathbf{y}, \mathbf{p})$ satisfy C.

Definition 1.6 (favorite's preferred auction under restrictions).

A favorite's preferred auction under set of restrictions C (later $FPA(C)$) is a feasible under C auction, which maximizes favorite's interim utility in FPE, namely,

$$FPA(C) = \arg \max_{FA(C)} U_1(v^*, \psi^*(FA(C)))$$

The concept of favoritism is formulated in the next definitions.

Definition 1.7 (intra-auction favoritism).

The auction allows intra-auction favoritism if there exist an equilibrium $\psi \in \Psi(A)$, two bidders i, j and a vector of values $\mathbf{v}$, such that $v_i \geq v_j$ and $U_i(v_i|\boldsymbol{\beta}^*(\mathbf{v})) < U_j(v_j|\boldsymbol{\beta}^*(\mathbf{v}))$.

This definition means that intra-auction favoritism exists if there exist an equilibrium and two bidders such that one of them has a greater value and at the same time a lower level of ex-post utility in this equilibrium compared to the other. It also implies that all bidders with the same values should obtain the same utilities. If intra-auction favoritism is possible, it means that the designer can discriminate bidders by their identities within the same auction.

Definition 1.8 (inter-auction favoritism).

The auction allows inter-auction favoritism if the favorite's preferred auction depends on the favorite's value v^*.

In other words, for two different values of the favorite the choice of the auction format will differ. Thus, even if intra-auction favoritism is not possible, the designer could favor one bidder by a particular choice of a mechanism.

Definition 1.9 (perfect favoritism).

Perfect favoritism is possible under set of restrictions C if there exists a feasible auction $FA(C)$ such that in any equilibrium in undominated strategies $\psi \in \Psi(FA(C))$ and for any $v^* \in V_1$ the following holds

$$U_1(v^*, \psi) \geq v^*$$

Thus, perfect favoritism is possible when the designer is always able to guarantee her favorite the interim utility greater than or equal to his value of the good. One trivial example is an allocation of the good to the favorite independent of bids. Another example is when rather than allocating a good she sends him a transfer $p_i > v_i$. Of course, these examples may not be feasible under an appropriate set of constraints. Next, we discuss what the designer can do under different sets of restrictions C.

1.4 Unrestricted Favoritism

First, suppose that $C = \oslash$. Thus, no restrictions are imposed on the designer's choice of an auction. In this case, the designer can simply give the object to her favorite for free. However, it is not the favorite's preferred auction and it is possible to construct an even better mechanism. The next proposition provides a characterization of FPA.

Claim 1.1 (favorite's preferred auction)**.**

If no restrictions are imposed on the designer, the favorite's preferred auction has a favorite's preferred equilibrium in dominant strategies ψ^* and treats the favorite and other bidders differently. The favorite obtains the object if nobody else obtains it and receives all collected revenue. All other bids are treated as in the optimal auction proposed by Myerson (1981), where a seller's reservation value is equal to v^*.

Proof. Since the designer is always able to transfer all collected revenue to his favorite, it is always possible to have an equality in (1.3) and hence $t_1(\mathbf{b}) = -\sum_{i\neq 1} t_i(\mathbf{b})$. Subsequently, problem (1.4) can be rewritten as

$$v^*\mathbb{E}_{\mathbf{v}_{-1}} y_1(\boldsymbol{\beta}^*(\mathbf{v})) - \mathbb{E}_{\mathbf{v}_{-1}} \sum\nolimits_{i\neq 1} t_i(\boldsymbol{\beta}^*(\mathbf{v})) \to \max_{FA}$$

This problem is essentially similar to a problem of profit maximization when the seller has a reservation value equal to v^* and all bidders aside from the favorite participate in the bidding. The result follows directly. □

The Myerson's optimal auction allocates the good to a bidder with the highest "ironed virtual value" $\phi_i(v_i)$[6], provided that this value is greater than the reservation value r of a seller. The winner should pay the amount that is equal to the lowest $\widehat{v}$, such that it lets him win, i.e. $\widehat{v}$ is the solution to $\phi_i(\widehat{v}) = \max(\{\phi_j(v_j)\}_{j\neq i}, r\}$ In the model of favoritism, we can think about a favorite's value as a reserve value of a designer and thus $r = v^*$. Hence, in FPA the favorite obtains the object if all other bidders have virtual values smaller than v^*, i.e. $\forall j \neq 1$, $\phi_j(v_j) < v^*$. Suppose that the bidder k wins in FPA. The smallest value $\widehat{v_k}$ that lets the him win the auction is always greater than or equal to v^*. Indeed, otherwise, since $\phi_k(v_k) < v_k$, we would have $\phi_k(\widehat{v_k}) < \widehat{v_k} < v^*$, which contradicts the fact that $\widehat{v}$ lets win the auction. Thus, when the favorite does not win the

[6] $\varphi_i(v_i) = v_i - \frac{1-F_i(v_i)}{f_i(v_i)}$ if it is increasing, otherwise $\varphi_i(v_i)$ is equal to a special "ironed" transformation of $v_i - \frac{1-F_i(v_i)}{f_i(v_i)}$, such that it makes it monotone.

FPA, he always receives a monetary transfer $\widehat{v} > v^*$ and hence even his ex-post utility is greater than his value.

Observation 1.1. *If the designer is unrestricted, then the perfect favoritism is possible.*

Since all of the collected money goes to the favorite, the actual revenue is always zero. In order to prevent the perfect favoritism and the zero revenue, restrictions on feasible auction should be imposed. To understand what would be the reasonable set of restrictions, I discuss what the designer uses to implement perfect favoritism if she is unrestricted. First, we observe from Claim 1.1 that the designer always wants to differentiate her favorite and all other bidders. This possibility should be excluded and the natural way to achieve this is to impose a restriction that requires the designer to treat all bidders equally, namely anonymity.

1.5 Anonymity

Let $\pi : \{1, ..., n\} \to \{1, ..., n\}$ be a permutation. Denote Θ as the set of all permutations of n elements. Later, for simplicity, I will also use expressions like $\pi(i) = j$, where I mean that the element in i-th position moves to j-th position when permutation π is applied. Denote by $\mathbf{b}_\pi = (b_{\pi(1)}, ..., b_{\pi(i)}, ..., b_{\pi(n)})$

Definition 1.10 (anonymity)**.**

A feasible auction FA is anonymous (feasible under c_A) if the names of the bidders do not matter, namely if any permutation of bids among bidders alters $(\mathbf{y}, \mathbf{t})$ symmetrically. Precisely, for any bidders $i, j \in N$, for any allocation $a_k \in A$ for any permutation $\pi \in \Theta$ and for any vector of bids $\mathbf{b} \in \mathbf{B}$:

$$\begin{aligned} B_i &= B_j, \\ y_i(b_{\pi(1)}, ..., b_{\pi(i)}, ..., b_{\pi(n)}) &= y_{\pi(i)}(b_1, ..., b_i, ..., b_n), \\ p_i(a_k, b_{\pi(1)}, ..., b_{\pi(i)}, ..., b_{\pi(n)}) &= p_{\pi(i)}(a_{\pi(k)}, b_1, ..., b_i, ..., b_n). \end{aligned}$$

This definition means that if after a permutation π a bidder i makes a bid that an agent $\pi(i)$ has made before the permutation, he should have the same probability of winning the auction and the same transfer at any allocation a_k as the agent $\pi(i)$ before the permutation at the allocation

$a_{\pi(k)}$. Note that this also implies that

$$
\begin{aligned}
& t_i(b_{\pi(1)}, ..., b_{\pi(i)}, ..., b_{(n)}) = \sum_{j=0}^{n} p_i(a_j, \mathbf{b}_\pi) y_j(\mathbf{b}_\pi) = \\
= & \sum_{j=0}^{n} p_{\pi(i)}(a_{\pi(j)}, \mathbf{b}) y_{\pi(j)}(\mathbf{b}) = \sum_{j=0}^{n} p_{\pi(i)}(a_j, \mathbf{b}) y_j(\mathbf{b}) = \\
= & t_{\pi(i)}(b_1, ..., b_i, ..., b_n).
\end{aligned}
$$

Hence, expected transfers are also symmetric with respect to a permutation. To understand how it works, consider an example with three bidders, and a vector of bids (b_1, b_2, b_3). Consider now the permuted vector of bids (b_2, b_3, b_1). By the anonymity restriction, the probability that bidder 1 wins bidding b_2, when bidder 2 bids b_3 and bidder 3 bids b_1, should be equal to the probability that bidder 2 wins bidding b_2 and his opponents bidding b_1 and b_3. Consider also the allocation a_3, i.e. the third bidder wins the good bidding b_1. Accordingly, the transfer to bidder 1 in this allocation given that he bids b_2, and bidder 2 bids b_3 must be equal to the transfer to bidder 2 when he bids b_2 in the allocation, where bidder 1 wins and bids b_1, with bidder 3 making a bid b_3.

This restriction holds strong importance. Without anonymity, the designer can simply give the object to her favorite for free. By contrast, when anonymity is imposed, the designer is no longer able to discriminate bidders directly by making different rules for different bidders. However, as shown in Deb and Pai (2017), the anonymity restriction often does not truly restrict the designer in the ability to implement the auction that she wants. Suppose that the designer wants to implement the nonanonymous allocation rule, such that it allocates the object to a bidder with the highest index I_i, where $I_i(v_i)$ is some increasing function of a bidder's value.

They show that there exists an anonymous auction and an equilibrium of this auction such that it implements the same allocation and the same expected payments as the original auction. One of their main results is also that the designer is able to implement in a symmetric way, particularly the optimal auction, which is not anonymous if the distributions of agents' values differ. Indeed, since the optimal auction allocates the object to a bidder with the highest ironed virtual value, we can define $I_i(v_i) = \phi_i(v_i)$. In terms of allocation rule, FPA only differs from the optimal auction in the index function for the first agent, namely $I_1(v_1) = v_1$. Hence, this implementability result also holds in our model and the designer can implement FPA as an anonymous mechanism.

In Appendix 1.B, we show ex-post implementability for the case of symmetric bidders. This theorem states that for symmetric bidders it is possible to construct an auction that has an equilibrium such that the outcome of this equilibrium is *ex-post identical* in terms of allocation and payments to the equilibrium of FPA. This equilibrium has the property that among all bidders with values smaller or equal than $\widehat{v}$ only the favorite participates and bids his true value v^*.If there is no other bidder with a value greater than $\widehat{v}$, the favorite wins the object and pays zero; otherwise, the highest bid wins the auction and this bidder pays a maximum of the second highest bid and $\widehat{v}$. This payment goes to the favorite. The intuition behind this result is that when such an auction is announced, all standard bidders with values lower than $\widehat{v}$ know that if they participate they cannot end up with a profit in the case when there is somebody else with a value below $\widehat{v}$, who participates. In this case, the revenue for the designer is also equal to zero.

This construction above — as well as one by Deb and Pai (2017) — has the weakness that there could be many equilibria of the symmetric auction and we emphasize one particular equilibrium where the favorite is preferred. In fact, it is assumed that the designer can choose among different equilibria. There are also $n-1$ similar equilibria where one of bidders participates and the others, including the favorite, do not participate. Since our notion of perfect favoritism requires that the favorite obtains sufficiently high utility in any equilibrium, the construction above does not allow preventing the perfect favoritism. Hence, at this point one could think that perfect favoritism is not possible if anonymity restriction is imposed. However, it is not true and, as we show below, the perfect favoritism is still possible; namely, there exists an auction such that it has only one equilibrium in undominated (in our case, it would even be dominant!) strategies that provides the favorite with the level of utility higher than his value.

Thus, anonymity restriction itself is not sufficient for the absence of favoritism. It is clear that the opportunity to exclude the participation of other bidders has to be disabled. Thus, we consider dominant strategy incentive compatibility restriction.

1.6 Dominant Strategy Incentive Compatibility

Definition 1.11 (dominant strategy incentive compatibility).

A feasible auction FA is dominant strategies incentive compatible (DIC, feasible under c_{DIC}) if for any bidder there exists a strategy $\beta_i^*(v_i)$ that provides higher utility than any other strategy independent of how the other bidders play, namely for all $\{\beta_j(v_j)\}, j = 1, ...n$:

$$v_i \mathbb{E}_{\mathbf{v}_{-i}} y_i(\beta_1(v_1), ..., \beta_i^*(v_i), ..., \beta_n(v_n)) + \mathbb{E}_{\mathbf{v}_{-i}} t_i(\beta_1(v_1), ..., \beta_i^*(v_i), ..., \beta_n(v_n)) \geq$$
$$\geq \quad v_i \mathbb{E}_{\mathbf{v}_{-i}} y_i(\beta_1(v_1), ..., \beta_i(v_i), ..., \beta_n(v_n)) + \mathbb{E}_{\mathbf{v}_{-i}} t_i(\beta_1(v_1), ..., \beta_i(v_i), ..., \beta_n(v_n))$$

Since the inequality should hold for all $\beta_j(v_j)$, it should also hold for any constant strategies,

$\forall j \neq i, \forall v_j : \beta_j(v_j) = b_j$. In turn, if the equality holds for any bids b_j plugged instead of $\beta_j(v_j)$, this means that it would hold in expectation. Thus, the $\beta_i^*(v_i)$ is a dominant strategy for a bidder i if and only if for any $b_j \in \{\oslash\} \cup B_i$ and for any $\beta_i(v_i)$

$$
\begin{aligned}
& v_i y_i(b_1, ..., \beta_i^*(v_i), ..., b_n) + t_i(b_1, ..., \beta_i^*(v_i), ..., b_n) \geq \\
\geq \; & v_i y_i(b_1, ..., \beta_i(v_i), ..., b_n) + t_i(b_1, ..., \beta_i(v_i), ..., b_n)
\end{aligned}
$$

Although dominant strategy implementation is robust in the sense that the behavior of each player does not depend on what others do, it can have more than one dominant strategy[7]. However, in the auction setting with bidders who have private values and linear utilities, the dominant strategy is unique if it exists. The next result shows this:

Lemma 1.1 (uniqueness of dominant strategy).

For any FA, there could be at most one dominant strategy in the sense that if there are other dominant strategies they also provide the same allocation and transfers, namely for any two dominant strategies of each player $\beta_i^*(v)$, $\beta_i^{**}(v)$ and for any bids of other bidders $\mathbf{b}_{-i}$ the following holds:

$$
\begin{aligned}
y_i(b_1, ..., \beta_i^*, ..., b_n) &= y_i(b_1, ..., \beta_i^{**}, ..., b_n), \\
t_i(b_1, ..., \beta_i^*, ...b_n) &= t_i(b_1, ..., \beta_i^{**}, ...b_n),
\end{aligned}
$$

Proof. See Appendix 1.A. □

The next simple lemma is also crucial for our further results and it only holds for anonymous auctions.

Lemma 1.2 (universality of dominant strategy).

If $b^*(v)$, $v \in V_i$, is a dominant strategy for a bidder i in some anonymous auction $FA(C)$, it is also a dominant strategy for any other bidder j with any value $v_j \in V_i \cap V_j$.

Proof. See Appendix 1.A. □

Later on, when we talk about "equilibrium" we mean the unique equilibrium in dominant strategies where all bidders use the same strategy. It is also convenient to consider direct auctions. An auction is called direct if for any bidder $i \in N$ the allowed bidding set is equal to a union of sets of possible values, namely $B_i = \underset{j \in N}{\cup} V_j$ for any i. Subsequently, describing direct auctions,

[7] Here, I mean a weakly dominant strategy. If there exists a strictly dominant strategy, it is unique.

instead of $(\mathbf{B}, \mathbf{y}, \mathbf{p})$, I use simplified notation $(\mathbf{y}, \mathbf{t})$, keeping in mind that $\mathbf{B} = \underset{i \in N}{\times} (\underset{j \in N}{\cup} V_j \}$ and $t_i(\mathbf{v}) = \sum_{j=0}^{n} p_i(a_j, \mathbf{b}) y_j(\mathbf{b})$. The classical revelation principle claims that without loss of generality it is possible to restrict attention to direct mechanisms in which truth-telling is a Bayes-Nash equilibrium. However, under anonymity restriction, it is not possible to directly apply the revelation principle and preserve this restriction for a direct auction. Note that anonymity restriction imposes constraints on allocation and transfers based on bids $\mathbf{b}$, not the values. If anonymity is the only restriction, namely $C = \{c_A\}$, then the anonymity with respect to bids does not imply the anonymity with respect to values of the direct mechanism. To illustrate this idea, consider the auction from Proposition 1. This auction is anonymous with respect to bids, although the bidding behavior is different for different bidders. Thus, bidders with the same values can make different bids in the auction. Hence, the class of anonymous direct auctions is smaller than the class of all anonymous auctions. Hence, while considering anonymous auctions, we cannot simply restrict our attention to direct anonymous auctions. However, under additional DIC restriction, I can show the equivalence between anonymity with respect to bids of the original auction and anonymity with respect to values of the corresponding direct auction.

Theorem 1.1 (anonymity with respect to valuations)**.**

Anonymity with respect to bids of any DIC auction implies anonymity with respect to values of the corresponding direct auction.

Proof. Suppose that each agent has a dominant strategy $\beta_i^*(v)$ in the original anonymous auction. In the corresponding direct auction, then:

$$\begin{aligned} & y_i(v_{\pi(1)}, ..., v_{\pi(n)}) = y_i(\beta_1^*(v_{\pi(1)}), ..., \beta_n^*(v_{\pi(n)})) = \\ = \ & y_i(\beta_{\pi(1)}^*(v_{\pi(1)}), ..., \beta_{\pi(n)}^*(v_{\pi(n)})) = \\ = \ & y_{\pi(i)}(\beta_1^*(v_1), ..., b_n^*(v_n)) = y_{\pi(i)}(v_1, ..., v_n), \end{aligned}$$

where the first equality follows from Lemma 1.1, the second equality follows from Lemma 1.2, the third equality is due to anonymity and the final one is again due to Lemma 1.1. The similar logic holds for transfers. □

In other words, for any feasible auction that is DIC and anonymous, the corresponding direct auction is also anonymous. Thus, we do not exclude any feasible auctions when instead of using original anonymous DIC auctions we consider corresponding anonymous direct auctions. Maskin and Laffont (1979) characterizate all DIC direct mechanisms and show that the necessary and sufficient conditions for bidders reporting their true values are as follows:

$$1)\ y_i(\mathbf{v}) \text{ is nondecreasing in } v_i \text{ for all } \mathbf{v}_{-i}, \tag{1.5}$$

$$2)\ v_i y_i(\mathbf{v}) + t_i(\mathbf{v}) = h_i(\underline{v}_i, \mathbf{v}_{-i}) + \int_{\underline{v}_i}^{v_i} y_i(v_1, ..., \underset{i}{q}, ...v_n)dq \tag{1.6}$$

where $h_i(\underline{v}_i, \mathbf{v}_{-i})$ are some arbitrary functions that do not depend on the bidder i's value. Using this characterization, we can consider auctions where all bidders report their true values.

We should note that if DIC is the only restriction, i.e. $C = \{c_{DIC}\}$, then it is never binding for the construction of the favorite's optimal auction, namely $FPA(\{c_{DIC}\}) = FPA$. Indeed, the favorite's optimal auction is dominant strategy incentive compatible, since the favorite does not participate in the bidding and his opponents have a dominant strategy to bid their true values in the optimal auction. As Theorem 1.1 shows, imposing DIC and anonymity is indeed a binding restriction that allows implementing only "true" anonymous auctions. However, I show below that despite Theorem 1.1, $anonymity + DIC$ do not prevent even perfect favoritism. It is almost always possible to send revenue to the favorite.

Theorem 1.2 (transferring revenue to the favorite).

For any direct feasible anonymous and dominant strategy incentive compatible auction $(\mathbf{y}', \mathbf{p}')$ *that generates the equilibrium revenue* $R(v^*, \mathbf{v}_{-1}) = -\sum_{i=1}^{n} t'_i(v^*, \mathbf{v}_{-1})$ *there exists another direct feasible anonymous and dominant strategy incentive compatible auction* $(\mathbf{y}'', \mathbf{p}'')$ *that has the same allocation rule* $\mathbf{y}''(.) = \mathbf{y}'(.)$ *and such a transfer rule* $\mathbf{p}''(.)$ *that implements the same equilibrium transfers for all bidders except the favorite, namely* $t''_j(v^*, \mathbf{v}_{-1}) = t'_j(v^*, \mathbf{v}_{-1})$ *for any* $j \neq 1$*, and the favorite's equilibrium transfer is such that* $t''_1(v^*, \mathbf{v}_{-1}) = t'_1(v^*, \mathbf{v}_{-1}) + R(v^*, \mathbf{v}_{-1})$ *almost always.*

Proof. See Appendix 1.A. □

In other words, it is almost always possible to transfer all collected revenue to the favorite even in an anonymous and DIC auction. Almost always means that the statement is true for all realizations of values, except those where one or more bidders' values coincide with the favorite's value v^*. However, since the distributions of values are atomless and the number of bidders is finite, the probability of such event is equal to zero. The intuition behind this surprising result is that the designer manipulates the transfer rule taking into account that the favorite bids v^*. Then, the auction is constructed in such a way that for any value it is a dominant strategy to bid the true value and the auction transfers all collected revenue to the bidder who submitted a bid v^*, namely to the favorite. Since the probability that there is more than one bidder having a value v^* (in this case, it is not possible to transfer money to the favorite) is zero, such events do not affect the favorite's utility. In order to maintain dominant strategy incentive compatibility, the designer should promise substantially high payments to bidders in the case when nobody bids v^*. In general, if there is no such bidder that reports v^*, the auction would not be budget balanced. However, since the bidder reports v^*, this is not an equilibrium path and hence the auction would be budget balanced in equilibrium for every realization of opponents' values.

Now, consider a standard second-price auction. It is anonymous and DIC. Hence, Theorem 1.2 implies the following:

Corollary 1.1 (transferring revenue in the second-price auction)**.**
If the designer is restricted to using only anonymous and DIC auctions, then there exists a feasible auction that implements the same allocation rule for all bidders and the same transfers in equilibrium for all bidders except the favorite as in the second-price auction. Instead, the favorite receives all collected revenue and has the ex-post utility greater than his value v^* in equilibrium almost always.

It is possible to send the revenue collected in a second-price auction to the favorite for almost all opponents' bids. In the constructed auction, the favorite wins if and only if he has the highest value and almost always pays nothing in this case. If the value of the favorite is not the highest, then the bidder with the highest value obtains the good and the favorite receives monetary transfer equal to the second highest value. Thus, in all cases when the favorite's value is not the highest or the second highest one, the utility obtained by the favorite is equal to the second highest value and strictly exceeds v^*. Since the equilibrium strategy is unique due to Lemma 1.1, the following is true:

Corollary 1.2. *Anonymity and DIC together do not prevent perfect favoritism.*

Thus, even a strong combination of anonymity and dominant strategy incentive compatibility that allows to implement only those rules that are symmetric with respect to real values does not prevent even perfect favoritism. It is important to notice here that unlike the mechanisms discussed in the previous section this construction has the unique equilibrium in dominant strategies. Although my notion of favoritism assumes that the designer can choose among equilibria in undominated strategies, corollary 1.1 implies that the designer can construct an auction that has the unique equilibrium in undominated strategies and in this equilibrium the favorite obtains utility weakly greater than his value for (almost) any realization of opponent values.

Furthermore, notice that I do not claim that the proposed auction is the constrained favorite's preferred auction. In fact, the designer can do even more for his favorite by imposing reserve prices, similar to the unconstrained case. However, the problem of finding the favorite's preferred auction is complicated in this case, because there are many anonymity constraints that have to be satisfied.

The auction proposed in corollary 1.1 is efficient, namely the good is always allocated to the bidder with the highest value. Thus, if the designer is restricted to using only efficient auctions, she can achieve perfect favoritism while implementing efficient auctions. Thus, we can formulate the following corollary.

Corollary 1.3. *It is possible to achieve efficiency and perfect favoritism simultaneously.*

In order to reduce favoritism, it is important to prevent the designer from sending all revenue to her favorite. Since anonymity and dominant strategy incentive compatibility do not restrict the designer's ability to transfer money to her favorite, an additional restriction should be imposed.

1.7 Non-Positive Transfers

Definition 1.12 (non-positive transfers).

A feasible auction FA satisfies non-positive transfers (NT, feasible under c_{NT}) if for any vector of bids $\mathbf{b} \in \mathbf{B}$ and any allocation $a \in A$

$$\mathbf{p}(a, \mathbf{b}) \leq \mathbf{0}$$

This restriction is crucial for preventing favoritism. We see from Theorem 1.2 that the designer always wants to transfer all collected revenue to her favorite. Even anonymity and DIC are insufficient to prevent the designer from doing this. It is clear that to prevent favoritism this possibility should be excluded. The natural way to do this is to impose a restriction that allows the designer to only collect money from the agents but not to give it. In other words, the principal may want to prohibit positive transfers.

If NT is the only restriction, namely $C = \{c_{NT}\}$, then the best that the designer can do is to allocate the good to her favorite for sure, independent of all bids. Imposing anonymity restriction jointly with NT, that is $C = \{c_A, c_{NT}\}$, does not particularly help. Again, with the result of Deb and Pai (2017) the designer is still able to allocate the good to her favorite (in some equilibrium) without extracting money from him. $C = \{c_{DIC}, c_{NT}\}$ works in the same way as $C = \{c_{NT}\}$, since allocating the good to the favorite independent of the bids is trivially incentive compatible. However, the combination of all three constraints, $C = \{c_A, c_{DIC}, c_{NT}\}$ substantially limits the scope of favoritism. In this case, as I show below, the designer has to use stochastic mechanisms and pool bidders having values in some regions to one specific value. I provide a complete solution to the problem in the case with two bidders. In the case with many bidders, it seems impossible to obtain an analytical solution due to the increased number of anonymity constraints that have to be satisfied. In general, there are $n!$ constraints only due to anonymity. Since the problem of maximizing the favorite's utility is asymmetric, it is incredibly difficult to take all of them into account. However, even the case with two bidders is sufficiently rich to shed some light on what is happening here.

There are two bidders, with bidder 1 being a favorite and bidder 2 being his opponent. For this case, we are able to characterize the $FPA(\{c_A, c_{DIC}, c_{NT}\})$ for any continuously differentiable distribution of the opponent's value $F(v)$, $v \in [0, \overline{v}]$. Note that I allow the case when the favorite's value is greater than any possible value of his opponent and thus $v^* > \overline{v}$ is possible. In order to formulate the main result, I need some additional notations. Denote $G : \mathbb{R}_+ \to \mathbb{R}_+$,

$$G_x(z) = \left\{ \begin{array}{c} conc_x \langle F(z) \rangle, \text{ if } z \le x \le \overline{v} \\ conc_{\overline{v}} \langle F(z) \rangle, \text{ if } z \le \overline{v} < x \\ 1 + \left(\lim\limits_{q \to x^-} \frac{dG_x(q)}{dz} \right) * (z - x), \text{ if } z, \overline{v} > x \\ 1, \text{ if } z, x > \overline{v} \end{array} \right\}^8$$

where $conc_x \langle F(z) \rangle$ is the lowest function that is concave, weakly greater than $F(v)$ and takes a value equal to 1 at the point $z = x$. It is illustrated in figure 1.1 for the case $x < \overline{v}$. Denote $g_x(z) := dG_x(z)/dz$[9]

Proposition 1.1 (FPA under anonymity, DIC and NT)**.**

Assume that there are only two bidders. The favorite's preferred auction under anonymity, DIC and NT allocates the object to a bidder with the lowest $g_{v^*}(v_i)$. In the case of equality, a simple lottery is used to determine a winner. Transfers are computed by (1.6) with $h_i(\underline{v}_i, \mathbf{v}_{-i}) = 0$.

Proof. See Appendix 1.A. □

The auction described in Proposition 1 has a very clear economic description and is easy to implement. I call it a second-price auction with pooling. It is possible to think about a standard second-price auction with a slight modification; namely, there are intervals on the value domain such that if a bidder reports a value in one of these regions, he is treated as a bidder having

[8] $\lim\limits_{q \to x^-}$ is the limit from the left at the point x. We use it to define $G_x(z)$ for values beyond the domain of F.

[9] If $z = q \in \{\overline{v}, x\}$ then $G_x(z)$ is not differentiable. In this case, let the derivative $g_x(q)$ equal the limit from the left of $g_x(z)$ at the point $z = q$, i.e. $g_x(q) = \lim\limits_{z \to q^-} \frac{dG_x(z)}{dz}$.

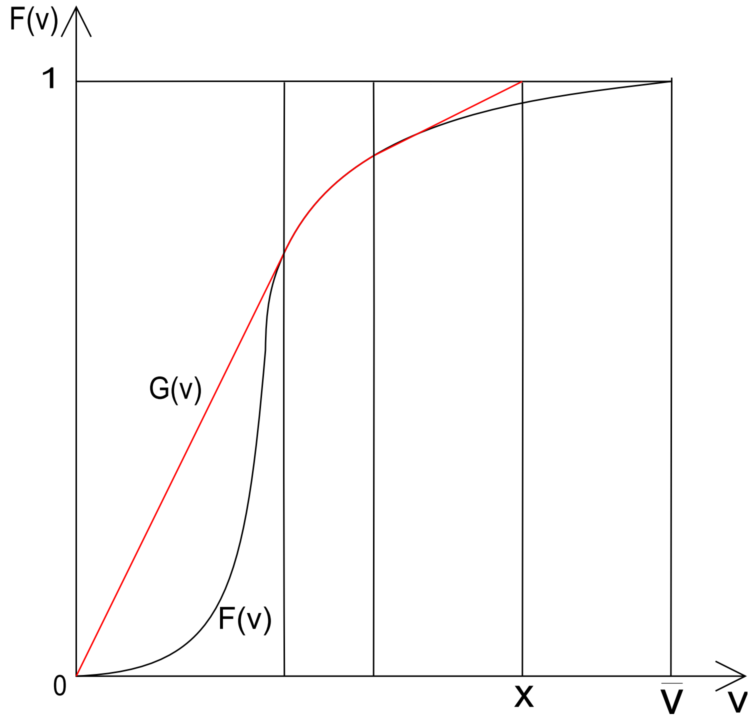


Figure 1.1: Concave envelope

value in the middle of this interval. For example, assume that there is only one such "pooling" or "lottery" interval (a, b). Suppose the second bidder has a value $v_2 \in (a, b)$. If a value of the favorite v^* is greater than b, he obtains the object and pays $(a+b)/2$. If $v^* \in (a, b)$ both bidders have an equal chance $1/2$ of winning the object. In the case of a win, the winner pays a, i.e. the left bound of the interval. If $v^* < a$, the second bidder obtains the object and pays v^*. The structure of pooling and contest regions is illustrated in figure 1.2.

There are two reasons why pooling arises in the solution. The first one is that it is a way for the designer to give the object to her favorite when the opponent's value is higher. In order to better understand this, I can formulate the following proposition:

Proposition 1.2 (pooling at the top)**.**

For any $v^* < \overline{v}$, there exists a cutoff $\widehat{v} < v^*$, such that the $FPA(\{c_A, c_{DIC}, c_{NT}\})$ pools all bidders with values above $\widehat{v}$. This cutoff $\widehat{v}(v^*)$ is a monotone increasing function of the favorite's value.

Proof. See Appendix 1.A. □

Thus, the designer prefers to use lottery if a value of a second bidder is higher than a value of her favorite or lower, but sufficiently close. In the first case, it gives a chance to allocate the object to her favorite and in the second case it reduces payments in the case of win.

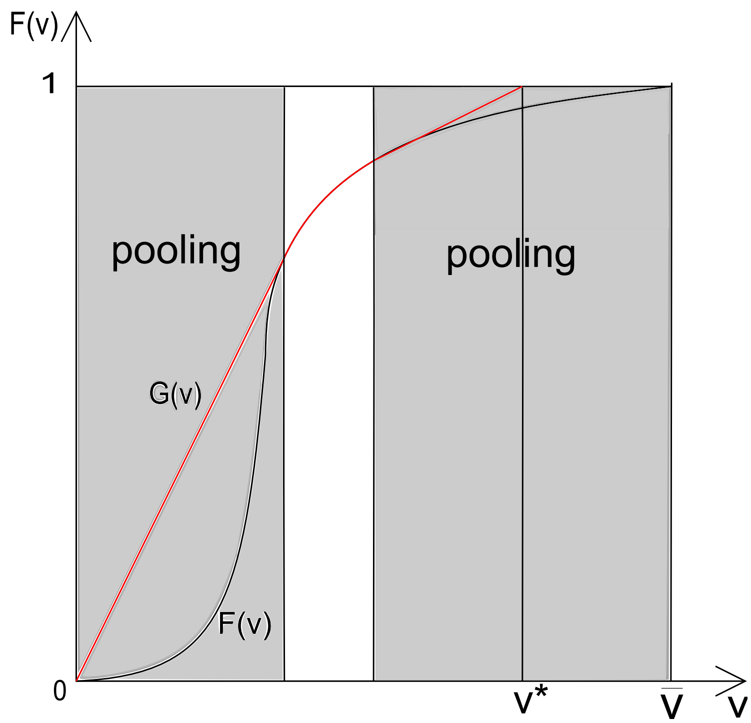


Figure 1.2: Second-price auction with pooling

Example 1.2. $F(v) = \sqrt{v}$, $v \in [0, 1]$, $v^* = 0.75$

By propositions 1.1 and 1.2, we know that since the distribution function is concave, the restricted favorite's preferred auction is a second-price auction with one pooling region at the top. The pooling cutoff $\widehat{v}$ can be computed from $\sqrt{\widehat{v}} + \frac{1}{2\sqrt{\widehat{v}}}(v^* - \widehat{v}) = 1$, which gives $\widehat{v} = 0.25$. Thus, the result of the favorite's optimal auction is as follows: if the opponent has a value lower than 0.25, *the favorite obtains the object and pays an amount equal to his opponent's value. If the opponent has a value higher than* 0.25, *there will be a lottery among two bidders and the winner pays* 0.25. *The expected utility of the favorite in the favorite's preferred auction is* 0.458, *which is larger than the utility* 0.433 *of a standard second-price auction and* 0.375 *of a standard lottery.*

The second reason why pooling may be optimal is that it reduces expected payments made by the favorite when his value is substantially higher than a value of his competitor. Indeed, in the regions where pooling is used the graph of the cumulative distribution function lies below the straight line and hence the average value in each such region is smaller than the middle value. Suppose that a value of the second bidder belongs to that region. In the case of no pooling, the first bidder would have to pay in expectation the amount that is equal to the average value. When pooling is used he pays only the amount equal to the middle value and this reduces payments. We can also observe how the $FPA(\{c_A, c_{DIC}, c_{NT}\})$ depends on the favorite's value.

Proposition 1.3 (comparative statics).

If $v^* < \overline{v}$, then the only difference between $FOA(\{c_A, c_{DIC}, c_{NT}\})$ for different values is the size of a pooling region above the cutoff function $\widehat{v}(v^*)$. For any $v^* > \overline{v}$, $FPA(\{c_A, c_{DIC}, c_{NT}\})$ is the same as the one for $v^* = \overline{v}$.

Proof. See Appendix 1.A. □

Thus, if the favorite's value becomes smaller, the designer wants to increase pooling in the region of high values and keep the same allocation rule for low realizations of values. Thus, the change of the favorite's value has only a local effect on the auction design. If the favorite has a value higher than any possible value of his opponent, then the optimal auction does not depend on the specific value.

One can easily see that intra-auction favoritism and perfect favoritism are not possible under restrictions of anonymity, DIC and NT. However, inter-auction favoritism can be successively used by the designer to make the auction better for her favorite. In the next section, I show how any form of favoritism can be prevented by adding one additional constraint.

1.8 Deterministic Auctions

Definition 1.13 (deterministic auctions).

A feasible auction FA is deterministic (DA, feasible under c_{DA}) if for any two bidders $i \neq j$ and for any two bids b_i, b_j submitted by these bidders, such that $b_i \neq b_j$, the allocation is such that $y_i \in \{0,1\}$ and $y_j \in \{0,1\}$.

This restriction does not allow the designer to use any randomization in the case when submitted bids are different. For example, the second price with probability 1. However, a second-price auction with pooling is not DA, because it uses randomization to determine the winner.

Definition 1.14 (second-price auction with a generalized reserve price).

An auction is called a second-price auction with a generalized reserve price if the following is satisfied:

$$\begin{cases} \left\{ \begin{array}{c} y_i(\mathbf{b}) = 1 \\ t_i(\mathbf{b}) = -\max_{j\neq i}(b_j, r(\mathbf{b}_{-i})) \end{array} \right\}, & \text{if } b_i > \max\limits_{j\neq i}(b_j, r(\mathbf{b}_{-i})) \\ y_i(\mathbf{b}) = t_i(\mathbf{b}) = 0, & \text{if } b_i < \max\limits_{j\neq i}(b_j, r(\mathbf{b}_{-i})) \end{cases}$$

where $r: \mathbb{R}^{n-1} \to \mathbb{R}$ is a componentwise symmetric function [10].

[10] In the zero probability case, when there are two or more bids that are equal and the highest among all bids,

The difference between a second-price auction with a generalized reserve price and a standard second-price auction is only that the reserve price is not the same for different bidders but rather for each player it may depend on bids made by his opponents. If a generalized reserve price is a constant function, i.e. $r(\mathbf{x}) = const\ \forall \mathbf{x} \in \mathbb{R}^{n-1}$, we obtain a second-price auction with a standard reserve price. In a special case of a zero reserve price, we get a standard second-price auction. The next important result characterizes the set of all feasible auctions.

Theorem 1.3 (auctions feasible under full set of constraints).

Any feasible anonymous, DIC, deterministic auction with nonpositive transfers is a second-price auction with a generalized reserve price.

Proof. See Appendix 1.A. □

This result shows that the set of all auctions feasible under $C = \{c_A, c_{DIC}, c_{NT}, c_{DA}\}$ is only a very specific class of auctions, described above. Hence, the ability to favor some participant is substantially limited. The next result shows that there is actually no scope for favoritism in this case.

Proposition 1.4 (no favoritism).

For any favorite's value v^*, the favorite's preferred auction feasible under restrictions of anonymity, strategyproofness, nonpositive transfers and determinism is a standard second-price auction.

Proof. See Appendix 1.A. □

From Theorem 1.3, we know that all the designer can do is to choose some auction from a class of second-price auctions with a generalized reserve price. Proposition 1.4 shows that the reserve price that makes the favorite better off is zero. Hence, $FPA(\{c_A, c_{DIC}, c_{NT}, c_{DA}\})$ is a standard second-price auction.

Thus, I have shown that if the designer is allowed to use only anonymous, dominant strategy incentive compatible, deterministic auctions such that bidders never obtain money from it, then any kind of favoritism is impossible. The best the designer can do is always choose a second-price auction independent of her favorite's value and value distributions. Note also that although this set of restrictions substantially limits the freedom to choose the auction format, the revenue maximizing auction is still available for the designer if the agents are symmetric. Indeed, in this case the revenue maximizing auction will be the second-price auction with a reserve price that

only a symmetric lottery can be used to determine the winner who obtains the object and pays his bid if it is greater than a generalized reserve price. Note also that due to symmetry the reserve price is the same for all agents with the highest bids.

can be implemented under $\{c_A, c_{DIC}, c_{NT}, c_{DA}\}$. Such a set of restrictions guarantees collected revenue as the revenue in a second-price auction even if the designer only cares about the favorite.

1.9 Hierarchy of Restrictions

Based on the previous results, we can understand how the restrictions interact with each other, namely how each restriction helps to prevent favoritism depending on the other restrictions in place.

First, DA is a binding restriction if and only in the situation without this restriction the favorite's optimal auction uses lotteries. As we have seen above, it can be only the case when anonymity and DIC are imposed jointly. Next, without anonymity, the requirement of dominant strategies does not restrict the designer, because in this case the situation is equivalent to a situation where the designer has her own value of v^* and there is no favorite. In this case, the optimal mechanism is DIC even without imposing DIC. The opposite direction is also true, namely without the DIC restriction, the requirement of anonymity does not restrict the designer. This follows from Deb and Pai (2017) argument. Hence, we conclude that the anonymity restriction is binding if and only if the DIC restriction is imposed and vice versa. From the previous discussion, we have seen that NT restriction binds if DA is not imposed. Since DA is binding if and only if a combination of *anononymity* $+ DIC$ is present, we are left to consider only the case with *anonymity* $+ DIC + DA$ as restrictions to fully understand the role of NT. Note that in this case the allocation rule is uniquely determined, since due to *anonymity* and DA the allocation rule has to be such that the highest bid wins the auction for sure, which jointly with DIC implies that the bidder with the highest value wins the auction. Hence, the favorite's preferred auction is the one described in corollary 1.1, where the designer transfers all collected revenue to her favorite in equilibrium. Thus, NT always reduces the scope of favoritism independent of other restrictions imposed. Thus, we obtain the following result:

Proposition 1.5 (hierarchy of restrictions)**.**

The set of restrictions comprising anonymity, DIC, NT, DA forms a hierarchy with NT at the top, $DIC + NT$ in the middle and DA at the bottom. NT restricts the scope of favoritism independent of whether other constraints are imposed. DIC reduces the scope of favoritism if and only if anonymity is imposed and vice versa. DA reduces the scope of favoritism if and only if a combination anonymity $+ DIC$ is imposed.

1.10 Discussion

The fact that the designer knows not only the identity of the favorite but also his value is the main driving force of favoritism in choosing the particular auction format. If the designer did not know the value of the favorite, a combination of anonymity and dominant strategy incentive compatibility would turn the problem of favoritism to a problem of buyers' welfare maximization irrespective of their identities. Hence, anonymity and dominant strategy incentive compatibility would be a sufficient condition to prevent any kind of favoritism.

My results are robust to the imperfect knowledge of the designer. In particular, if the designer's belief v_d^* about the favorite's value is sufficiently precise, namely there exist such ε and δ such that $\Pr(|v_d^* - v^*| > \varepsilon) < \delta$, she can provide him the interim utility $U_d(v^*)$ such that $U_d(v^*) > U(v^*) - O(\max\{\varepsilon, \delta\})$, where $U(v^*)$ is the utility of the favorite in the perfect knowledge case. When ε and δ are sufficiently small, the favorite's expected utility approaches the perfect information case. Thus, our results are not simply an artifact of precise information about the favorite's value.

For a set of restrictions which that comprises anonymity, DIC and non-positive transfers, I have analyzed the case with only two bidders. The favorite's preferred auction for many agents and one favorite would have the similar properties and would be a second-price auction with pooling. For many agents, pooling in general is partial, namely not all bidders above some cutoffs are pooled, although due to increased number of anonymity restrictions it is much more difficult to compute. In order to observe what happens when we increase the number of bidders, we can compare a standard lottery and a second-price auction. The expected utility of the favorite from participation in a lottery is $U_1(v^*) = v^*/n$ and from participation in a second-price auction is $U_1(v^*) = F^n(v^*) * (v^* - E[v_{(1)}|v_{(1)} \leq v^*])$, where $v_{(1)}$ is the first-order statistic out of $(n - 1)$ variables. Obviously, both expressions go to zero when n increases, although the speed of convergence is $1/n$ in the case of a lottery and $F^n(v^*)$ in the second-price auction. Since for any $v^* < \overline{v}$ we have $F^n(v^*) < v^*/n$ we can conclude that when the number of bidders increases, all of them would prefer a lottery to an auction[11].

1.11 Conclusion

In this chapter, I have analyzed the problem of favoritism in auctions from a mechanism design perspective. In my model, the designer has one favorite among the bidders, whose value is known

[11] This does not imply that a lottery is socially preferable to an auction. See Condorelli (2012) for a description of a socially optimal mechanism.

to the designer. I have characterized feasible auctions that the designer can implement to maximize the utility of her favorite under different sets of restrictions on these auctions. Deb and Pai (2017) have shown that assuming that the designer can choose between different undominated equilibria, anonymity is not a binding restriction for the designer. I have shown that even if the designer is restricted not only by anonymity but also by dominant strategy incentive compatibility, it is insufficient to prevent perfect favoritism. Namely, the designer is almost always able to transfer all collected revenue to her favorite in any auction. Hence, it is possible to guarantee him the interim utility greater than or equal to his value in the unique equilibrium of the constructed auction. To prevent this possibility, I additionally impose the non-positive transfers restriction. Subsequently, the designer cannot discriminate bidders within any auction. However, although intra-auction favoritism is not possible, the inter-auction favoritism could still be possible, whereby the designer chooses different auction formats for different favorite's values. I have shown that the favorite's preferred auction is a second-price auction with pooling where the designer commits to not distinguishing some value reports. The size of the pooling region for the highest values depends on the favorite's value. Thus, the designer uses inter-auction favoritism. Finally, I have shown that it is possible to completely prevent any form of favoritism if the designer is restricted to using only deterministic auctions in addition to anonymity, dominant strategy incentive compatibility and non-positive transfers restrictions. In this case, any feasible mechanism is a second-price auction with a generalized reserve price, whereby the reserve price for each bidder depends on bids submitted by other bidders. The favorite's preferred auction in this class is a standard second-price auction without any reserve price.

My results imply that while delegating the decision about the auction format choice to the designer, the principal should care about how much freedom should be given to the designer and in what way this freedom can be limited. If the final goal of the principal is revenue maximization, then along with anonymity and dominant strategy incentive compatibility, restrictions of non-positive transfers and deterministic auctions should be imposed. Non-positive transfers would help to prevent discrimination of bidders via transfers. Determinism is used to sustain competition, since without it the designer would like to make it less intensive by using lotteries.

Traditional problems of mechanism design (revenue maximization, efficiency maximization, social welfare maximization) are symmetric and hence they have symmetric solutions. I have considered essentially asymmetric problems and have found symmetric (anonymous) solutions for them. Thus, my results can also serve as a mathematical approach to solving such kind of problems.

1.A Appendix A

Proof of Lemma 1.1. Similar to Maskin and Laffont (1979), if $\beta_i^*(v_i)$ is a dominant stratefy for bidder i, then for $b_i = \beta_i^*(v_i)$ and for any $b_j \in \{\oslash\} \cup B_j$, $j \neq i$:

$$U_i(v_i|\mathbf{b}) = \int_{\underline{v}_i}^{v_i} y_i(b_1, ..., \beta_i^*(q), ..., b_n))dq + h_i(\underline{v}_i, \mathbf{b}_{-i}).$$

Since $\beta_i^*(v_i)$ and $\beta_i^{**}(v_i)$ are both dominant strategies:

$$\begin{aligned} U_i(v_i|b_1, ..., \beta_i^*(v_i), ..., b_n) &\geq U_i(v_i|b_1, ..., \beta_i^{**}(v_i), ..., b_n), \\ U_i(v_i|b_1, ..., \beta_i^{**}(v_i), ..., b_n) &\geq U_i(v_i|b_1, ..., \beta_i^*(v_i), ..., b_n). \end{aligned}$$

Hence,

$$U_i(v_i|b_1, ..., \beta_i^{**}(v_i), ..., b_n) = U_i(v_i|b_1, ..., \beta_i^*(v_i), ..., b_n).$$

Taking a derivative of both sides with respect to v_i we obtain for any v_i:

$$y_i(b_1, ..., \beta_i^*(v_i), ..., b_N) = y_i(b_1, ..., \beta_i^{**}(v_i), ..., b_N).$$

Then,

$$\begin{aligned} &t_i(b_1, ...\beta_i^*(v_i), ...b_N) = U_i(v_i|b_1, ..., \beta_i^*(v_i), ..., b_N) - v_i y_i(b_1, ..., \beta_i^*(v_i), ..., b_N) = \\ &= U_i(v_i, |b_1, ..., \beta_i^{**}(v_i), ..., b_N) - v_i y_i(b_1, ..., \beta_i^{**}(v_i), ..., b_N) = t_i(b_1, ..., \beta_i^{**}(v_i), ...b_N). \end{aligned}$$

□

Proof of Lemma 1.2. Suppose that $\beta^*(v)$ is a dominant strategy for an agent i and consider some value v from the intersection of possible values sets for bidders i and j :

$$U_i(v|b_1, ..., \underset{i}{\beta^*(v)}, ..., b_N) \geq U_i(v|b_1, ..., \underset{i}{\beta(v)}, ..., b_N) \text{ for any } \beta(v) \text{ and } b_k \in \{\oslash\} \cup B_k,\ k \neq i.$$

This means that for any $\widetilde{b}$:

$$\begin{aligned} & vy_i(b_1, ..., \underset{i}{\beta^*(v)}, ..., \underset{j}{\widetilde{b}}, ..., b_N) + t_i(b_1, ..., \underset{i}{\beta^*(v)}, ..., \underset{j}{\widetilde{b}}, ..., b_N) \geq \\ \geq\ & vy_i(b_1, ..., \underset{i}{\beta(v)}, ..., \underset{j}{\widetilde{b}}, ..., b_N(v_N)) + t_i(b_1, ..., \underset{i}{\beta(v)}, ..., \underset{j}{\widetilde{b}}, ..., b_N). \end{aligned}$$

If we switch bids of agents i and j, by anonymity agent j should have the same allocation as an agent i had before. Hence, the previous inequality can be rewritten as:

$$\begin{aligned} & vy_j(b_1,...,\underset{i}{\widetilde{b}},...,\underset{j}{\beta^*(v)},...,b_N)+t_i(b_1,...,\underset{i}{\widetilde{b}},...,\underset{j}{\beta^*(v)},...,b_N) \geq \\ \geq\ & vy_j(b_1,...,\underset{i}{\widetilde{b}},...,\underset{j}{\beta(v)},...,b_N)+t_i(b_1,...,\underset{i}{\widetilde{b}},...,\underset{j}{\beta(v)},...,b_N). \end{aligned}$$

Hence,

$$U_j(v|b_1,...,\underset{j}{\beta^*(v)},...,b_N) \geq U_j(v|b_1,...,\underset{j}{\beta(v)},...,b_N).$$

□

Proof of Theorem 1.2. Step 1 (Construction of $h''_i(\underline{v}_i,\mathbf{v}_{-i})$).

Consider some anonymous and DIC auction that has the allocation rule $\mathbf{y}'(\mathbf{v})$ and the transfer rule $\mathbf{t}'(\mathbf{v})$. The new constructed auction also has to be DIC. By (1.6) functions $\{h'_i(\underline{v}_i,\mathbf{v}_{-i})\}_{i=1}^n$, $\{h''_i(\underline{v}_i,\mathbf{v}_{-i})\}_{i=1}^n$ have to satisfy:

$$t'_i(\mathbf{v}) = -v_iy'_i(\mathbf{v})+h'_i(\underline{v}_i,\mathbf{v}_{-i})+\int_{\underline{v}_i}^{v_i} y'_i(v_1,...,\underset{i}{q},...v_n)dq \tag{1.7}$$

$$t''_i(\mathbf{v}) = -v_iy''_i(\mathbf{v})+h''_i(\underline{v}_i,\mathbf{v}_{-i})+\int_{\underline{v}_i}^{v_i} y''_i(v_1,...,\underset{i}{q},...v_n)dq \tag{1.8}$$

It is required that the new allocation rule is the same as before. Accordingly, for any vector of reported values $\mathbf{v}$, we must have $y'_i(\mathbf{v}) = y''_i(\mathbf{v})$. However, transfers should be (almost always) the same only in equilibrium. In equilibrium, the favorite always reports v^*. Hence, only vectors $\mathbf{v} = (v^*,\mathbf{v}_{-1})$ can be on equilibrium path. For any i and for any $\mathbf{v}_{-i}$, define

$$h''_i(\underline{v}_i,\mathbf{v}_{-i}) := h'_i(\underline{v}_i,\mathbf{v}_{-i}) \tag{1.9}$$

if at least one component of $\mathbf{v}_{-i}$ is equal to v^* and

$$\begin{aligned} h''_i(\underline{v}_i,\mathbf{v}_{-i}) := {} & v^*y''_i(v_1,...,\underset{i}{v^*},...v_n)-\int_{\underline{v}_i}^{v^*} y''_i(v_1,...,\underset{i}{q},...,v_n)dq+ \\ & +\sum_{j\neq i} v_jy''_j(v_1,...,\underset{i}{v^*},...v_n)-\sum_{j\neq i}\int_{\underline{v}_j}^{v_j} y''_j(v_1,...,\underset{i}{v^*},...,\underset{j}{q},...,v_n)dq-\sum_{j\neq i} h''_j(\underline{v}_j,\mathbf{v}_{-j}|v_i=v^*) \end{aligned} \tag{1.10}$$

if none of $\mathbf{v}_{-i}$ components is equal to v^*, where $h'_j(\underline{v}_j,\mathbf{v}_{-j}|v_i = v^*)$ means that the value of component v_i in $\mathbf{v}_{-j}$ is replaced by v^*.

Step 2. (Computing transfers).

Equation (1.8) then uniquely defines $t''_i(\mathbf{v})$ given $y''_i(\mathbf{v})$ and $h''_i(\underline{v}_i,\mathbf{v}_{-i})$. Thus, if $\mathbf{v}_{-i}$ has a component equal to v^*, then $h''_i(\underline{v}_i, \mathbf{v}_{-i}) = h'_i(\underline{v}_i, , \mathbf{v}_{-i})$ and, hence,

$$t''_i(\mathbf{v}) = t'_i(\mathbf{v}). \tag{1.11}$$

If all components of $\mathbf{v}_{-i}$ are different from v^*, plugging the expression (1.10) to (1.8), using $\mathbf{y}'' = \mathbf{y}'$ and $h''_j(\underline{v}_j, \mathbf{v}_{-j}|v_i = v^*) = h'_j(\underline{v}_j, \mathbf{v}_{-j}|v_i = v^*)$, $j \neq i$ we obtain

$$\begin{aligned}
& t''_i(\mathbf{v}) = -v_i y'_i(\mathbf{v}) + \int_{v^*}^{v_i} y'_i(v_1, ..., \underset{i}{q}, ...v_n)dq + v^* y'_i(v_1, ..., \underset{i}{v^*}, ...v_n) + \\
& + \sum_{j\neq i} v_j y'_j(v_1, ..., \underset{i}{v^*}, ...v_n) - \sum_{j\neq i} \int_{\underline{v}_j}^{v_j} y'_j(v_1, ..., \underset{i}{v^*}, ..., \underset{j}{q}, ..., v_n)dq - \sum_{j\neq i} h'_j(\underline{v}_j, \mathbf{v}_{-j}|v_i = v^*) = \\
= \; & -v_i y'_i(\mathbf{v}) + \int_{v^*}^{v_i} y'_i(v_1, ..., \underset{i}{q}, ...v_n)dq + v^* y'_i(v_1, ..., \underset{i}{v^*}, ...v_n) + \\
& + \sum_{j\neq i} v_j y'_j(v_1, ..., \underset{i}{v^*}, ...v_n) - \sum_{j\neq i} t'_j(v_1, ..., \underset{i}{v^*}, ..., v_n) - \sum_{j\neq i} v_j y'_j(v_1, ..., \underset{i}{v^*}, ..., v_n) = \\
= \; & -v_i y'_i(\mathbf{v}) + \int_{v^*}^{v_i} y'_i(v_1, ..., \underset{i}{q}, ...v_n)dq + v^* y'_i(v_1, ..., \underset{i}{v^*}, ...v_n) - \sum_{j\neq i} t'_j(v_1, ..., \underset{i}{v^*}, ..., v_n),
\end{aligned} \tag{1.12}$$

where we also used that (1.7) implies $\sum_{j\neq i} \int_{\underline{v}_j}^{v_j} y'_j(v_1, ..., \underset{i}{v^*}, ..., \underset{j}{q}, ..., v_n)dq + \sum_{j\neq i} h'_j(\underline{v}_j, \mathbf{v}_{-j}|v_i = v^*) = \sum_{j\neq i} t'_j(v_1, ..., \underset{i}{v^*}, ..., v_n) + \sum_{j\neq i} v_j y'_j(v_1, ..., \underset{i}{v^*}, ..., v_n)$. Now, we need to verify that the constructed auction satisfies anonymity and in equilibrium it almost always implements the described transfers.

Step 3. (Check anonymity of $(\mathbf{y}'', \mathbf{t}'')$).

Since $(\mathbf{y}', \mathbf{t}')$ is an anonymous auction and $\mathbf{y}'' = \mathbf{y}'$, the allocation rule is trivially symmetric. Now, consider $\mathbf{t}''(\mathbf{v})$. If $\mathbf{v}_{-i}$ has a component equal to v^*, then $t''_i(\mathbf{v}) = t'_i(\mathbf{v})$. Since $t'_i(\mathbf{v})$ is symmetric, then $t''_i(\mathbf{v})$ is also symmetric. If all components of $\mathbf{v}_{-i}$ are different from v^*, then $t''_i(\mathbf{v})$ is described by expression (1.12), which does not depend on $\{\underline{v}_i\}_{i=1}^n$ and has only symmetric functions inside. Thus, anonymity is satisfied.

Step 4. (Equilibrium transfers).

In equilibrium the favorite reports v^*. Hence, (1.11) implies that $t''_i(v^*, \mathbf{v}_{-1}) = t'_i(v^*, \mathbf{v}_{-1})$ for all bidders, except the favorite. Since the number of bidders is finite and the distributions are strictly increasing the probability that some other bidder is going to report v^* is zero. Thus, the favorite's transfer in equilibrium is almost always described by (1.12) and plugging $v_1 = v^*$, we obtain $t''_1(v^*, \mathbf{v}_{-1}) = -\sum_{j\neq 1} t'_j(v^*, \mathbf{v}_{-1}) = t'_1(v^*, \mathbf{v}_{-1}) + R(v^*, \mathbf{v}_{-1})$. The no-deficit requirement is trivially satisfied in equilibrium, because the constructed auction transfers all revenue to the favorite making the budget balanced. This completes the proof. □

Lemma 1.3. *If $\{c_A, c_{DIC}, c_{NT}\} \subset C$, then in any direct $FA(C)$: $h_i(0, \mathbf{v}_{-i}) = 0$ for any i, $\mathbf{v}_{-i}$*

Proof of Lemma 1.3. Suppose that bidder i has a value $v_i = 0$. Then, by (1.6)

$$t_i(\mathbf{v}) = U_i(\mathbf{v}) = h_i(0, \mathbf{v}_{-i})$$

Since $c_{NT} \in C$, we should have $h_i(0, \mathbf{v}_{-i}) = t_i(\mathbf{v}) \leq 0$ for any $\mathbf{v}_{-i}$. Simultaneously, $U_i(\mathbf{v})$ has to be positive otherwise it would not be the dominant strategy to report the true value and bidder i could exclude himself from participation. Hence, $h_i(0, \mathbf{v}_{-i}) \geq 0$ should also hold for any $\mathbf{v}_{-i}$. Combining last two inequalities we have $h_i(0, \mathbf{v}_{-i}) = 0$. □

Proof of Proposition 1.1. Using characterization (1.6) and Lemma 1.3, transfers $t_i(v_1, v_2)$ are fully determined by the allocation rule $y_i(v_1, v_2)$. Since the function G_x has a different form depending on a relationship between v^* and $\overline{v}$, we consider possible cases separately.

Case 1. $v^* \leq \overline{v}$.

By the anonymity restriction we need to specify an allocation rule only on the cone $\Gamma = \{\mathbf{v} = (v_1, v_2) \in [0, \overline{v}]^2 : v_1 \geq v_2\}$. Indeed, suppose we have specified some allocation rule on Γ. Then, for all reported values $(v_1, v_2) \notin \Gamma$ we have $v_2 > v_1$. Since for $v_2 > v_1$ the bid vector $(v_2, v_1) \in \Gamma$, we know the allocation probabilities $y_1(v_2, v_1)$ and $y_2(v_2, v_1)$. Then, by anonymity we have the allocation for $(v_1, v_2) \notin \Gamma$ as $y_1(v_1, v_2) = y_2(v_2, v_1)$ and $y_2(v_1, v_2) = y_1(v_2, v_1)$.

To illustrate our proof we plot for convenience simultaneously two things on the same figure. The first one is a graph of a distribution function $F(v)$ of the opponent's value. The second one is the value space (v_1, v_2). The auction described in the statement implies that the whole value space is cut into a certain number of triangles and rectangles (see figure 1.3 as an example). I use R_i to talk about region i on the figure 1.3. The rectangles can be only of two types: 1) interior rectangles, like R_2, in general there could be many of them; and 2) at most one boundary rectangle with values $v_1 \geq v^*$ inside, like R_4. For all pairs (v_1, v_2) inside each such rectangle $y_1(v_1, v_2) = 1$ and $y_2(v_1, v_2) = 0$. Triangles can be of three types: 1) interior triangles like R_3, 2) the unique boundary triangle containing $v_1 = v_2 = 0$, like R_1, 3) the unique boundary triangle with values $v_1 \geq v^*$ inside, like R_5. If a triangle is the region, where $g_{v^*}(v_1)$ is constant (R_1, R_5 on figure 1.3), then $y_1(v_1, v_2) = y_2(v_1, v_2) = 1/2$, for all pairs (v_1, v_2) inside this triangle. If a triangle lies in the region, where $g_{v^*}(v_1)$ is strictly decreasing (R_3 on Figure 3), then $y_1(v_1, v_2) = 1$ and $y_2(v_1, v_2) = 0$. Our task is to prove that the described allocation is indeed optimal for the favorite having a value v^*.

Using a notation $k(v_1, v_2) := y_1(v_1, v_2) + y_2(v_1, v_2)$, where $0 \leq k(v_1, v_2) \leq 1$, we can rewrite

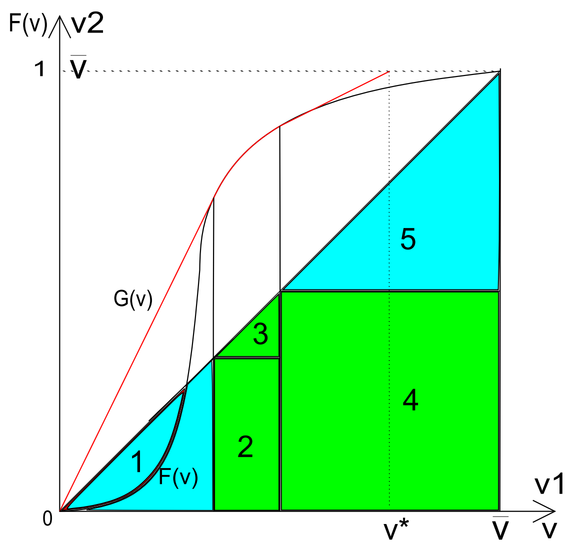


Figure 1.3: Illustration of a proof

the interim utility of the first agent:

$$
\begin{aligned}
U_1(v^*) = \int_0^{\overline{v}} (v^* y_1(v^*, v_2) + t_1(v^*, v_2)) f(v_2) dv_2 = \int_0^{\overline{v}} (\int_0^{v^*} y_1(v_1, v_2) dv_1) f(v_2) dv_2 = \\
= \int_0^{v^*} \int_{v_2}^{v^*} y_1(v_1, v_2) f(v_2) dv_1 dv_2 + \int_0^{v^*} \int_{v_1}^{\overline{v}} y_1(v_1, v_2) f(v_2) dv_2 dv_1 = \\
= \int_0^{v^*} \int_{v_2}^{v^*} y_1(v_1, v_2) f(v_2) dv_1 dv_2 + \int_0^{v^*} \int_{v_1}^{\overline{v}} y_2(v_2, v_1) f(v_2) dv_2 dv_1 = \\
= \int_0^{v^*} \int_{v_2}^{v^*} y_1(v_1, v_2) f(v_2) dv_1 dv_2 + \int_0^{v^*} \int_{v_2}^{\overline{v}} y_2(v_1, v_2) f(v_1) dv_1 dv_2 = \\
= \int_0^{v^*} \int_{v_2}^{v^*} y_1(v_1, v_2) f(v_2) dv_1 dv_2 + \int_0^{v^*} \int_{v_2}^{\overline{v}} [k(v_1, v_2) - y_1(v_1, v_2)] f(v_1) dv_1 dv_2, \qquad (1.13)
\end{aligned}
$$

where the equality in the first line follows from (1.6) and Lemma 1.3, the next one is changing the order of integration, then we apply anonymity, and finally we switch notations of v_1 and v_2 in the second summand. Here, we can notice that it is always optimal to put $k(v_1, v_2) = 1$ for any v_1, v_2. This means that it is never optimal to throw the object away. From now onwards, I will skip the term $\int_0^{v^*} \int_{v_2}^{\overline{v}} k(v_1, v_2) f(v_1) dv_1 dv_2$, which is constant in the $FPA(C)$. Now, we need to maximize (1.13) subject to monotonicity constraints (1.5).

I prove that even separately in each of the described regions, i.e. neglecting global monotonicity constraints, it is not possible to change an allocation rule to increase utility of the favorite. Denote by $z_1, z_2, ...$ such points where $g_{v^*}(v_1)$ changes its type from linear to strictly concave and vice versa. Suppose that there exists any interior or boundary triangle with $(0,0)$ inside, called R_1, such that $g_{v^*}(v_1)$ is linear for any $v_1 \in R_1$ and $y_1(v_1, v_2) \neq 1/2$ for some $(v_1, v_2) \in R_1$. Due to anonymity on the diagonal $y_1(v, v) = 1/2$ for any v and due to monotonicity $y_1(v_1, v_2) \geq 1/2$ in each of the regions. Thus, $y_1(v_1, v_2) > 1/2$ is only possible in the low-right corner of the triangle R_1, which I denote by $A_1 \subset R_1$. But if it is the case, we can reduce $y_1(v_1, v_2)$ by a small $\varepsilon > 0$[12]. In R_1 the following holds: $\int\int_{A_1} \varepsilon f(v_2) dv_1 dv_2 < \int\int_{A_1} \varepsilon f(z_1) dv_1 dv_2$ and $\int\int_{A_1} \varepsilon f(v_1) dv_1 dv_2 > \int\int_{A_1} \varepsilon f(z_1) dv_1 dv_2$. The change in utility is:

$$
\begin{aligned}
\Delta U_1 &= -\int\int_{A_1} \varepsilon f(v_2) dv_1 dv_2 + \int\int_{A_1} \varepsilon f(v_1) dv_1 dv_2 > \\
&> -\int\int_{A_1} \varepsilon f(z_1) dv_1 dv_2 + \int\int_{A_1} \varepsilon f(z_1) dv_1 dv_2 > 0
\end{aligned}
$$

Hence, it is not possible to improve upon $y_1(v_1, v_2) = 1/2$ in the region R_1.

[12]Strictly speaking, we cannot always reduce allocation probability by ε everywhere, since it could prove to be lower than $1/2$ and violate monotonicity constraint. Thus, in the points where it occurs, we only reduce by $y_1(v_1, v_2) - 1/2$. Hence, the decrease is $\min\{\varepsilon, y_1(v_1, v_2) - 1/2\}$. But it matters only in the region with at least one dimension of order ε and hence it would be a second-order effect, which we can neglect.

Now consider any interior rectangle R_2. I claim that it is always optimal to give the object to the first agent. I use the similar logic as above. Assume that it is not true and there exists a subset $A_2 \subset R_2$: for any $(v_1, v_2) \in A_2$ we have $y_1(v_1, v_2) < 1$. Due to monotonicity, it could only be the upper-left corner. Now we increase probability of allocation to the first agent by ε in A_2[13]. In R_2 the following inequalities hold $\int\int_{A_2} \varepsilon f(v_2)dv_1dv_2 > \int\int_{A_2} \varepsilon f(z_1)dv_1dv_2$ and $\int\int_{A_1} \varepsilon f(v_1)dv_1dv_2 < \int\int_{A_1} \varepsilon f(z_1)dv_1dv_2$. Hence, the utility change is:

$$\begin{aligned}\Delta U_1 &= \int\int_{A_2} \varepsilon f(v_2)dv_1dv_2 - \int\int_{A_2} \varepsilon f(v_1)dv_1dv_2 > \\ &> \int\int_{A_2} \varepsilon f(z_1)dv_1dv_2 - \int\int_{A_2} \varepsilon f(z_1)dv_1dv_2 > 0\end{aligned}$$

So it is never optimal to put $y_1(v_1, v_2) < 1$ anywhere in R_2 , i.e. in the $FPA(C)$ the first agent always get the object in R_2.

While considering any interior or boundary triangle R_3 such that $g_{v^*}(v_1)$ is strictly increasing for any $v_1 \in R_3$, we notice that for any point $(v_1, v_2) \in R_3$ the following relation holds: $f(v_1) < f(v_2)$. Hence, from (1.13) it is optimal even pointwise in R_3 to make $y(v_1, v_2)$ as high as possible, i.e. $y(v_1, v_2) = 1$.

Boundary rectangle R_4 and boundary triangle R_5 such that $(v^*, v_2) \in R_4 \cap R_5$ for any $v_2 \le v^*$ are specific regions. The logic of a proof is a modified logic of the proof for regions R_1 and R_2. We start from R_4 and assume that for some $A_4 \subset R_4$ it is optimal to allocate the good to the favorite with a probability $y_1(v_1, v_2) < 1$. Again, it could only be the upper-low corner of the rectangle. We again increase probability of allocation in A_4 by ε[14]. The change in utility is:

$$\Delta U_1 = \int\int_{A_4 \cap \{v_1 \le v^*\}} \varepsilon f(v_2)dv_1dv_2 - \int\int_{A_4} \varepsilon f(v_1)dv_1dv_2$$

In this region $\int\int_{A_4 \cap \{v_1 \le v^*\}} \varepsilon f(v_2)dv_1dv_2 > \int\int_{A_4 \cap \{v_1 \le v^*\}} \varepsilon f(z_2)dv_1dv_2$ and $\int\int_{A_4} \varepsilon f(v_1)dv_1dv_2 < \int\int_{A_4 \cap \{v_1 \le v^*\}} \varepsilon f(z_2)dv_1dv_2$. Hence, $\Delta U > 0$ and $y_1(v_1, v_2) = 1$ must be optimal.

In R_5 we need to show that $y_1(v_1, v_2) = 1/2$ is optimal. By contrast, assume that there is $A_5 \subset R_5$ in the low-right corner where $y_1(v_1, v_2) > 1/2$. As before, reduce allocation probability by ε[15]. Since $\int\int_{A_5 \cap \{v_1 \le v^*\}} \varepsilon f(v_2)dv_1dv_2 < \int\int_{A_5 \cap \{v_1 \le v^*\}} \varepsilon f(z_2)dv_1dv_2$ and $\int\int_{A_5} \varepsilon f(v_1)dv_1dv_2 > \int\int_{A_5 \cap \{v_1 \le v^*\}} \varepsilon f(z_2)dv_1dv_2$, the utility change is:

$$\Delta U_1 = -\int\int_{A_5 \cap \{v_1 \le v^*\}} \varepsilon f(v_2)dv_1dv_2 + \int\int_{A_5} \varepsilon f(v_1)dv_1dv_2 > 0$$

[13] $\min\{\varepsilon, 1 - y_1(v_1, v_2)\}$
[14] $\min\{\varepsilon, 1 - y_1(v_1, v_2)\}$
[15] $\min\{\varepsilon, y_1(v_1, v_2) - 1/2\}$.

Thus, $y_1(v_1, v_2) = 1/2$ is optimal in R_5.

Since we worked with each region independently, this proof holds for any number and any combination of these regions. Since for any distribution function we can divide the subset of values below the diagonal $v_1 = v_2$ into regions of described values, we can apply the above logic to any distribution function and corresponding partition.

To complete the proof, we must show that the global monotonicity conditions are satisfied. Indeed, $y_1(v_1, v_2) \in \{1, 1/2\}$ for any $v_1 > v_2$. The regions where $y_1(v_1, v_2) = 1/2$ are only the triangles close to the diagonal. Thus, the proposed auction is indeed monotone. Transfers are chosen according to (1.6) taking into account that by Lemma 1.3 we have $h_i(0, \mathbf{v}_{-i}) = 0$.

Case 2. $v^* > \overline{v}$.

The idea here is to consider the characterization for the case $v^* = \overline{v}$ which follows from the previous case, and then to show for $v^* > \overline{v}$ that for all (v_1, v_2) such that $v_1 \in (\overline{v}, v^*]$ and $v_2 < v_1$ the optimal allocation is $y(v_1, v_2) = 1$, and for all (v_1, v_2) such that $(v_1, v_2) \in [0, \overline{v}] \times [0, \overline{v}]$ the allocation remains unchanged.

Indeed, suppose we consider $v^* > \overline{v}$. Then, similarly to the previous case we obtain the following.

$$\begin{aligned}
U_1(v^*) &= \int_0^{\overline{v}} (v^* y_1(v^*, v_2) + t_1(v^*, v_2)) f(v_2) dv_2 = \int_0^{\overline{v}} (\int_0^{v^*} y_1(v_1, v_2) dv_1) f(v_2) dv_2 = \\
&= \int_0^{\overline{v}} \int_{v_2}^{v^*} y_1(v_1, v_2) f(v_2) dv_1 dv_2 + \int_0^{\overline{v}} \int_{v_1}^{\overline{v}} y_1(v_1, v_2) f(v_2) dv_2 dv_1 = \\
&= \int_0^{\overline{v}} \int_{v_2}^{v^*} y_1(v_1, v_2) f(v_2) dv_1 dv_2 + \int_0^{\overline{v}} \int_{v_1}^{\overline{v}} y_2(v_2, v_1) f(v_2) dv_2 dv_1 = \\
&= \int_0^{\overline{v}} \int_{v_2}^{v^*} y_1(v_1, v_2) f(v_2) dv_1 dv_2 + \int_0^{\overline{v}} \int_{v_2}^{\overline{v}} y_2(v_1, v_2) f(v_1) dv_1 dv_2 = \\
&= \int_0^{\overline{v}} \int_{v_2}^{v^*} y_1(v_1, v_2) f(v_2) dv_1 dv_2 + \int_0^{\overline{v}} \int_{v_2}^{\overline{v}} [k(v_1, v_2) - y(v_1, v_2)] f(v_1) dv_1 dv = \\
&= U_1(\overline{v}) + \int_0^{\overline{v}} \int_{\overline{v}}^{v^*} y_1(v_1, v_2) f(v_2) dv_1 dv_2
\end{aligned}$$

Hence, it is optimal to have $y_1(v_1, v_2) = 1$ if $v_1 > \overline{v}$. At the same time, it does not violate monotonicity constraint. Thus, the optimal allocation for $v_1 \leq \overline{v}$ when $v^* > \overline{v}$ should coincide with the allocation for $v_1 \leq \overline{v}$, when $v^* = \overline{v}$. The auction described in the statement implements exactly this allocation[16]. Once again, transfers can be computed according to (1.6) taking into account that by Lemma 1.3 we have $h_i(0, \mathbf{v}_{-i}) = 0$. □

Proof of Proposition 1.2. Since distribution $F(v)$ is atomless, we have $F(v) < 1$ for any $v < \overline{v}$.

[16] The allocation for $(v_1, v_2) : \overline{v} < v_1 \leq v_2$ does not affect the utility of the favorite. For definiteness sake, in the statement we have specified $y(v_1, v_2) = 1$ for $\overline{v} < v_1 \leq v_2$.

Thus, $1 = G_{v^*}(v^*) > F(v^*)$ for any $v^* < \overline{v}$. Since $G_{v^*}(v)$ and $F(v)$ are different at $v = v^*$, it means that $v = v^*$ belongs to a subset where $G_{v^*}(v)$ is linear, i.e. there is a pooling interval $(\widehat{v}, \widehat{\widehat{v}})$ such that $v^* \in (\widehat{v}, \widehat{\widehat{v}})$. $\widehat{v}$ is a point, tangent line from which goes directly to $(v^*, 1)$. By construction $g_{v^*}(v) = const$ for $v > v^*$ and hence $v > v^*$ is a pooling region. This implies that all values $v > \widehat{v}$ must be pooled.

To show monotonicity of $\widehat{v}$ as a function of v^*, suppose that it is not true, i.e. there exist v_1^* and v_2^* such that $v_1^* < v_2^*$ and $\widehat{v}(v_1^*) > \widehat{v}(v_2^*)$. By definition of G_{v^*} the following holds: $G_{v_2^*}(\widehat{v}(v_1^*)) \geq F(\widehat{v}(v_1^*))$. Since $G_{v_1^*}(v_1^*) - G_{v_1^*}(\widehat{v}(v_1^*)) = 1 - F(\widehat{v}(v_1^*)) > G_{v_2^*}(v_1^*) - G_{v_2^*}(\widehat{v}(v_1^*))$ we must have $g_{v_1^*}(\widehat{v}(v_1^*)) > g_{v_2^*}(\widehat{v}(v_1^*))$. Then

$$
\begin{aligned}
G_{v_1^*}(\widehat{v}(v_2^*)) &= G_{v_1^*}(\widehat{v}(v_1^*)) - (\widehat{v}(v_1^*) - \widehat{v}(v_2^*))g_{v_1^*}(\widehat{v}(v_1^*)) \\
&= F(\widehat{v}(v_1^*)) - (\widehat{v}(v_1^*) - \widehat{v}(v_2^*))g_{v_1^*}(\widehat{v}(v_1^*)) \\
&< F(\widehat{v}(v_1^*)) - (\widehat{v}(v_1^*) - \widehat{v}(v_2^*))g_{v_2^*}(\widehat{v}(v_1^*)) = F(\widehat{v}(v_2^*))
\end{aligned}
$$

However, $G_{v_1^*}(\widehat{v}(v_2^*)) < F(\widehat{v}(v_2^*))$ is impossible by construction of $G_{v_1^*}$. Thus, $\widehat{v}(v^*)$ has to be monotone. □

Proof of Proposition 1.3. The result follows from the proof of Proposition 1.1. Suppose $v^* < \overline{v}$. If the favorite's value changes, the corresponding change of the $FPA(\{c_A, c_{DIC}, c_{NT}\})$ is related to the change of the function $G_{v^*}(v)$. The only change of this function happens on the subset $[\widehat{v}(v^*), \overline{v}]$, which is a pooling region. For all favorite's values above the maximal possible value of his opponent, the function $G_{v^*}(v)$ is the same function for all v^*, which brings the same $FPA(\{c_A, c_{DIC}, c_{NT}\})$ for all $v^* > \overline{v}$. □

Proof of Theorem 1.3. Step 1:

First, consider a value v_i of a bidder i such that $v_i < \max_{j\neq i}\{v_j\}$. Then, due to DA, anonymity and monotonicity condition (1.5), the bidder i should receive the object with zero probability and $y_i(v_1, ..., \underset{i}{q}, ...v_n) = 0$ for all $q \leq v_i$. Indeed, to show this, suppose that $y_i(v_1, ..., v_i, ..., v_n) = 1$ for some $v_i < \max_{j\neq i}\{v_j\}$. Then, monotonicity implies that $y_i(v_1, ..., \max_{j\neq i}\{v_j\}, ...v_n) = 1$. However, due to anonymity the bidder k who has the value $v_k = \max_{j\neq i}\{v_j\}$ should also have probability of assigning the good equal to one. Thus, we obtain that $y_i(v_1, ..., \max_{j\neq i}\{v_j\}, ...v_n) = y_k(v_1, ..., \max_{j\neq i}\{v_j\}, ...v_n) = 1$, which contradicts feasibility. Thus, all bidders whose value is not the highest one should receive the good with zero probability, namely if $v_i < \max_{j\neq i}\{v_j\}$, then $y_i(\mathbf{v}) = 0$.

Step 2:

From Step 1, it follows that for any realization of values there could be only two possible cases: 1) the bidder with the highest value obtains the object for sure, 2) nobody gets the object.

Monotonicity constraint (1.5) implies that if for some vector of values bidder i receives the good, he should also receive the good when he has a higher value keeping values of his opponents fixed. Thus, for any $FA(C)$ there is a cutoff r_i for each bidder, which can depend on other bidders' values, such that the bidder obtains the object with a probability of 1 if and only if his value is 1) the greatest among values of other bidders and 2) greater or equal than the cutoff r_i. Thus, $y_i(\mathbf{v}) = 1$ if and only if $v_i > \max_{j \neq i}(v_j, r_i)$, otherwise $y_i(\mathbf{v}) = 0$.

Step 3:

Now we need to understand how these values $\{r_i\}_{i=1}^n$, or essentially reserve values, are constructed. First, notice that for each bidder i his reserve value r_i can depend on his opponents' bids. Hence, r_i can depend on $v_1, ..., v_{i-1}, v_{i+1}, ..., v_n$. By anonymity, the allocation probability for a bidder i should not be affected by any permutation of other players' bids. Hence, r_i has to be a symmetric function of $n-1$ variables $r_i : \mathbb{R}^{n-1} \to \mathbb{R}$. Then, again due to anonymity, since the allocation rule must be symmetric among bidders, then for any bidders i and j and for any $\mathbf{x} \in \mathbb{R}^{n-1}$ the function $r_i(\mathbf{x})$ and $r_j(\mathbf{x})$ have to be equal, $r_i(\mathbf{x}) = r_j(\mathbf{x})$. Hence, the reserve value function should be common for all bidders: $r_1(\mathbf{x}) = ... = r_n(\mathbf{x}) = r(\mathbf{x})$.

Step 4:

Take any bidder $i : v_i < \max_{j \neq i}\{v_j\}$. Then, from Step 1 we have $y_i(v_1, ..., \underset{i}{q}, ...v_n) = 0$ for all $q \leq v_i$. Hence, from (1.6) we have

$$t_i(\mathbf{v}) = U_i(\mathbf{v}) = h_i(\underline{v}_i, \mathbf{v}_{-i})$$

Since transfers have to be non-positive, it follows that $\forall \mathbf{v}_{-i}$: $h_i(\underline{v}_i, \mathbf{v}_{-i}) \leq 0$. However, simultaneously to satisfy DIC, utility of bidder i has to be at least non-negative, otherwise he could refrain himself from participation. Thus, it must be the case that $h_i(\underline{v}_i, \mathbf{v}_{-i}) \geq 0$. Combining the two inequalities we obtain $h_i(\underline{v}_i, \mathbf{v}_{-i}) = 0$. It means that transfers are uniquely defined when the allocation is chosen. Thus, plugging the obtained allocation rule and $h_i(\underline{v}_i, \mathbf{v}_{-i}) = 0$ to (1.6) we get $t_i(\mathbf{v}) = -\max_{j \neq i}(v_j, r(\mathbf{v}_{-i}))$ if and only if $v_i > \max_{j \neq i}(v_j, r(\mathbf{v}_{-i}))$, and $t_i(\mathbf{v}) = 0$ otherwise. Since we have $y_i(\mathbf{v}) = 1$ if and only if $v_i > \max_{j \neq i}(v_j, r(\mathbf{v}_{-i}))$, the statement follows. □

Proof of Proposition 1.4. From the proof of Theorem 3, the utility of any bidder under the full set of restrictions $C = \{c_A, c_{DIC}, c_{NT}, c_{DA}\}$ must be $U_i(\mathbf{v}) = \int_{\underline{v}_i}^{v_i} y_i(v_1, ..., \underset{i}{q}, ...v_n)dq$, where $y_i(\mathbf{v}) = 1$ if and only if $v_i > \max_{j \neq i}(v_j, r_i(\mathbf{v}_{-i}))$, otherwise $y_i(\mathbf{v}) = 0$. The choice of a reserve value function completely determines the auction format. Hence, the utility of each bidder including the favorite can be written as follows:

$$\begin{aligned} U_i(\mathbf{v}) &= \int_{\max_{j \neq i}(v_j, r_i(\mathbf{v}_{-i}))}^{v_i} y_i(v_1, ..., \underset{i}{q}, ...v_n)dq = \\ &= \max\{0, v_i - \max_{j \neq i}(v_j, r_i(\mathbf{v}_{-i}))\} \end{aligned}$$

Hence, making positive reserve prices can only reduce the utility of each bidder including the favorite. Thus, it is optimal to put zero reserve price, so $r_i(\mathbf{x}) = 0 \ \forall \mathbf{x} \in \mathbb{R}^{n-1}$. □

1.B Appendix B

Assumption 1.1 (symmetric bidders)**.** *$\forall i, j \in N : V_i = V_j = V$ and $F_i(v) = F_j(v) = F(v)$*

Proposition 1.6 (ex-post equivalence of unrestricted and anonymity-restricted favoritism)**.**

If bidders are symmetric and $C = \{c_A\}$ then there exists $FA(C) = (\widehat{B}, \widehat{\mathbf{y}}, \widehat{\mathbf{p}})$ such that it has an equilibrium $\widehat{\psi}$, in which $\widehat{B} = [\underline{v}, \overline{v}]$, $\widehat{M} = M^{FPA}$, $\widehat{\mathbf{p}}(a, \widehat{\mathbf{b}}(\mathbf{v})) = \mathbf{p}^{FPA}(a, \mathbf{b}^{FPA}(\mathbf{v}))$, $\widehat{\mathbf{y}}(\widehat{\mathbf{b}}(\mathbf{v})) = \mathbf{y}^{FPA}(\mathbf{b}^{FPA}(\mathbf{v}))$, where $\widehat{\mathbf{b}}(\mathbf{v})$ and $\mathbf{b}^{FPA}(\mathbf{v})$ stand for the biddings in the equilibrium $\widehat{\psi}$ of $FA(C)$ and ψ^* of FPA respectively.

Proof of Proposition 1.6. I prove the theorem by directly constructing the equivalent anonymous auction $FA(C)$. Consider the set of admissible bids equal to the set of possible values, $B = [\underline{v}, \overline{v}]$. Denote $\widehat{v}$ as the smallest value such that $v^* = \widehat{v} - \frac{1-F(\widehat{v})}{f(\widehat{v})}$. If no solution to this equation exists, assume $\widehat{v} = \overline{v}$. The allocation rule is such that the bidder with the highest bid wins, i.e. $\widehat{\mathbf{y}}(b_1, ..., b_i, ..., b_m) = (0, ..., \underset{i}{1}, ..., 0)$ if $b_i > b_j$ for any $j \neq i$. If $k \geq 2$ bidders make exactly the same bids, there is a symmetric lottery between them with $1/k$ being a probability of securing the good for each of them. Transfers $\widehat{\mathbf{p}}(\mathbf{b})$ are such that if there is only one bid on the interval $[\underline{v}, \widehat{v}]$, then this bidder pays nothing, although if there are two or more bidders who make bids from this interval, all of them should pay $\widehat{v}$. If the winning bid is greater than $\widehat{v}$, the payment is the maximum between the second highest bid and $\widehat{v}$. Subsequently, there is an equilibrium $\widehat{\psi}$, in which the favorite bids v^*, all bidders with values smaller than $\widehat{v}$ do not participate in the bidding and all bidders with values greater than $\widehat{v}$ participate and bid their true values. This equilibrium outcome is always the same as in the FPA. □

Chapter 2

Non-discriminatory Strategyproof Optimal Auction

2.1 Introduction

In a seminal paper, Myerson (1981) characterizes the revenue maximizing mechanisms for auctioning a single indivisible object to buyers who have independent and private valuations (IPV) of the object. Such a mechanism allocates the object to an agent with the highest "virtual valuation", which depends on this agent's actual valuation and her valuation distribution. If bidders are ex-ante symmetric, i.e., their valuations are drawn from the same distribution, revenue maximization can be achieved by implementing a second-price auction with a common reserve price. Second-price auction is an anonymous (symmetric) and dominant strategy incentive compatible (DIC), or strategyproof, mechanism. Anonymity means that the allocations and transfers depend only on bidders' bids but not on bidders' identities (names, races, nationalities, and etc.). Dominant strategy incentive compatibility However, when bidders are ex-ante asymmetric, then Myerson's optimal auction is not symmetric anymore. In reality, agents are often ex-ante asymmetric, for example, foreign firms and domestic firms can be characterized by different distributions. Meanwhile, mechanism designers can be restricted to use only symmetric mechanisms to avoid descrimination. Hence, the following natural question arises. Namely, what would be the optimal mechanism under the restriction of anonymity. The surprising answer to this question is given by Deb and Pai (2017). They demonstrate that the optimal mechanism is ex-ante implementable in a symmetric way. Precisely, there is a symmetric auction that has an equilibrium with the same ex-post allocation rule and interim utilities as that of the asymmetric optimal auction. However, this equivalence holds only in the sense of bayesian incentive compatibility (BIC). It means that even though the initial mechanism is dominant strategy incentive compatible (DIC), its implementation is only bayesian incentive compatible. Hence, many important properties of the mechanism are lost.[1] Even more crucial issue is that BIC implementation does not exclude multiplicity of equilibria. For example, the mechanisms from Deb and Pai (2017) generically have many equilibria and some of them could be symmetric equilibria. However, only one particular equilibrium is chosen. Hence, the results obtained by Deb and Pai (2017) crucually rely on the asumption that the designer can pick the preferred equilibrium. In this chapter, we are interested in finding the optimal mechanism that preserves anonymity and DIC together. This mechanism is robust and has a unique equilibrium in undominated strategies.

There is a literature on BIC-DIC equivalence. Manelli and Vincent (2010), Gershkov et al. (2013) show that in IPV models any Bayes-Nash equilibrium outcome can also be achived in expectation in some mechanism that implements dominant strategies. However, in this chapter, the anonymity restriction breaks this equivalence.

Azrieli and Jain (2018) generalize Deb and Pai (2017) from auction setting to a symmetric implementation of a general social choice function. However, Azrieli and Jain (2018) obtain this generalization by using abstract message spaces. In particular, they allow agents to report their names in messages. Then the designer can make the mechanism depend on the reported names. On need only to care that there exists and equilibrium where every bidder reports his name truthfully. At the same time, there could be many equilibria where agents strategically misreport their names.

The methedology used in this chapter is closely related to the methodology used in Chapter 1that considers a question of favoritism in auctions. In Chapter 1, the designer is interested in maximizing the utility of her favored bidder and is restricted by the anonymity and DIC constraints. In this chapter we find the auction which maximizes the revenue of the seller under the same two constraints. We have shown in Chapter 1 that anonymity and DIC constraints imply that anonymity of original auction transfers to anonymity of the corresponding direct auction. We also employ this characterization to construct the optimal anonymous DIC auction.

In the next section we present our model and show our main result that the optimal anonymous DIC mechanism is a second-price auction with specially constructed reserve prices. Each bidder's reserve price depends on the bids and the value distributions of her competitors. However, the constructions of the reserve prices are symmetric for the bidders and satisfy the anonymity restriction.

[1] Some of them are the following. Bidders should know each others' value distributions. It is possible that a bidder has to pay eventually without obtaining the object. Multiple equilibria may also arise.

2.2 Main Model

There is one indivisible object that can be sold to one of n buyers (bidders). Buyer i has a privately known valuation of the object, v_i, which is drawn indepently from a continuously differentiable destibution F_i on $[0, \overline{v}]$, $\overline{v} \in (0, +\infty]$. Each bidder i submits a bid $b_i \in \mathbb{R}_+$, and all bids are submitted simultaneously and independently of each other. The objective of the designer is to maximize her expected revenue. A mechanism $M := (\mathbf{y}, \mathbf{t})$ is a collection of an allocation rule $\mathbf{y} :\mathbb{R}^n_+ \to [0,1]^n$, $\mathbf{y}(\mathbf{b}) = (y_1(\mathbf{b}), ..., y_n(\mathbf{b}))$ and a transfer rule $\mathbf{t} :\mathbb{R}^n_+ \to \mathbb{R}^n$, $\mathbf{t}(\mathbf{b}) = (t_1(\mathbf{b}), ..., t_n(\mathbf{b}))$. For each bidder, $y_i(\mathbf{b})$ is the probability of bidder i getting the object and $t_i(\mathbf{b})$ is the transfer to agent i, where $\mathbf{b} := (b_1, ...b_n)$ is a vector of submitted bids. Bidder i's utility is

$$U_i = \begin{cases} v_i y(b) + t_i & \text{if } i \text{ obtains the object,} \\ t_i & \text{otherwise.} \end{cases} \tag{2.1}$$

The difference from the standard literature on revenue maximization is that we require the mechanism to be anonymous and implements dominant strategies.

Let $\pi : \{1, ..., n\} \to \{1, ..., n\}$ be a permutation. Denote Θ to be the set of all permutations of N elements. The expression $\pi(i) = j$ means that the element in i-th position is permuted to j-th position. We denote π^{-1} as the inverse of the permutation of π. To simplify notations, for any vector $\mathbf{x} = (x_1, ..., x_n)$ we use $\pi(\mathbf{x})$ to denote the vector obtained after permutation π is applied. Thus, $\pi(\mathbf{x}) := (x_{\pi(1)}, ..., x_{\pi(n)})$.

Definition 2.1. *A mechanism M is anonymous if under a permutation of bids, its allocations and payments are permuted. Formally, for any permutation π,*

$$\mathbf{y}(\pi(\mathbf{b})) = \pi(\mathbf{y}(\mathbf{b})) \text{ and } \mathbf{t}(\pi(\mathbf{b})) = \pi(\mathbf{t}(\mathbf{b})).$$

We now consider dominant strategy implementation. Although dominant strategy implementation is quite robust in the sense that the behavior of an agent does not depend on the strategies of her opponents, there could still be a problem of multiple equilibria in a mechanism, because each agent can have more than one dominant strategy. Lemmas 1.1 and 1.2 and Theorem 1.1 from Chapter 1are crucial for our results.

Together they imply that if we apply the revelation principle and consider only mechanisms in which bidders report their valuations directly, instead of considering mechanisms that implement dominant strategies, we can directly consider anonymous dominant strategy incentive compatible

(DIC) direct mechanisms.

The following standard lemma due to Maskin and Laffont (1979) characterizes all DIC direct mechanisms.

Lemma 2.1 (Maskin and Laffont (1979)). *A direct mechanism is dominant strategy incentive compatible if and only if for each agent:*

1. $y_i(\mathbf{v})$ *is nondecreasing in* v_i *for all* $\mathbf{v}_{-i}$.
2. *There exist functions* $\{C_i(\mathbf{v}_{-i})\}$ *such that*

$$v_i y_i(\mathbf{v}) + t_i(\mathbf{v}) = C_i(\mathbf{v}_{-i}) + \int_0^{v_i} y_i(v_1, ..., \underset{i}{q}, ...v_N)dq. \tag{2.2}$$

Thus, the designer's problem is:

$$\max_{\{y_i(.)\},\{t_i(.)\}} E[R] = E[-\sum_i t_i(\mathbf{v})] \tag{2.3}$$

$$\text{subject to:} \tag{2.4}$$

$$y_i(v_1, ..., v_N) \quad = \quad y_{\pi(i)}(v_{\pi(1)}, ..., v_{\pi(N)}) \tag{2.5}$$

$$t_i(v_1, ..., v_N) \quad = \quad t_{\pi(i)}(v_{\pi(1)}, ..., v_{\pi(N)}) \tag{2.6}$$

$$0 \leq y_i(\mathbf{v}) \leq 1, \;\; \sum_i y_i(\mathbf{v}) \leq 1 \tag{2.7}$$

$$y_i(\mathbf{v}) \text{ is nondecreasing in } v_i \text{ for all } \mathbf{v}_{-i} \tag{2.8}$$

$$v_i y_i(\mathbf{v}) + t_i(\mathbf{v}) = C_i(\mathbf{v}_{-i}) + \int_0^{v_i} y_i(v_1, ..., \underset{i}{q}, ...v_N)dq \tag{2.9}$$

We solve this problem under the following assumption on the monotonicity of cross hazard rates.[2]

Assumption 2.1. *For any* i, j *the function* $h_{i,j}(\cdot) := \frac{1-F_i(\cdot)}{f_j(\cdot)}$ *is decreasing.*

[2]This asumption is not new in the mechanism design literature. See, for example, Krähmer and Strausz (2015)

This assumption is a generalization of the standard decreasing inverse harzard rate assumption made in the mechanism design literature to the case in which the anonymity restriction is present. Then we can obtain the following result.

Proposition 2.1. *The optimal anonymous DIC auction is a second-price auction with different reserve prices for different bidders. Each bidder's reserve price depends on her opponents' bids and valuation distributions, and it is determined in the following equation:*

$$r_k = \frac{\sum\limits_{\pi\in\Theta}(1-F_{\pi^{-1}(k)}(r_k))\Pi_{j\neq\pi^{-1}(k)}f_j(v_{\pi(j)})}{\sum\limits_{\pi\in\Theta}f_{\pi^{-1}(k)}(r_k)\Pi_{j\neq\pi^{-1}(k)}f_j(v_{\pi(j)})} \tag{2.10}$$

The intuition behind this result is as follows. As we know from Theorem 1.1, the auction must treat agents' reported valuations in a symmetric way. Hence, any feasible auction must allocate the object to a bidder with the highest bid or pool some bids. In the optimal auction, pooling may arise only if some ironing procedure is necessary, as in Myerson (1981). By Asumption 1 we exclude those cases where ironing is necessary. Hence, the object should be allocated to a bidder with the highest bid. Then, by the DIC restriction, the winner should pay the second highest bid. The most interesting aspect of the mechanism is the optimal reserve prices. The reserve price for each bidder depends both on the actual valuations and the valuation distributions of all her opponents. However, the constructions of all reserve prices are symmetric.

Proof of Proposition 2.1.

$$\begin{aligned}
E[R] &= E[-\sum_i t_i(\mathbf{v})] \\
&= E[-\sum_i (U_i(\mathbf{v}) - v_i y_i(\mathbf{v}))] \\
&= E[-\sum_i (C_i(\mathbf{v} v_i y_i(\mathbf{v}))] \\
&= \int_{\mathbf{v}}\cdots\int(-\sum_i (C_i(\mathbf{v}_{-i}) + \int_0^{v_i} y_i(v_1,...,\underset{i}{q},...v_n)dq - v_i y_i(\mathbf{v}))f_1(v_1)...f_n(v_n)dv_1...dv_n.
\end{aligned}$$

It is standard that $\{C_i(\mathbf{v}_{-i})\}_{i=1,\dots,n}$ are set to be as low as possible to satisfy the individual rationality constraints. Due to Lemma 2.1, $C_i(\mathbf{v}_{-i}) \equiv 0$.

Using standard technique of integration by parts we obtain the following representation of revenue:

$$\int_{\mathbf{v}}\cdots\int\sum_i y_i(\mathbf{v})(v_i - \frac{1-F_i(v_i)}{f_i(v_i)})f_1(v_1)...f_n(v_n)dv_1...dv_n \tag{2.11}$$

Now we must take into account the anonymity constraint. Denote Θ as the set of all permutation functions. There are $n!$ elements in this set. Then, for any allocation rule, if we know the allocation at some vector $\mathbf{v}=(v_1, v_2..., v_n)$, by anonymity we also know the allocation at each vector $\pi(\mathbf{v}) = (v_{\pi(1)}, v_{\pi(2)}, ..., v_{\pi(n)})$, which is a permutation of $\mathbf{v}$.

In order to find the optimal allocation, we do pointwise maximization and then check that all monotonicity constraints hold. Suppose, we take some allocation $\mathbf{y}(\mathbf{v}) := (y_1(\mathbf{v}), ..., y_n(\mathbf{v}))$. Consider some bidders l, k and assume without loss of generality that $v_k > v_l$. We will first fix the allocation for other bidders and distribute the rest allocation probability between bidders l and k. Denote $\overline{q}(\mathbf{v}) := 1 - \sum_{j \neq k.l} y_j(\mathbf{v})$ and $q(\mathbf{v}) := y_l(\mathbf{v}) + y_k(\mathbf{v})$. By definition $q(\mathbf{v}) \leq \overline{q}(\mathbf{v})$.

Now in (2.11) we can consider only the terms associated with y_l and y_k, since the other probabilities are fixed. We substitute y_l by $q - y_k$ in the maximization function (2.11) and then have

$$\begin{aligned}
& \int_{\mathbf{v}} \cdots \int [y_k(\mathbf{v})(v_k - \frac{1-F_k(v_k)}{f_k(v_k)}) + \\
& + (q(\mathbf{v}) - y_k(\mathbf{v}))(v_l - \frac{1-F_l(v_l)}{f_l(v_l)})] f_1(v_1)...f_n(v_n) dv_1...dv_n \\
= \; & \int_{\mathbf{v}} \cdots \int [y_k(\mathbf{v})((v_k - \frac{1-F_k(v_k)}{f_k(v_k)}) - (v_l - \frac{1-F_l(v_l)}{f_l(v_l)})) + \\
& + q(\mathbf{v})(v_l - \frac{1-F_l(v_l)}{f_l(v_l)})] f_1(v_1)...f_n(v_n) dv_1...dv_n.
\end{aligned}$$

To maximize properly the above term under anonymity constraint, we need to maximize the integral of the following term, which considers together all points which are permutations of each other:

$$\begin{aligned}
& q(\mathbf{v}) \sum_{\pi \in \Theta} [v_l - \frac{1-F_{\pi^{-1}(l)}(v_l)}{f_{\pi^{-1}(l)}(v_l)}] \Pi_j f_j(v_{\pi(j)}) \\
& + y_k(\mathbf{v}) (\sum_{\pi \in \Theta} [v_k - \frac{1-F_{\pi^{-1}(k)}(v_k)}{f_{\pi^{-1}(k)}(v_k)}] \Pi_j f_j(v_{\pi(j)}) - \sum_{\pi \in \Theta} [v_l - \frac{1-F_{\pi^{-1}(l)}(v_l)}{f_{\pi^{-1}(l)}(v_l)}]) \Pi_j f_j(v_{\pi(j)}.
\end{aligned} \tag{2.12}$$

We omit for a while the term associated with $q(\mathbf{v})$ and focus only on the term associated with $y_k(\mathbf{v})$. Now, consider a permutation $S \in \Theta$ such that it switches the positions of k and l, without affecting the other positions. Thus, $S(j) = j$ for all $j \neq k, l$, $S(k) = S^{-1}(k) = l$, and $S(l) = S^{-1}(l) = k$. Since in the last expression we have summands with respect to all possible permutations, the value of the expression will not change if we replace $\pi(\cdot)$ by $S\pi(.) := S(\pi(.))$ in the second summand. We also notice that $S\pi^{-1}(.)$, the inverse permutation of $S\pi(.)$, is equal to

$\pi^{-1}(S(.))$. Thus, we have

$$
\begin{aligned}
& y_k(\sum_{\pi\in\Theta}[v_k - \frac{1-F_{\pi^{-1}(k)}(v_k)}{f_{\pi^{-1}(k)}(v_k)}]\Pi_j f_j(v_{\pi(j)}) - \sum_{\pi\in\Theta}[v_l - \frac{1-F_{\pi^{-1}(S(l))}(v_l)}{f_{\pi^{-1}(S(l))}(v_l)}]\Pi_j f_j(v_{S(\pi(j))})) \\
= \; & y_k(\sum_{\pi\in\Theta}[v_k - v_l]\Pi_j f_j(v_{\pi(j)}) + \sum_{\pi\in\Theta}[\frac{1-F_{\pi^{-1}(S(l))}(v_l)}{f_{\pi^{-1}(S(l))}(v_l)}\Pi_j f_j(v_{S(\pi(j))})) - \\
& -\frac{1-F_{\pi^{-1}(k)}(v_k)}{f_{\pi^{-1}(k)}(v_k)}\Pi_j f_j(v_{\pi(j)})] \\
= \; & y_k(\sum_{\pi\in\Theta}[v_k - v_l]\Pi_j f_j(v_{\pi(j)}) + \sum_{\pi\in\Theta}[\frac{1-F_{\pi^{-1}(k)}(v_l)}{f_{\pi^{-1}(k)}(v_l)}\Pi_j f_j(v_{S(\pi(j))})) - \\
& -\frac{1-F_{\pi^{-1}(k)}(v_k)}{f_{\pi^{-1}(k)}(v_k)}\Pi_j f_j(v_{\pi(j)})] \\
= \; & y_k(\sum_{\pi\in\Theta}[v_k - v_l]\Pi_j f_j(v_{\pi(j)}) + \big((1-F_{\pi^{-1}(k)}(v_l))f_{\pi^{-1}(l))}(v_k) \\
& -(1-F_{\pi^{-1}(k)}(v_k))f_{\pi^{-1}(l)}(v_l)\big)\Pi_{j\neq\pi^{-1}(k),\pi^{-1}(l)} f_j(v_{\pi(j)}) \\
= \; & y_k(\sum_{\pi\in\Theta}[v_k - v_l]\Pi_j f_j(v_{\pi(j)}) + \\
& +(\frac{1-F_{\pi^{-1}(k)}(v_l)}{f_{\pi^{-1}(l)}(v_l)} - \frac{1-F_{\pi^{-1}(k)}(v_k)}{f_{\pi^{-1}(l)}(v_k)})f_{\pi^{-1}(l)}(v_l)f_{\pi^{-1}(l)}(v_k)\Pi_{j\neq\pi^{-1}(k),\pi^{-1}(l)} f_j(v_{\pi(j)})
\end{aligned}
$$

Since by assumption $v_k > v_l$ and $\frac{1-F_i(v)}{f_j(v)}$ is decreasing, it is optimal to set $y_k = q$ and $y_l = 0$. Now we need to find the optimal value for q. Notice that (2.12) has the following form:

$$
q(\mathbf{v})\sum_{\pi\in\Theta}[v_k - \frac{1-F_{\pi_{-1}(k)}(v_k)}{f_{\pi_{-1}(k)}(v_k)}]\Pi_j f_j(v_{\pi(j)}) \tag{2.13}
$$

Hence, $q(\mathbf{v})$ should be equal to $\overline{q}(\mathbf{v})$ if $\sum_{\pi\in\Theta}[v_k - \frac{1-F_{\pi^{-1}(k)}(v_k)}{f_{\pi^{-1}(k)}(v_k)}]\Pi_j f_j(v_{\pi(j)}) \geq 0$ and zero otherwise. Keeping $\{v_j\}_{j\neq k}$ constant, consider $\Phi(v_k) \equiv \sum_{\pi\in\Theta}[v_k - \frac{1-F_{\pi^{-1}(k)}(v_k)}{f_{\pi^{-1}(k)}(v_k)}]\Pi_j f_j(v_{\pi(j)})$. We will show that this function is monotonely increasing and there exists a unique point r_k such that $\Phi_k(r_k) = 0$. If $\Phi_k(r_k) = 0$ then

$$
\sum_{\pi\in\Theta}[r_k - \frac{1-F_{\pi^{-1}(k)}(r_k)}{f_{\pi^{-1}(k)}(r_k)}]\Pi_j f_j(v_{\pi(j)}) = 0 \tag{2.14}
$$

Hence, r_k satisfies the following equation:

$$
r_k = \frac{\sum_{\pi\in\Theta}(1-F_{\pi^{-1}(k)}(r_k))\Pi_{j\neq\pi^{-1}(k)} f_j(v_{\pi(j)})}{\sum_{\pi\in\Theta} f_{\pi^{-1}(k)}(r_k)\Pi_{j\neq\pi^{-1}(k)} f_j(v_{\pi(j)})} \tag{2.15}
$$

If the derivative $\Phi_k^{'}$ of the function Φ_k has the same sign the uniqueness of r_k is guaranteed.

Since all inverse hazard rates are decreasing by assumption, the derivative of Φ_k is always positive:

$$\Phi_k'(r_k) = \sum_{\pi\in\Theta}[1 - \left(\frac{1 - F_{\pi^{-1}(k)}(r_k)}{f_{\pi^{-1}(k)}(r_k)}\right)']\Pi_j f_j(v_{\pi(j)}) > 0 \tag{2.16}$$

Thus, $q(\mathbf{v}) = \overline{q}(\mathbf{v})$ if $v_k > r_k$ and $q(\mathbf{v}) = \mathbf{0}$ if $v_k \leq r_k$. Essentially, when choosing the allocation between bidders k, l, we compare their valuations and allocate the good to the bidder with a higher valuation, provided that this valuation is also higher than the reserve price determined from (2.15). Following this procedure we choose the buyer with the highest valuation and he receives the good if his valuation is higher than his reserve price. The price that she needs to pay is then the maximum between the second highest valuation and the reserve price for this buyer. All monotonicity constraints are trivially satisfied in such an allocation. The only constraint left to be varified is that the reserve prices are symmetric for all bidders. To show this, consider two bidders m, n and show that r_m and r_n depend on the bids of their respective opponents in the same way. To show this, we use the same trick as above. In the summation, we replace π with $S\pi$, where S is such permutation that switches the positions of m and n. Consider bidder m when bidder n has a valuation $v_n = \widehat{v}$. Then the following equalities hold:

$$\begin{aligned}
r_m &= \frac{\sum_{\pi\in\Theta}(1 - F_{\pi^{-1}(m)}(r_m))\Pi_{j\neq\pi^{-1}(m)} f_j(v_{\pi(j)})}{\sum_{\pi\in\Theta} f_{\pi^{-1}(m)}(r_m)\Pi_{j\neq\pi^{-1}(m)} f_j(v_{\pi(j)})} = \\
&= \frac{\sum_{\pi\in\Theta}(1 - F_{\pi^{-1}(S(m))}(r_m))\Pi_{j\neq\pi^{-1}(S(m))} f_j(v_{S(\pi(j))})}{\sum_{\pi\in\Theta} f_{\pi^{-1}(S(m))}(r_m)\Pi_{j\neq\pi^{-1}(S(m))} f_j(v_{S(\pi(j))})} = \\
&= \frac{\sum_{\pi\in\Theta}(1 - F_{\pi^{-1}(n)}(r_m))\Pi_{j\neq\pi^{-1}(n)} f_j(v_{S(\pi(j))})}{\sum_{\pi\in\Theta} f_{\pi^{-1}(n)}(r_m)\Pi_{j\neq\pi^{-1}(n)} f_j(v_{S(\pi(j))})} = \\
&= \frac{\sum_{\pi\in\Theta}(1 - F_{\pi^{-1}(n)}(r_m)) f_{\pi^{-1}(m)}(v_{S(\pi(\pi^{-1}(m)))})\Pi_{j\neq\pi^{-1}(n),\pi^{-1}(m)} f_j(v_{S(\pi(j))})}{\sum_{\pi\in\Theta} f_{\pi^{-1}(n)}(r_m) f_{\pi^{-1}(m)}(v_{S(\pi(\pi^{-1}(m)))})\Pi_{j\neq\pi^{-1}(n),\pi^{-1}(m)} f_j(v_{S(\pi(j))})} = \\
&= \frac{\sum_{\pi\in\Theta}(1 - F_{\pi^{-1}(n)}(r_m)) f_{\pi^{-1}(m)}(v_n)\Pi_{j\neq\pi^{-1}(n),\pi^{-1}(m)} f_j(v_{\pi(j)})}{\sum_{\pi\in\Theta} f_{\pi^{-1}(n)}(r_m) f_{\pi^{-1}(m)}(v_n)\Pi_{j\neq\pi^{-1}(n),\pi^{-1}(m)} f_j(v_{\pi(j)})} = \\
&= \frac{\sum_{\pi\in\Theta}(1 - F_{\pi^{-1}(n)}(r_m)) f_{\pi^{-1}(m)}(\widehat{v})\Pi_{j\neq\pi^{-1}(n),\pi^{-1}(m)} f_j(v_{\pi(j)})}{\sum_{\pi\in\Theta} f_{\pi^{-1}(n)}(r_m) f_{\pi^{-1}(m)}(\widehat{v})\Pi_{j\neq\pi^{-1}(n),\pi^{-1}(m)} f_j(v_{\pi(j)})}
\end{aligned}$$

The last expression represents exactly the reserve price for a bidder n, if bidder m has a valuation $\widehat{v}$. Hence, the constructed mechanism is indeed anonymous. □

Chapter 3

Head Starts and Doomed Losers: Contest via Search

3.1 Introduction

> "[U]nfortunately, for every Apple out there, there are a thousand other companies . . . like Woolworth, Montgomery Ward, Borders Books, Blockbuster Video, American Motors and Pan Am Airlines, that once 'ruled the roost' of their respective industries, to only get knocked off by more innovative competitors and come crashing down." (*Forbes*, January 8, 2014)

This chapter studies innovation contests, which are widely observed in a variety of industries. In many innovation contests, some firms have head starts: One firm has a more advanced existing technology than its rivals at the outset of a competition. The opening excerpt addresses a prominent phenomenon that is often observed in innovation contests: Companies with a head start ultimately lose a competition in the long run. It seems that having a head start sometimes results in being trapped. The failure of Nokia, the former global mobile communications giant, to compete with the rise of Apple's iPhone is one example. James Surowiecki (2013) pointed out that Nokia's focus on (improving) hardware, its existing technology, and neglect of (innovating) software contributed to the company's downfall. In his point of view, this was "a classic case of a company being enthralled (and, in a way, imprisoned) by its past success" (*New Yorker Times*, September 3, 2013).

Motivated by these observations, we investigate the effects of head starts on firms' competition strategies and payoffs in innovation contests. Previous work on innovation contests focuses on reduced form games and symmetric players, and previous work on contests with head starts considers all-pay auctions with either sequential bidding or simultaneous bidding. By contrast, we consider a stochastic contest model in which one firm has a superior existing innovation at the outset of the contest and firms' decisions are dynamic. The main contribution of our study is the identification of the long-run effects of a head start. In particular, in a certain range of the head start value, the *head start firm* becomes the ultimate loser in the long run and its competitor (or competitors) benefits greatly from its initial apparent "disadvantage". The key insight to the above phenomenon is that a large head start (e.g., a patent) indicates a firm's demise as an innovator.

Specifically, the model we develop in Section 3.3 entails two firms and one fixed prize. At the beginning of the game, each firm may or may not have an initial innovation. Whether a firm has an initial innovation, as well as the value of the initial innovation if this firm has one, is common knowledge. If a firm conducts a search for innovations, it incurs a search cost. As long as a firm continues searching, innovations arrive according to a Poisson process. The value of each innovation is drawn independently from a fixed distribution. The search activity and innovation process of each firm are privately observed. At any time point before a common deadline, each firm decides whether to stop its search process. At the deadline, each firm releases its most effective innovation to the public, and the one whose released innovation is deemed superior wins the prize.

First, we consider equilibrium behavior in the benchmark case, in which no firm has any innovation initially, in Section 3.4. We divide the deadline-cost space into three regions (as in figure 3.1). For a given deadline, (1) if the search cost is relatively high, there are two equilibria, in each of which one firm searches until it discovers an innovation and the other firm does not search; (2) if the search cost is in the middle range, each firm searches until it discovers an innovation; (3) if the search cost is relatively low, each firm searches until it discovers an innovation with a value above a certain positive cut-off value. In the third case, the equilibrium cut-off value strictly increases as the deadline extends and the arrival rate of innovations increases, and it strictly decreases as the search cost increases.

We then extend the benchmark case to include a head start: The head start firm is assigned a better initial innovation than its competitor, called the latecomer. Section 3.5 considers equilibrium behavior in the case with a head start and compares equilibrium payoffs across firms, and Section 3.6 analyzes the effects of a head start on each firm's equilibrium payoff.

Firms' equilibrium strategies depend on the value of the head starter's initial innovation (head start). Our main findings concern the case in which the head start lies in the middle range. In this range, the head starter loses its incentive to search because of its high initial position. The latecomer takes advantage of that and searches more actively, compared to when there is no head start.

An immediate question is: who does the head start favor? When the deadline is short, the latecomer does not have enough time to catch up, and thus the head starter obtains a higher expected payoff than the latecomer does. When the deadline is long, the latecomer is highly likely to obtain a superior innovation than the head starter, and thus the latecomer obtains a higher expected payoff. In the latter case, the latecomer's initial apparent "disadvantage", in fact, puts it in a more favorable position than the head starter. When the deadline is sufficiently long, the head starter is doomed to lose the competition with a payoff of zero because of its unwillingness to search, and all benefits of the head start goes to the latecomer.

Then, does the result that the latecomer is in a more favorable position than the head starter when the deadline is long imply that the head start hurts the head starter and benefits the latecomer in the long run? Focus on the case in which the latecomer does not have an initial innovation. When the search cost is relatively low, the head start, in fact, always benefits the head starter, but the benefit ceases as the deadline extends. It also benefits the latecomer when the deadline is long. When the search cost is relatively high, the head start could potentially hurt the head starter.

If the head start is large, neither firm will conduct a search, because the latecomer is deterred from competition. In this scenario, no innovation or technological progress is created, and the head starter wins the contest directly. If the head start is small, both firms play the same equilibrium strategy as they do when neither firm has an initial innovation. In both cases, the head start benefits the head starter and hurts the latecomer.

Section 3.7.1 extends our model to include stages at which the firms sequentially have an option to discard their initial innovation before the contest starts. Suppose that both firms' initial innovations are of values in the middle range and that the deadline is long. If the head starter can take the first move in the game, it can increase its expected payoff by discarding its initial innovation and committing to search. When search cost is low, by sacrificing the initial innovation, the original head starter actually makes the competitor the new head starter; this new head starter has no incentive to discard its initial innovation or to search any more. It is possible that by discarding the head start, the original head starter may benefit both firms. When search cost is high, discarding the initial innovation is a credible threat to the latecomer, who will find the apparent leveling of the playing field discouraging to conducting a high-cost search. As a result, the head starter suppresses the innovation progress.

In markets, some firms indeed give up head starts (Ulhøi, 2004), and our result provides a partial explanation of this phenomenon. For example, Tesla gave up its patents for its advanced technologies on electric vehicles at an early stage of its business.[1] While there may be many reasons for doing so, one significant reason is to maintain Tesla's position as a leading innovator in the electronic vehicle market.[2] As Elon Musk (2014), the CEO of Tesla, wrote,

> technology leadership is not defined by patents, which history has repeatedly shown to be small protection indeed against a determined competitor, but rather by the ability of a company to attract and motivate the world's most talented engineers.[3]

Whilst Tesla keeps innovating to win a large share of the future market, its smaller competitors have less incentive to innovate since they can directly adopt Tesla's technologies. One conjecture which coincides with our result is that "Tesla might be planning to distinguish itself from the competitors it helps . . . by inventing and patenting better electric cars than are available today" (*Discovery Newsletter*, June 13, 2014).

Section 3.7.2 considers intermediate information disclosure. Suppose the firms are required to reveal their discoveries at an early time point after the starting of the contest, how would firms compete against each other? If the head start is in the middle range, before the revelation point, the head starter will conduct a search, whereas the latecomer will not. If the head starter obtains a very good innovation before that point, the latecomer will be deterred from competition. Otherwise, the head starter is still almost certain to lose the competition. Hence, such an information revelation at an early time point increases both the expected payoff to the head starter and the expected value of the winning innovation.

Section 3.8 compares the effects of a head start to those of a cost advantage and points out a significant difference. A cost advantage reliably encourages a firm to search more actively for innovations, whereas it discourages the firm's competitor.

Section 3.9 concludes this chapter. The overarching message this chapter conveys is that a market regulator who cares about long-run competitions in markets may not need to worry too much about the power of the current market dominating firms if these firms are not in excessively high positions. In the long run, these firms are to be defeated by latecomers. On the other hand, if the dominating firms are in excessively high positions, which deters entry, a regulator can intervene the market.

[1] Toyota also gave up patents for its hydrogen fuel cell vehicles at an early stage.

[2] Another reason is to help the market grow faster by the diffusion of its technologies. A larger market increases demand and lowers cost.

[3] See "All Our Patent Are Belong To You," June 12, 2014, on http://www.teslamotors.com/blog/all-our-patent-are-belong-you.

3.2 Literature

There is a large literature on innovation contests. Most work considers reduced form models (Fullerton and McAfee, 1999; Moldovanu and Sela, 2001; Baye and Hoppe, 2003; Che and Gale, 2003).[4] Head starts are studied in various forms of all-pay auctions. Leininger (1991), Konrad (2002), and Konrad and Leininger (2007) model a head start as a first-mover advantage in a sequential all-pay auction and study the first-mover's performance. Casas-Arce and Martinez-Jerez (2011), Siegel (2014), and Seel (2014) model a head start as a handicap in a simultaneous all-pay auction and study the effect on the head starter. Kirkegaard (2012) and Seel and Wasser (2014) also model a head start as a handicap in a simultaneous all-pay auction but study the effect on the auctioneer's expected revenue. Segev and Sela (2014) analyzes the effect a handicap on the first mover in a sequential all-pay auction. Unlike these papers, we consider a framework in which players' decisions are dynamic.

The literature considering settings with dynamic decisions is scarce, and most studies focus on symmetric players. The study by Taylor (1995) is the most prominent.[5] In his symmetric T-period private search model, there is a unique equilibrium in which players continue searching for innovations until they discover one with a value above a certain cut-off. We extend Taylor's model to analyze the effects of a head start and find the long-run effects of the head start, which is our main contribution.

Seel and Strack (2013, 2016) and Lang et al. (2014) also consider models with dynamic decisions. Same as in our model, in these models each player also solves an optimal stopping problem. However, the objectives and the results of these papers are different from ours. In the models of Seel and Strack (2013, 2016), each player decides when to stop a privately observed Brownian motion with a drift. In their earlier model, there is no deadline and no search cost and a process is forced to stop when it hits zero. They find that players do not stop their processes immediately even if the drift is negative. In their more resent model, each search incurs a cost that depends on the stopping time. This more recent study finds that when noise vanishes the equilibrium outcome converges to the symmetric equilibrium outcome of an all-pay auction. Lang et al. (2014) consider a multi-period model in which each player decides when to stop a privately observed stochastic points-accumulation process. They find that in equilibrium the distribution

[4]Also see, for example, Hillman and Riley (1989), Baye et al. (1996), Krishna and Morgan (1998), Che and Gale (1998), Cohen and Sela (2007), Schöttner (2008), Bos (2012), Siegel (2009, 2010, 2014), Kaplan et al. (2003), and Erkal and Xiao (2015).

[5]Innovation contests were modeled as a race in which the first to reach a defined finishing line gains a prize, e.g., Loury (1979), Lee and Wilde (1980), and Reinganum (1981, 1982).

over successes converges to the symmetric equilibrium distribution of an all-pay auction when the deadline is long.

Our study also contributes to the literature on information disclosure in innovation contests. Aoyagi (2010), Ederer (2010), Goltsman and Mukherjee (2011), and Wirtz (2013) study how much information on intermediate performances a contest designer should disclose to the contestants. Unlike what we do, these papers consider two-stage games in which the value of a contestant's innovation is its total outputs from the two stages. Bimpikis et al. (2014) and Halac et al. (2017) study the problem of designing innovation contests, which includes both the award structures and the information disclosure policies. Halac et al. (2017) consider a model in which each contestant searches for innovations, but search outcomes are binary. A contest ends after the occurrence of a single breakthrough, and a contestant becomes more and more pessimistic over time if there has been no breakthrough. Bimpikis et al. (2014) consider a model which shares some features with Halac et al. (2017). In the model, an innovation happens only if two breakthroughs are achieved by the contestants, the designer decides whether to disclose the information on whether the first breakthrough has been achieved by a contestant, and intermediate awards can be used. In both models, contestants are symmetric. In contrast, the contestants in our model are always asymmetric. Rieck (2010) studies information disclosure in the two-period case of Taylor's (1995) model. In contrast to our finding, he shows that the contest designer prefers concealing the outcome in the first stage. Unlike all the above papers, Gill (2008), Yildirim (2005), and Akcigit and Liu (2015) address the incentives for contestants, rather than the designer, to disclose intermediate outcomes.

Last but most importantly, our study contributes to the literature on the relationship between market structure and incentive for R&D investment. The debate over the effect of market structure on R&D investment dates back to Schumpeter (1934, 1942).[6] Due to the complexity of the R&D process, earlier theoretical studies tend to focus on one facet of the process. Gilbert and Newbery (1982), Fudenberg et al. (1983), Harris and Vickers (1985a,b, 1987), Judd (2003), Grossman and Shapiro (1987), and Lippman and McCardle (1987) study preemption games. In these models, an incumbent monopolist has more incentive to invest in R&D than a potential entrant. In fact, a potential entrant sees little chance to win the competition, because of a lag at the starting point of the competition, and is deterred from competition. In our model, the intuition for the result in the case of a large head start is similar to this "preemption effect", except that no firm invests in our case.

[6]See Gilbert (2006) for a comprehensive survey.

By contrast, Arrow (1962) and Reinganum (1983, 1985) show, in their respective models, that an incumbent monopolist has less incentive to innovate than a new entrant.[7] The cause for this is what is called the "replacement effect" by Tirole (1997). While an incumbent monopolist can increase its profit by innovating, it has to lose the profit from the old technology once it adopts a new technology. This effectively reduces the net value of the new technology to the incumbent. It is then natural that a firm who has a lower value of an innovation has less incentive to innovate, which is exactly what happens in our model with asymmetric costs. On the other hand, our main result, on medium-sized head start, has an intuition very similar to the "replacement effect". Rather than a reduction in the value of an innovation to the head starter, a head start decreases the increase in the probability of winning from innovating. In both Reinganum's models and our model, an incumbent could have a lower probability of winning than a new entrant. However, different from her models, in our model an incumbent (head starter) can also have a lower expected payoff than a new entrant (latecomer).

3.3 The Model

Firms and Tasks

There are two risk neutral firms, Firm 1 and Firm 2, competing for a prespecified prize, normalized to 1, in the contest. Time is continuous, and each firm searches for innovations before a deadline T. At the deadline T, each firm releases to the public the best innovation it has discovered, and the firm who releases a superior innovation wins the prize. If no firm has discovered any innovation, the prize is retained. If there is a tie between the two firms, the prize is randomly allocated to them with equal probability.

At any time point $t \in [0, T)$ before the deadline, each firm decides whether to continue searching for innovations. If a firm continues searching, the arrival of innovations in this firm follows a Poisson process with an arrival rate of λ. That is, the probability of discovering m innovations in an interval of length δ is $\frac{e^{-\lambda\delta}(\lambda\delta)^m}{m!}$. The values of innovations are drawn independently from a distribution F, defined on $(0, 1]$ with $F(0) := \lim_{a\to 0} F(a) = 0$. F is continuous and strictly increasing over the domain.

Each firm's search cost is $c > 0$ per unit of time. We assume that $c < \lambda$, because if $c > \lambda$ the cost is so high that no firm is going to conduct a search. To illustrate this claim, suppose Firm 2 does not search, Firm 1 will not continue searching if it has an innovation with a value above 0,

[7] Doraszelski (2003) generalizes the models of Reinganum (1981, 1982) to a history-dependent innovation process model and shows, in some circumstances, the catching-up behavior in equilibrium.

whereas Firm 1's instantaneous gain from searching at any moment when it has no innovation is

$$\lim_{\delta \to 0} \frac{\sum_{m=1}^{+\infty} \frac{e^{-\lambda\delta}(\lambda\delta)^m}{m!} - c\delta}{\delta} = \lambda - c,$$

which is negative if $c > \lambda$.

Information

The search processes of the two firms are independent and with recall. Whether the opponent firm is actively searching is unobservable; whether a firm has discovered any innovation, as well as the values of discovered innovations, is private information until the deadline T.

For convenience, we say a firm is in a **state** $a \in [0, 1]$ at time t if the value of the best innovation it has discovered by time t is a, where $a = 0$ means that the firm has no innovation. The **initial states** of Firm 1 and Firm 2 are denoted by a_1^I and a_2^I, respectively. Firms' initial states are commonly known.

Strategies

In our model, each firm's information on its opponent is not updated. Hence, the game is static, although the firms' decisions are dynamic. Then, the solution concept we use is Nash equilibrium. In accordance with the standard result from search theory that each firm's optimal strategy is a constant cut-off rule, we make the following assumption.[8]

Assumption 3.1. *We focus on equilibria that consist of* ***constant cut-off rules****: Denote $\hat{a}_i^*$ as an equilibrium strategy. $\hat{a}_i^* \in \mathcal{S}_i := \{-1\} \cup [a_i^I,\ 1]$.*

If a constant cut-off strategy $\hat{a}_i^* \in [a_i^I,\ 1]$ is played, at any time point $t \in [0, T)$, Firm i stops searching if it is in a state above $\hat{a}_i^*$ and continues searching if in a state below or at $\hat{a}_i^*$.[9] The strategy $\hat{a}_i = -1$ represents that Firm i does not conduct a search.

Suppose both firms have no initial innovation. Without this assumption, for any given strategy played by a firm's opponent, there is a constant cut-off rule being the firm's best response. Such a cut-off value being above zero is the unique best response strategy, ignoring elements associated with zero probability events. However, in the cases in which a firm is indifferent between continuing searching and not if it is in state 0, this firm has (uncountably) many best response strategies.

[8] See Lippman and McCall (1976) for the discussion on optimal stopping strategies for searching with finite horizon and recall.

[9] Once Firm i stops searching at some time point it shall not search again later.

The above assumption helps us to focus on the two most natural strategies: not to search at all and to search with 0 as the cut-off.[10] A full justification for this assumption is provided in the appendix.

Let $\tilde{P}[a|\hat{a}_i, a_i^I]$ denote the probability of Firm i ending up in a state below a if it adopts a strategy $\hat{a}_i$ and its initial state is a_i^I; let $E[cost|\hat{a}_i]$ denote Firm i's expected cost on search if it adopts a strategy $\hat{a}_i$. Firm i's ex ante expected utility is

$$U_i = \int_0^1 P[a|\hat{a}_{-i}, a_{-i}^I] dP[a|\hat{a}_i, a_i^I] - E[cost|\hat{a}_i].$$

Now, we are ready to study equilibrium behavior. Before solving the head start case, we first look at the case with no initial innovation.

3.4 The Symmetric-Firms Benchmark ($a_1^I = a_2^I = 0$)

In this section, we look at the benchmark case, in which both firms start with no innovation. It is in the spirit of Taylor's (1995), except that it is in continuous time. The equilibrium strategies are presented below.[11]

Theorem 3.1. *Suppose $a_1^I = a_2^I = 0$.*

i. If $c \in [\frac{1}{2}\lambda(1 + e^{-\lambda T}), \lambda)$, there are two equilibria, in each of which one firm searches with 0 as the cut-off and the other firm does not search.

ii. If $c \in [\frac{1}{2}\lambda(1 - e^{-\lambda T}), \frac{1}{2}\lambda(1 + e^{-\lambda T}))$, there is a unique equilibrium, in which both firms search with 0 as the cut-off.

iii. If $c \in (0, \frac{1}{2}\lambda(1 - e^{-\lambda T}))$, there is a unique equilibrium, in which both firms search with a^* as the cut-off, where $a^* > 0$ is the unique value that satisfies

$$\frac{1}{2}\lambda[1 - F(a^*)]\left[1 - e^{-\lambda T[1-F(a^*)]}\right] = c. \tag{3.1}$$

Proof. See Appendix 3.A.2. □

---

[10] Without this assumption, there can be additional best response strategies of the following type: a firm randomizes between searching and not searching until a time $T' < T$ with cutoff 0 and stops at T' even if no discovery was made.

[11] When search cost is low, the equilibrium is unique even without Assumption 3.1. When search is high, without Assumption 3.1, there are additional symmetric equilibria of the following type: each firm randomizes between not participating and participating until a time $T' < T$ with 0 as the cutoff.

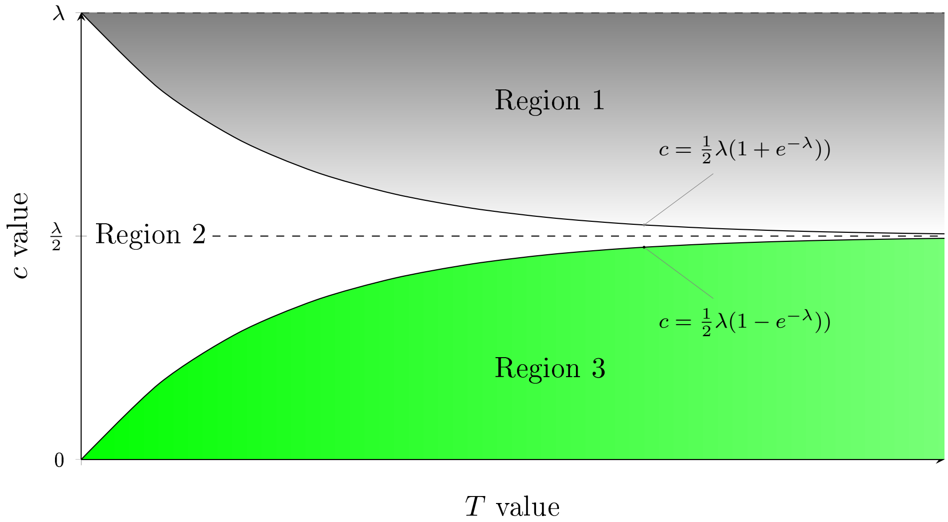


Figure 3.1

The result is illustrated in figure 3.1. The deadline-cost space is divided into three regions.[12] In Region 1, the search cost is so high that it is not profitable for both firms to innovate. In Region 2, both firms would like to conduct a search in order to discover an innovation with any value, but none has the incentive to spend additional effort to find an innovation with a high value. In Region 3, both firms exert efforts to find an innovation with a value above a certain level. In this case, a firm in the cut-off state is indifferent between continuing and stopping searching. This is represented by equation (3.1), in which $1 - e^{-\lambda T[1-F(a^*)]}$ is the probability of a firm's opponent ending up in a state above a^* and $\frac{1}{2}[1 - F(a^*)]$ is the increase in the probability of winning if the firm, in state a^*, obtains a new innovation. Hence, this equation represents that, in the cut-off state, the instantaneous increase in the probability of winning from continuing searching equals the instantaneous cost of searching. As T goes to infinity, $c = \frac{\lambda}{2}$ becomes the separation line for Case [i] and Case [iii].

Generally, there is no closed form solution for the cut-off value in Case [iii]. However, if the search cost is very low, we have a simple approximation for it.

Corollary 3.1. *Suppose $a_1^I = a_2^I = 0$. When c is small, $a^* \approx F^{-1}\left(1 - \sqrt{\frac{2c}{\lambda^2 T}}\right)$.*

Proof. First, we assume that $\lambda T[1 - F(a^*)]$ is small, and we come back to check that it is implied

[12]An " area" we say is the interior of the corresponding area.

by that c is small. Applying equation (3.1), we have

$$\begin{aligned}
&\frac{c}{\lambda} = \frac{1}{2}[1 - F(a^*)]\left[1 - e^{-\lambda T[1-F(a^*)]}\right] \approx \frac{1}{2}\lambda T[1 - F(a^*)]^2 \\
\Leftrightarrow &[1 - F(a^*)]^2 \approx \frac{2c}{\lambda^2 T} \\
\Leftrightarrow &a^* \approx F^{-1}\left(1 - \sqrt{\frac{2c}{\lambda^2 T}}\right) \quad \text{and} \quad \lambda T[1 - F(a^*)] \approx \sqrt{2cT}.
\end{aligned}$$

□

For later reference we, based on the previous theorem, define a function $a^* : (0, \lambda) \times [0, +\infty) \to [0, 1]$ where

$$a^*(c, T) = \begin{cases} 0 & \text{for } c \in [\frac{1}{2}\lambda(1 - e^{-\lambda T}), \lambda) \\ \text{the } a^* \text{ that solves (3.1)} & \text{for } c \in (0, \frac{1}{2}\lambda(1 - e^{-\lambda T})). \end{cases}$$

A simple property which will be used in later sections is stated below.

Lemma 3.1. *In Region* 3*, $a^*(c, T)$ is strictly increasing in T (and λ) and strictly decreasing in c.*

There are two observations. One is that $a^*(c, T) = 0$ if $c \geq \frac{\lambda}{2}$. The other is that $a^*(c, T)$ converges to $F^{-1}(1 - \frac{2c}{\lambda})$ as T goes to infinity if $c < \frac{\lambda}{2}$, which derives from taking the limit of equation (3.1) w.r.t. T. Let us denote a^*_L as the limit of $a^*(c, T)$ w.r.t. T:

$$a^*_L := \lim_{T \to +\infty} a^*(c, T) = \begin{cases} 0 & \text{for } c \geq \frac{\lambda}{2}, \\ F^{-1}(1 - \frac{2c}{\lambda}) & \text{for } c < \frac{\lambda}{2}. \end{cases}$$

We end this section by presenting a full rent dissipation property of the contest when the deadline approaches infinity.

Lemma 3.2. *Suppose $a^I_1 = a^I_2 = 0$. If $c < \frac{\lambda}{2}$, each firm's expected payoff in equilibrium goes to* 0 *as the deadline T goes to infinity.*

Proof. See Appendix 3.A.2. □

The intuition is as follows. The instantaneous increase in the expected payoff from searching for a firm who is in state $a^*(c, T)$, the value of the equilibrium cut-off, is 0 (it is indifferent between continuing searching and not). If the deadline is finite, a firm in a state below $a^*(c, T)$ has a positive probability of winning even if it stops searching. Hence, the firms have positive rents in the contest. As the deadline approaches infinity, there is no difference between being in a state below $a^*(c,T)$ and at $a^*(c,T)$, because the firm will lose the contest for sure if it does not search. In either case the instantaneous increase in the expected payoff from searching is 0. Hence, in the limit the firms' rents in the contest are fully dissipated.

Though the equilibrium expected payoff goes to 0 in the limit, it is not monotonically decreasing to 0 as the deadline approaches infinity, because each firm's expected payoff converges to 0 as the deadline approaches 0 as well.[13]

3.5 Main Results: Exogenous Head Starts ($a_1^I > a_2^I$)

In this section, we add head starts into the study. Without loss of generality, we assume that Firm 1 has a better initial innovation than does Firm 2 before competition begins, i.e., $a_1^I > a_2^I$. We first derive the equilibrium strategies, and then we explore equilibrium properties.

3.5.1 Equilibrium Strategies

Theorem 3.2. *Suppose $a_1^I > a_2^I$.*

1. *For $a_1^I > F^{-1}(1-\frac{c}{\lambda})$, there is a unique equilibrium, in which no firm searches, and thus Firm 1 wins the prize.*

2. *For $a_1^I = F^{-1}(1-\frac{c}{\lambda})$, there are many equilibria. In one equilibrium, both firms do not search. In the other equilibria, Firm 1 does not search and Firm 2 searches with a value $\hat{a}_2 \in [a_2^I, a_1^I]$ as the cut-off.*

3. *For $a_1^I \in (a^*(c,T), F^{-1}(1-\frac{c}{\lambda}))$, there is a unique equilibrium, in which Firm 1 does not search and Firm 2 searches with a_1^I as the cut-off .*

4. *For $a_1^I = a^*(c,T)$, there are two equilibria. In one equilibrium, both firms search with a_1^I as the cut-off. In the other equilibrium, Firm 1 does not search and Firm 2 searches with a_1^I as the cut-off.*

5. *For $a_1^I \in (0, a^*(c,T))$, there is a unique equilibrium, in which both firms search with $a^*(c,T)$ as the cut-off.*

Proof. See Appendix 3.A.3. □

Remark. *Case [4] and [5] exist only when $c \leq \frac{1}{2}\lambda[1-e^{-\lambda T}]$ (Region 3).*

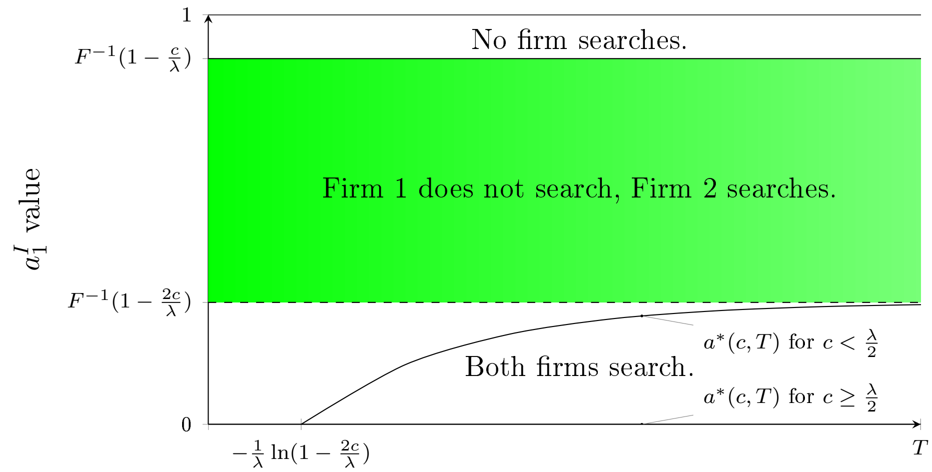


Figure 3.2: Thresholds (when $c < \frac{\lambda}{2}(1-e^{-\lambda T})$).

The thresholds in the theorem are depicted in figure 3.2. The leading case is Case [3], when the head start is in the middle range. A head start reduces the return of a search, in terms of the increase in the probability of winning. Having a sufficiently large initial innovation, Firm 1 loses incentive to search because the marginal increase in the probability of winning from searching for Firm 1 is too small compared to the marginal cost of searching, whether Firm 2 searches or not. Firm 2 takes advantage of that and commits to search until it discovers an innovation better than Firm 1's initial innovation. Hence, compared to its equilibrium behavior in the benchmark case, Firm 2 is more active in searching (in terms of a higher cut-off value) when Firm 1 has a medium-sized head start, and a larger value of head start forces Firm 2 to search more actively.

In Case [1], Firm 1's head start is so large that Firm 2 is deterred from competition because Firm 2 has little chance to win if it searches. Firm 1 wins the prize without incurring any cost. Moreover, it is independent of the deadline T.

In Case [5], in which Firm 1's head start is small, the head start has no effect on either firm's equilibrium strategy, and both firms search with $a^*(c,T)$ as the cut-off, same as in the benchmark case. The only effect of the head start is an increase in Firm 1's probability of winning (and a decrease in Firm 2's).

In brief, a comparison of Theorem 3.2 and Theorem 3.1 shows that a head start of Firm 1 does not alter its own equilibrium behavior but Firm 2's. The effect on Firm 2's equilibrium strategy is

[13] In fact, by taking the derivative of (3.12) (as in the appendix) w.r.t. T, one can show that the derivative at $T=0$ is $\lambda - c > 0$ and that, if $c < \frac{\lambda}{2}$, (3.12) is increasing in T for $T < \min\{\frac{1}{\lambda}\ln\frac{\lambda}{c}, \frac{1}{\lambda}\ln\frac{\lambda}{\lambda-2c}\}$ and decreasing in T for $T > \max\{\frac{1}{\lambda}\ln\frac{\lambda}{c}, \frac{1}{\lambda}\ln\frac{\lambda}{\lambda-2c}\}$.

not monotone in the head start of Firm 1. The initial state of Firm 2, the latecomer, is irrelevant to the equilibrium strategies. Figure 3.3 illustrates how each firm's equilibrium strategy changes as the value of the initial innovation of Firm 1, the head starter, varies.

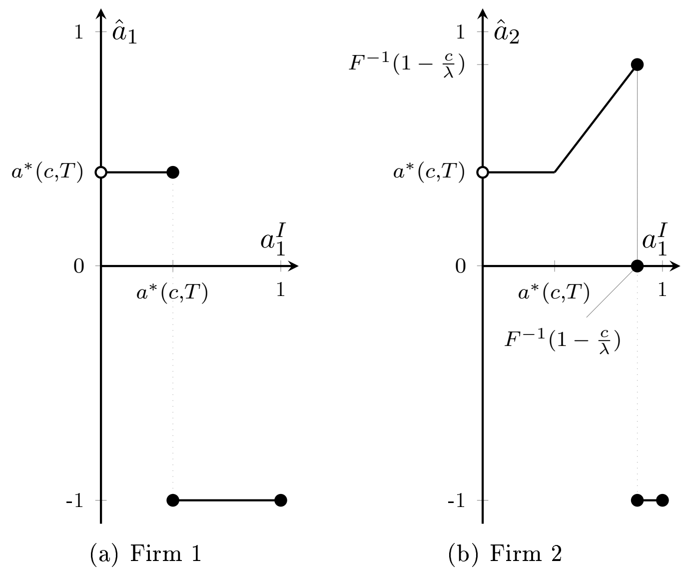


Figure 3.3: Firms' equilibrium cut-off values as the value of Firm 1's initial innovation, a_1^I, varies.

Figure 3.4 illustrates Firm 2'the best responses (when it has no initial innovation) to Firm 1's strategies for various values of Firm 1's initial states. The case in which Firm 1 has a high-value initial innovation is significantly different from the case in which Firm 1 has no initial innovation.

Turning back to Case [3] in the previous result, we notice that the lower bound for this case to happen does not converge to the upper bound as the deadline approaches infinity, i.e., $a_L^* < F^{-1}(1-\frac{c}{\lambda})$. The simplest but most interesting result of this study, the case of "head starts and doomed losers", derives.

Corollary 3.2. *Suppose* $a_1^I \in (a_L^*, F^{-1}(1-\frac{c}{\lambda}))$.

1. *Firm 2's (Firm 1's) probability of winning increases (decreases) in the deadline.*

2. *As T goes to infinity, Firm 2's probability of winning goes to* 1, *and Firm 1's goes to* 0.

Proof. In equilibrium Firm 1 does not search and Firm 2 searches with a_1^I as the cut-off. Firm 2's probability of winning is thus

$$1 - e^{-\lambda T[1-F(a_1^I)]},$$

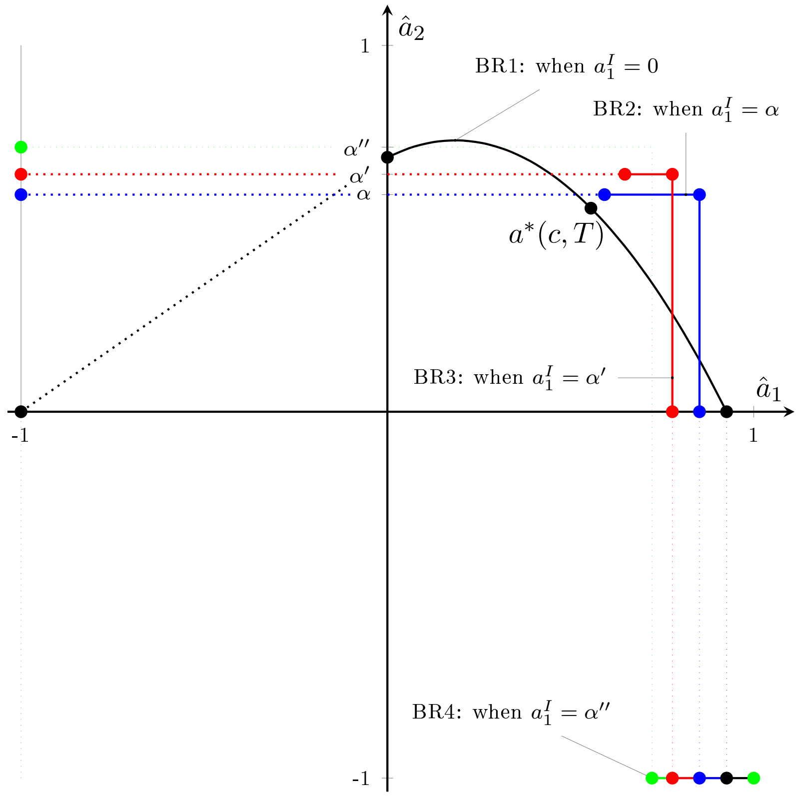


Figure 3.4: Best response projections for Firm 2 as a_1^I takes values in $\{-1, \alpha, \alpha', \alpha''\}$ ($c < \frac{\lambda}{2}(1 - e^{-\lambda T})$ and $a_2^I = 0$).

BR1 represents Firm 2's best responses when no firm has an initial innovation. If Firm 1 does not search, Firm 2 would search with 0 as the cut-off. If Firm 1 searches with 0 as the cut-off, Firm 2 would search with a cut-off higher than the equilibrium cut-off. As Firm 1 further rises its cut-off, Firm 2 would first rise its cut-off and then lower its cut-off. When the deadline is long, Firm 2 would not search if Firm 1's cut-off is high. BR2 and BR3 represent Firm 2's best responses when Firm 1 has an initial innovation with a value slightly above $a^*(c,T)$, the equilibrium cut-off when there is no initial innovation. In this case, if Firm 1 does not search, Firm 2's best response is to search with the value Firm 1's initial innovation as the cut-off. If Firm 1 searches with a cut-off slightly above the value of its initial innovation, Firm 2's best response is still to search with the value of Firm 1's initial innovation as the cut-off. Once Firm 1's cut-off is greater than a certain value, Firm 2 would not search. BR4 represents Firm 2's best responses when Firm 1 has a high-value initial innovation. In this case, if Firm 1 does not search, Firm 2 would still search with the value of Firm 1's initial innovation as the cut-off; if Firm 1 searches, Firm 2 would have no incentive to search. On the other hand, when the value of Firm 1's initial innovation is above $a^*(c,T)$, Firm 1's best response to any strategy of Firm 2 is not to search.

which is increasing in T, and it converges to 1 as T goes infinity. Firm 1's probability of winning is $e^{-\lambda T[1-F(a_1^I)]}$, which is, in the contrast, decreasing in T, and it converges to 0 as T approaches infinity. □

This property results from our assumption that search processes are with recall. The larger the head start is, the smaller the marginal increase in the probability of winning from searching is, given any strategy played by the latecomer. Hence, even if the head starter knows that in the long run the latecomer will almost surely obtain an innovation with a value higher than its initial innovation, the head starter is not going to conduct a search as long as the instantaneous increase in the probability of winning is smaller than the instantaneous cost of searching.

3.5.2 Payoff Comparison across Firms

A natural question arises: which firm does a head start favor? Will Firm 1 or Firm 2 achieve a higher expected payoff? To determine that, we need a direct comparison of the two firms' expected payoffs. When $a_1^I \in (a^*(c,T), F^{-1}(1-\frac{c}{\lambda}))$, the difference between the payoffs of Firm 1 and Firm 2 is[14]

$$D^F(T, a_1^I) := e^{-\lambda T[1-F(a_1^I)]} - (1 - e^{-\lambda T[1-F(a_1^I)]})(1 - \frac{c}{\lambda[1-F(a_1^I)]}). \tag{3.2}$$

The head start of Firm 1 favors Firm 1 (Firm 2) if $D^F(T, a_1^I) > (<) 0$.

$D^F(T, a_1^I)$ is increasing in a_1^I and decreasing in T. Hence, a longer deadline tends toward to favor Firm 2 when the head start is in the middle range. Since

$$D^F(0, a_1^I) = 1 > 0$$

and

$$\lim_{T\to\infty} D^F(T, a_1^I) = -(1 - \frac{c}{\lambda[1-F(a_1^I)]}) < 0 \text{ for any } a^I < F^{-1}(1-\frac{c}{\lambda}),$$

there must be a unique $\hat{T}(a_1^I) > 0$ such that $DE(\hat{T}(a_1^I), a_1^I) = 0$. The following result derives.

Proposition 3.1. *For $a_1^I \in (a_L^*, F^{-1}(1-\frac{c}{\lambda}))$, there is a unique $\hat{T}(a_1^I) > 0$ such that Firm 1 (Firm 2) obtains a higher expected payoff if $T < (>) \hat{T}(a_1^I)$.*

[14] $1 - e^{-\lambda T[1-F(a_1^I)]}$ is Firm 2's probability of obtaining an innovation better than Firm 1's initial innovation, a_1^I, and $\frac{1}{\lambda[1-F(a_1^I)]}$ is the unconditional expected interarrival time of innovations with a value higher than a_1^I. The second term in $D^F(T, a_1^I)$ thus represents the expected payoff of Firm 2.

That is, for any given value of the head start in the middle range $(a_L^*, F^{-1}(1-\frac{c}{\lambda}))$, the head start favors the latecomer (head starter) if the deadline is long (short). The effects of a head start do not vanish as the deadline approaches infinity. In fact, as the deadline approaches infinity, the head start eventually pushes the whole share of the surplus to Firm 2.

Lemma 3.3. *As the deadline increases to infinity,*

1. *for $a_1^I \in (0, a_L^*)$, both Firms' equilibrium payoffs converge to 0;*
2. *for $a_1^I \in (a_L^*, F^{-1}(1-\frac{c}{\lambda}))$, Firm 1's equilibrium payoff converges to 0, whereas Firm 2's equilibrium payoff converges to $1-\frac{c}{\lambda[1-F(a_1^I)]} \in (0, \frac{1}{2})$.*

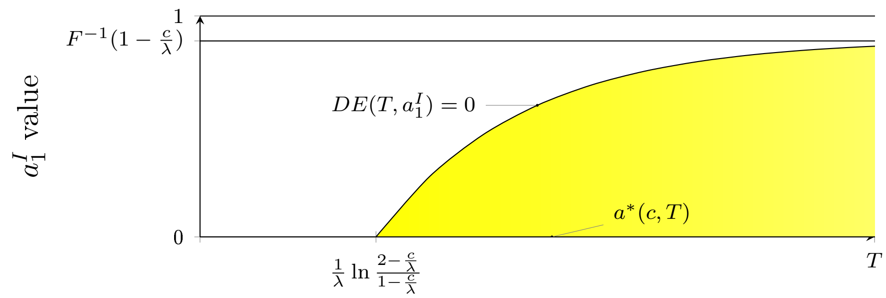
Proof. [1] follows from Lemma 3.2. [2] follows from Corollary 3.2 and the limit of Firm 2's expected payoff

$$(1-e^{-\lambda T[1-F(a^I)]})\left(1-\frac{c}{\lambda[1-F(a_1^I)]}\right) \tag{3.3}$$

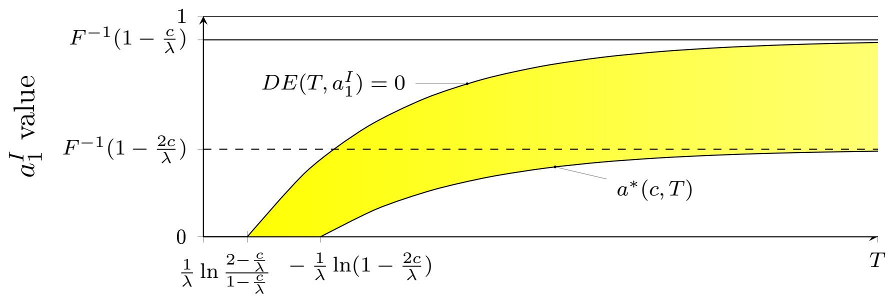
w.r.t. T. □

A comparison to Lemma 3.2 shows that, just as having no initial innovation, when Firm 1 has an innovation whose value is not of very high, its expected payoff still converges to 0 as the deadline becomes excessively long. When there is no initial innovation, the expected total surplus for the firms (i.e., the sum of the two firms' expected payoff) converges to 0. In contrast, when there is a head start with a value above a_L^*, the expected total surplus is strictly positive even when the deadline approaches infinity. However, as it approaches infinity, this total surplus created by the head start of Firm 1 goes entirely to Firm 2, the latecomer, if the head start is in the middle range. The intuition is as follows. For Firm 1, it is clear that its probability of winning converges to 0 as the deadline goes to infinity. For Firm 2, we first look at the case that $a_1^I = F^{-1}(1-\frac{c}{\lambda})$. In this case Firm 2 is indifferent between searching and not searching, and thus the expected payoff is 0. As the deadline approaches infinity, both expected cost of searching and the probability of winning converges to 1, if Firm 2 conducts a search. Then, if a_1^I is below $F^{-1}(1-\frac{c}{\lambda})$ (but above a_L^*), as the deadline approaches infinity, Firm 2's probability of winning still goes to 1, but the expected cost of searching drops to a value below 1 because it adopts a lower cut-off for stopping. Hence, Firm 2's expected payoff converges to a positive value.

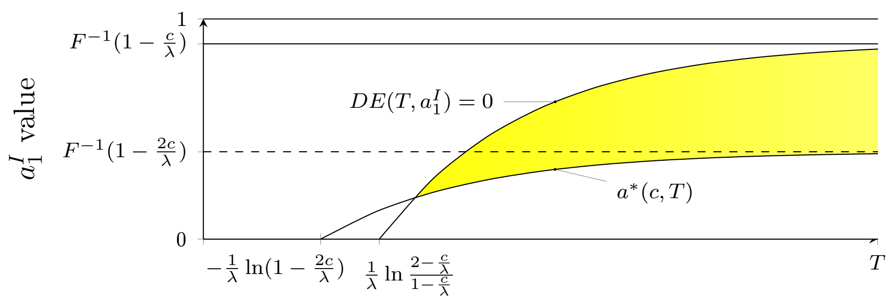
The relationship between the rank order of the two Firms' payoffs and the deadline is illustrated in figure 3.5, in each of which Firm 2 obtains a higher expected payoff at each point in the colored area. (a) is for the cases in which $c \leq \frac{\lambda}{2}$. In these cases a longer deadline tends to favor the latecomer. (b) and (c) are for the cases in which $c > \frac{\lambda}{2}$. In these cases, the rank order is not generally monotone in the deadline and the head start.



(a) $c \in [\frac{\lambda}{2}, \lambda]$



(b) $c \in [(1-\frac{\sqrt{2}}{2})\lambda, \frac{\lambda}{2}]$



(c) $c \in [0, (1-\frac{\sqrt{2}}{2})\lambda]$

Figure 3.5: When (a_1^I, T) lies in the colored area, the head start favors the latecomer.

We notice in all the figures that if the deadline is sufficiently short, a head start is ensured to bias toward Firm 1, whereas if it is long, only a relatively large head start biases toward Firm 1.

3.6 Effects of Head Starts on Payoffs

In this section, we study the effects of a head start on both firms' payoffs. Suppose Firm 2 has no initial innovation, who does a head start of Firm 1 benefit or hurt? The previous comparison between Theorem 3.1 and Theorem 3.2 already shows that a head start a_1^I benefits Firm 1 and hurts Firm 2 if $a_1^I < a^*(c,T)$ or $a_1^I > F^{-1}(1-\frac{c}{\lambda})$. In the former case, which happens only when $a^*(c,T) > 0$ (Region 3 of figure 3.1), both firms search with $a^*(c,T)$ as the cut-off, the same as when there is no head start, and the head start increases Firm 1's probability of winning and decreases Firm 2's. As the deadline goes to infinity, the expected payoffs to both firms converge to 0, with the effect of the head start disappearing. In the latter case, Firm 1 always obtains a payoff of 1, and Firm 2 always 0.

The interesting case occurs then when the head start is in the middle range, $a_1^I \in (a^*(c,T), F^{-1}(1-\frac{c}{\lambda}))$, which will be the focus in the remaining parts of this study. To answer the above question regarding the head start being in the middle range, we first analyze the case that point (c,T) lies in Regions 2 and 3 (in figure 3.1), and then we turn to analyze the case of Region 1.

3.6.1 Regions 2 and 3

In the previous section, we showed that for T being sufficiently long, Firm 1 is almost surely going to lose the competition if a_1^I is in the middle range. Although it seems reasonable that in this case a head start may make Firm 1 worse off, the following proposition shows that this conjecture is not true.

Proposition 3.2. *Suppose $a_2^I = 0$. In Regions 2 and 3, in which $c < \frac{1}{2}\lambda(1+e^{-\lambda T})$, a head start $a_1^I > 0$ always benefits Firm 1, compared to the equilibrium payoff it gets in the benchmark case.*

To give the intuition, we consider the case of $a^*(c,T) > 0$. Suppose Firm 1 has a head start $a_1^I = a^*(c,T)$. As shown in Case [4] of Theorem 3.2, we have the following two equilibria: in one equilibrium both firms search with $a^*(c,T)$ as the cut-off; in another equilibrium Firm 1 does not search and Firm 2 searches with $a^*(c,T)$ as the cut-off. Firm 1 is indifferent between these two equilibria, hence its expected payoffs from both equilibria are $e^{-\lambda T[1-F(a^*(c,T))]}$, the probability of Firm 2 finding no innovation with a value higher than $a^*(c,T)$. However, Firm 1's probability of winning increases in its head start, hence a larger head start gives Firm 1 a higher expected payoff.

The above result itself corresponds to expectation. What unexpected is the mechanism through which Firm 1 gets better off. As a head start gives Firm 1 a higher position, we would expect that it is better off by (1) having a better chance to win and (2) spending less on searching. Together with Theorem 3.2, the above proposition shows that Firm 1 is better off purely from an increase in the probability of winning when $a_1^I < a^*(c,T)$; purely from spending nothing on searching when $a_1^I \in (a^*(c,T), F^{-1}(1-\frac{c}{\lambda}))$ (though there could be a loss from a decrease in the probability of winning); from an increase in the probability of winning and a reduction in the cost of searching when $a_1^I > F^{-1}(1-\frac{c}{\lambda})$.

In contrast to the effect of a head start of Firm 1 on Firm 1's own expected payoff, the effect on Firm 2's expected payoff is not clear-cut. Instead of giving a general picture of the effect, we present some properties in the following.

Proposition 3.3. *Suppose $a_2^I = 0$.*

1. *A head start $a_1^I \in (0, F^{-1}(1-\frac{c}{\lambda}))$ hurts Firm 2 if the deadline T is sufficiently small.*

2. *If $c < \frac{\lambda}{2}$, a head start $a_1^I \in (a_L^*, F^{-1}(1-\frac{c}{\lambda}))$ benefits Firm 2 if the deadline is sufficiently long.*

Proof. See Appendix 3.A.3. □

Case [1] occurs because a head start of Firm 1 reduces Firm 2's probability of winning and may increase its expected cost of searching. Case [2] follows from Propositions 3.2 and 3.1. Because a head start of Firm 1 always benefits Firm 1 and a long deadline favors Firm 2, a head start must also benefit Firm 2 if the deadline is long.[15]

Figure 3.6 illustrates how Firm 2's equilibrium payoff changes as Firm 1's head start increases. In particular, a head start of Firm 1 slightly above $a^*(c,T)$, the equilibrium cut-off when there is no initial innovation, benefits Firm 2 if $a^*(c,T)$ is low. Some more conditions under which a head start benefits or hurts the latecomer are given below.

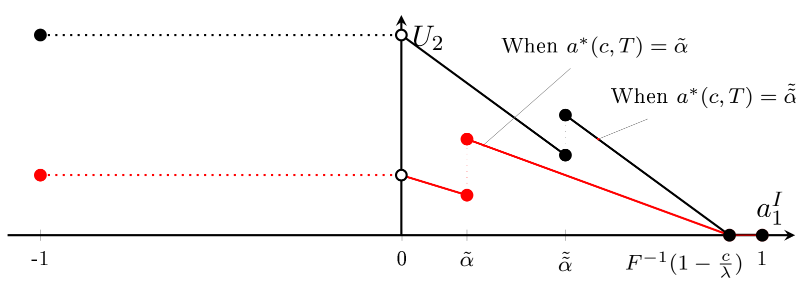


Figure 3.6: Firm 2's equilibrium payoffs as $a^*(c,T)$ varies.

Proposition 3.4. *In Region 2 and 3, in which $c < \frac{1}{2}\lambda(1+e^{-\lambda T})$,*

[15] Alternatively, it also follows from Lemmas 3.2 and 3.3. If the deadline is very long and the head start of Firm 1 is in the middle range, Firm 2's payoff converges to 0 same as in the benchmark case and some positive value in head start case.

1. *if*

$$(1 - e^{-\lambda T[1-F(a^*(c,T))]}) - \frac{1}{2}(1 - e^{-2\lambda T}) > 0, \tag{3.4}$$

there exists a $\tilde{a}_1^I \in (a^*(c,T), F^{-1}(1-\frac{c}{\lambda}))$ *such that the head start* a_1^I *hurts Firm 2 if* $a_1^I \in (\tilde{a}_1^I, F^{-1}(1-\frac{c}{\lambda}))$ *and benefits Firm 2 if* $a_1^I \in (a^*(c,T), \tilde{a}_1^I)$*;*

2. *if* (3.4) *holds in the opposite direction, any head start* $a_1^I \in (a^*(c,T), F^{-1}(1-\frac{c}{\lambda}))$ *hurts Firm* 2.

Proof. See Appendix 3.A.3. □

The first term on the left side of inequality (3.4) is Firm 2's probability of winning in the equilibrium in which Firm 2 searches and Firm 1 does not search in the limiting case that Firm 1 has a head start of $a^*(c,T)$. The second term, excluding the minus sign, is Firm 1's probability of winning when there is no head start. The expected searching costs are the same in both cases. The following corollary shows some scenarios in which inequality (3.4) holds.

Corollary 3.3. *In Region 2, when* $a^*(c,T) = 0$*, inequality* (3.4) *holds.*

This shows that for search cost lying in the middle range, a head start of Firm 1 must benefit Firm 2, if it is slightly above 0. The simple intuition is as follows. When Firm 1 has such a small head start, Firm 2's cut-off value of searching increases by only a little bit, and thus the expected cost of searching also increases slightly. However, the increase in Firm 2's probability of winning is very large, because Firm 1, when having a head start, does not search any more. Thus, in this case Firm 2 is strictly better off.

Lastly, even though Firm 1 does not search when the head start $a_1^I > a^*(c,T)$, it seems that a low search cost may benefit Firm 2. On the contrary, a head start of Firm 1 would always hurt Firm 2 when the search cost is sufficiently small.

Corollary 3.4. *For any fixed deadline* T*, if the search cost is sufficiently small, inequality* (3.4) *holds in the opposite direction.*

Proof. As c being close to 0, $a^*(c,T)$ is close to 1, and thus the term on left side of inequality (3.4) is close to $-\frac{1}{2}(1 - e^{-2\lambda T}) < 0$. □

That is because when c is close to 0, $a^*(c,T)$ is close to 1, and the interval in which Firm 1 does not search while Firm 2 searches is very small, and thus the chance for Firm 2 to win is too low when $a_1^I > a^*(c,T)$, even though the expected cost of searching is low as well.

3.6.2 Region 1

Since there are multiple equilibria in the benchmark case when (c, T) lies in Region 1, whether a head start hurts or benefits a firm depends on which equilibrium we compare to. If we compare the two equilibria in each of which Firm 1 does not search and Firm 2 searches, then the head start benefits Firm 1 and hurts Firm 2. If we compare to the other equilibrium in the benchmark case, the outcome is not clear-cut.

Proposition 3.5. *Suppose $a_2^I = 0$. In Region 1, in which $c > \frac{1}{2}\lambda(1 + e^{-\lambda T})$ and $a^*(c, T) = 0$, for $a_1^I \in (0, F^{-1}(1 - \frac{c}{\lambda}))$, if*

$$(1 - e^{-\lambda T})(1 - \frac{c}{\lambda}) - e^{-\lambda T[1 - F(a_1^I)]} < 0, \tag{3.5}$$

Firm 1's equilibrium payoff is higher than its expected payoff in any equilibrium in the benchmark case. If the inequality holds in the opposite direction, Firm 1's equilibrium payoff is lower than its payoff in the equilibrium in which Firm 1 searches and Firm 2 does not search in the benchmark case.

This result is straightforward. The first term on the left side of inequality (3.5) is Firm 1's expected payoff in the equilibrium in which Firm 1 searches and Firm 2 does not in the benchmark case and the second term, excluding the minus sign, is its expected payoff when there is no head start.

Moreover, the left hand side of inequality (3.5) strictly increases in T, and it reaches -1 when T approaches 0 and $1 - \frac{c}{\lambda}$ when T approaches infinity. The intermediate value theorem insures that inequality (3.5) holds in the opposite direction for the deadline T being large.

As a result of the above property, when the head start is small and the deadline is long, in an extended game in which Firm 1 can publicly discard its head start before the contest starts, there are two subgame perfect equilibria: in one equilibrium, Firm 1 does not discard its head start and Firm 2 searches with the Firm 1's initial innovation value as the cut-off; in the other equilibrium, Firm 1 discards the head start and searches with 0 as the cut-off and Firm 2 does not search. Hence, there is the possibility that Firm 1 can improve its expected payoff if it discards its head start.

Last, we discuss Firm 2's expected payoff. The result is also straightforward.

Proposition 3.6. *Suppose $a_2^I = 0$. In Region 1, in which $c > \frac{1}{2}\lambda(1 + e^{-\lambda T})$, for $a_1^I \in (0, F^{-1}(1 - \frac{c}{\lambda}))$, Firm 2's equilibrium payoff is*

- *less than its expected payoff in the equilibrium in which Firm* 1 *does not search and Firm* 2 *searches in the benchmark case, and*

- *higher than the payoff in the equilibrium in which Firm* 1 *searches and Firm* 2 *does not search in the benchmark case.*

Proof. Compared to the equilibrium in which Firm 2 searches in the benchmark case, in the equilibrium when Firm 1 has a head start, Firm 2 has a lower expected probability of winning and a higher expected cost because of a higher cut-off, and and thus a lower expected payoff. But this payoff is positive. □

3.7 Extended Dynamic Models ($a_1^I > a_2^I$)

In this section, we study two extended models.

3.7.1 Endogenous Head Starts

We first study how the firms would play if each firm has the option to discard its initial innovation before the contest starts. Formally, a game proceeds as below.

Model*:

- Stage 1: Firm i decides whether to discard its initial innovation.

- Stage 2: Firm i's opponent decides whether to discard its initial innovation.

- Stage 3: Upon observing the outcomes in the previous stages, both firms simultaneously start playing the contest as described before.

The incentive for a head starter to discard its initial innovation when the latecomer has no initial innovation has been studied in the previous section. The focus of the section is on the case in which both firms have an initial innovation in the middle range.[16] The main result of this section is as follows.

Proposition 3.7. *Suppose* $a_1^I, a_2^I \in (a_L^*, F^{-1}(1 - \frac{c}{\lambda}))$*. In Model* with Firm 1 having the first move, there is a* $\breve{T}(a_1^I, a_2^I) > 0$ *such that*

[16]As shown in Proposition 3.2, when (c, T) lies in Regions 2 and 3, the head starter with a medium-sized initial innovation has no incentive to discard its head start if the latecomer has a no innovation. The head starter would also have no incentive to discard its initial innovation when the latecomer has a low-value initial innovation.

- *if* $T > \breve{T}(a_1^I, a_2^I)$*, there is a unique subgame perfect equilibrium (SPE), in which Firm* 1 *discards its initial innovation and searches with* a_2^I *as the cut-off and Firm* 2 *keeps its initial innovation and does not search;*
- *if* $T < \breve{T}(a_1^I, a_2^I)$*, subgame perfect equilibria exist, and in each equilibrium Firm* 2 *searches with* a_1^I *as the cut-off and Firm* 1 *keeps its initial innovation and does not search.*

Proof. See Appendix 3.A.4. □

This proposition shows that in the prescribed scenario Firm 1 is better off giving up its initial innovation if the deadline is long.[17] The intuition is simple. For the deadline being long, Firm 1's expected payoff is low, because its probability of winning is low. By giving up its initial innovation, it makes Firm 2 the head starter, and thus Firm 1 obtains a higher expected payoff than before by committing to searching whereas Firm 2 has no incentive to search. Yet the reasoning for the case in which $c \leq \frac{\lambda}{2}$ differs from that for the case in which $c > \frac{\lambda}{2}$. After Firm 1 discards its initial innovation, Firm 2 turns to the new head start firm. In the former case, Firm 2 would then have no incentive to discard its initial innovation any more as shown in Proposition 3.2, and its dominant strategy in the subgame is not to search whether Firm 1 is to search or not. In the latter case, Firm 2 may have the incentive to discard its initial innovation and search if the deadline is

[17]Discarding a head start is one way to give up one's initial leading position. In reality, a more practical and credible way is to give away the head start innovation. A head start firm could give away it patent for its technology. By doing so, any firm can use this technology for free. That is, every firm's initial state becomes a_1^I. If firms can enter the competition freely, the value of the head start technology is approximately zero to any single firm, because everyone has approximately zero probability to win with this freely obtained innovation. For a head start being in the middle range, the market is not large enough to accommodate two firms to compete. Hence, to model giving away head starts with free entry to the competition, we can study a competition between two firms but with some modified prize allocation rules. Formally, the game proceeds as below.

Model**:

- Stage 1: Firm 1 decides whether to give away its initial innovation.
- Stage 2: Upon observing the stage 1 outcome, Firm 1 and Firm 2 simultaneously start playing the contest as described before.
- If Firm 1 gives away its initial innovation:
  - Both firms' states at time 0 become a_1^I.
  - The prize is retained if no firm is in a state above a_1^I at the deadline T.
  - The firm with a higher state, which is higher than a_1^I, at the deadline wins.
- If Firm 1 retains its initial innovation, the firm with a higher state at the deadline wins.

Suppose $a_1^I \in (a_L^*, F^{-1}(1 - \frac{c}{\lambda}))$. If the deadline is long, in Model** there are two subgame perfect equilibria. In one equilibrium, Firm 1 gives away its initial innovation and searches with a_1^I as the cut-off and Firm 2 does not search. In the other equilibrium, Firm 1 retains its initial innovation and does not search and Firm 2 searches with a_1^I as the cut-off. However, forward induction selects the first equilibrium as the refined equilibrium, because giving away a head start is a credible signal of Firm 1 to commit to search.

long. However, discarding the initial innovation is a credible threat for Firm 1 to deter Firm 2 from doing that.

Remark. *When the deadline is sufficiently long, by giving up the initial innovation, Firm* 1 *makes itself better off but Firm* 2 *worse off. However, if the deadline is not too long, by doing so, Firm* 1 *can benefit both firms. This is because the total expected cost of searching after Firm 1 discards its initial innovation is lower than before and hence there is an increase in the total surplus for the two firms. It is then possible that both firms get a share of the increase in the surplus. We illustrate that in the following example.*

Example 3.1. *Suppose F is the uniform distribution, $c = \frac{1}{3}$, $\lambda = 1$, $a_1^I = \frac{1}{2}$, and $a_2^I = \frac{1}{3}$. If Firm* 1 *discards its initial innovation, then its expected payoff would be $\frac{1}{2}(1 - e^{-\frac{T}{3}})$, and Firm* 2*'s expected payoff would be $e^{-\frac{T}{3}}$; if Firm* 1 *does not discard its initial innovation, then its expected payoff would be $e^{-\frac{T}{2}}$, and Firm* 2*'s expected payoff would be $\frac{1}{3}(1 - e^{-\frac{T}{2}})$.*

Firm 1 *would be better off by discarding its initial innovation if $T > 2.52$. If $T \in (2.52, 3.78)$, by discarding the initial innovation, Firm* 1 *makes both firms better off. If T is larger, then doing so would only make Firm* 2 *worse off.*

The previous result is conditional on Firm 1 having the first move. If Firm 2 has the first move, it may, by discarding its initial innovation, be able to prevent Firm 1 from discarding its own head start and committing to searching. However, if the deadline is not sufficiently long, Firm 1 would still have the incentive to discard its initial innovation.

Proposition 3.8. *Suppose $a_1^I, a_2^I \in (a_L^*, F^{-1}(1 - \frac{c}{\lambda}))$. In Model* with Firm 2 having the first move, there is a $\hat{T}(a_1^I, a_2^I) > 0$ such that*

- *for $T > \hat{T}(a_1^I, a_2^I)$, there is a unique SPE, in which Firm* 2 *discards its initial innovation and searches with a_2^I as the cut-off and Firm* 1 *keeps the initial innovation and does not search;*
- *for $T \in (\breve{T}(a_1^I, a_2^I), \hat{T}(a_1^I, a_2^I))$, there is a unique SPE, in which Firm* 2 *keeps its initial innovation and does not search and Firm* 1 *discards its initial innovation and searches with a_1^I as cut-off;*
- *for $T < \breve{T}(a_1^I, a_2^I)$, subgame perfect equilibria exist, and in each equilibrium Firm* 2 *searches with cut-off a_1^I and Firm* 1 *keeps its initial innovation and does not search.*

Proof. See Appendix 3.A.4. □

In the middle range of the deadline, even though Firm 2 can credibly commit to searching and scare Firm 1 away from competition by discarding its initial innovation, it is not willing to do so, yet Firm 1 would like to discard its initial innovation and commit to searching. This is because Firm 1's initial innovation is of a higher value than Firm 2's. The cut-off value of the deadline at which Firm 2 is indifferent between discarding the initial innovation to commit to searching, and keeping the initial innovation, is higher than that of Firm 1.

3.7.2 Intermediate Information Disclosure

In the software industry, it is common to preannounce with a long lag to launch (Bayus et al., 2001). Many firms do that by describing the expected features or demonstrating prototypes at trade shows. Many other firms publish their findings in a commercial disclosure service, such as Research Disclosure, or in research journals.[18] Suppose there is a regulator who would like to impose an intermediate information disclosure requirement on innovation contests. What are the effects of the requirement on firms' competition strategies and the expected value of the winning innovation. Specifically, suppose there is a time point $t_0 \in (0, T)$ at which both firms have to reveal everything they have, how would firms compete against each other?

When the head start a_1^I is larger the threshold $F^{-1}(1 - \frac{c}{\lambda})$, it is clear that no firm has an incentive to conduct any search. When the head start is below this threshold, if t_0 is very close to T, information revelation has little effect on the firms' strategies. Both firms will play approximately the same actions before time t_0 as they do when there is no revelation requirement. After time t_0, the firm in a higher state at time t_0 stops searching. The other firm searches with this higher state as the cut-off if this higher state is below $F^{-1}(1 - \frac{c}{\lambda})$, and stops searching as well if it is higher than $F^{-1}(1 - \frac{c}{\lambda})$.

Our main finding in this part regards the cases in which the head start is in the middle range and the deadline is sufficiently far from the information revelation point.[19] That is, firms have to reveal their progress at an early stage of a competition.

Proposition 3.9. *Suppose at a time point $t_0 \in (0, T)$ both firms have to reveal their discoveries. For $a_1^I \in (a_L^*, F^{-1}(1 - \frac{c}{\lambda}))$ and $a_2^I < a_1^I$, if $T - t_0$ is sufficiently large, there is a unique subgame perfect equilibrium, in which*

- *Firm 1 searches with $F^{-1}(1 - \frac{c}{\lambda})$ as the cut-off before time t_0 and stops searching from time t_0;*

[18]Over 90% of the world's leading companies have published disclosures in Research Disclosure's pages (see www.researchdisclosure.com).

[19]Generally, for the cases in which $a_1^I < a_L^*$, there are many subgame perfect equilibria, including two equilibria in each of which one firm searches with $F^{-1}(1 - \frac{c}{\lambda})$ as the cut-off between time 0 and time t_0 and the other firm does not.

- *Firm 2 does not search before time t_0 and searches with a_1^I as the cut-off from time t_0 if $a_1^I < F^{-1}(1-\frac{c}{\lambda})$.*

Proof. See Appendix 3.A.4. □

In the proof, we show that between time 0 and time t_0, the dominant action of Firm 2 is not to search, given that the equilibria in the subgames from time t_0 are described as in Theorem 3.2. If Firm 1 is in a state higher than the threshold $F^{-1}(1-\frac{c}{\lambda})$ at time t_0, Firm 2's effort will be futile if it searches before time t_0. If Firm 1 is in a state in between its initial state a_1^I and and the threshold $F^{-1}(1-\frac{c}{\lambda})$ at time t_0, Firm 2 has the chance to get into a state above the threshold $F^{-1}(1-\frac{c}{\lambda})$ and thus a continuation payoff of 1, but this instantaneous benefit only compensates the instantaneous cost of searching. Firm 2 also has the chance to get into a state above that of Firm 1 but below the threshold $F^{-1}(1-\frac{c}{\lambda})$ at time t_0, which results in a continuation payoff of approximately 0 if the deadline is sufficiently long, whereas it obtains a strictly positive payoff if it does not search before time t_0. It is thus not worthwhile for Firm 2 to conduct a search before time t_0. If Firm 2 does not search before time t_0, Firm 1 then has the incentive to conduct a search if the deadline is far from t_0. If it does not search, it obtains a payoff of approximately 0 when $T-t_0$ is sufficiently large. If it conducts a search before time t_0, the benefit from getting into a higher state can compensate the cost.

An early stage revelation requirement therefore hurts the latecomer and benefits the head starter. It gives the head starter a chance to get a high-value innovation so as to deter the latecomer from competition. It also increases the expected value of the winning innovation.

3.8 Discussion: Asymmetric Costs ($a_1^I = a_2^I = 0$, $c_1 < c_2$)

In this section, we show that, compared to the effects of head starts, the effects of cost advantages are simpler. A head start probably discourages a firm from conducting searching and can either discourage its competitor from searching or encourage its competitor to search more actively. In contrast, a cost advantage encourages a firm to search more actively and discourages its opponent.

We now assume that the value of pre-specified prize to Firm i, $i=1,2$, is V_i and that the search cost is for Firm i is C_i per unit of time. However, at each time point Firm i only makes a binary decision on whether to stop searching or to continue searching. Whether it is profitable to continue searching depends on the ratio of $\frac{C_i}{V_i}$ rather than the scale of V_i and C_i. Therefore, we can normalize the valuation of each player to be 1 and the search cost to be $\frac{C_i}{V_i} =: c_i$. W.l.o.g, we assume $c_1 < c_2$. For convenience, we define a function

$$I(a_i|a_j, c_i) := \lambda \int_{a_i}^{\bar{a}} [Z(a|a_j) - Z(a_i|a_j)] dF(a) - c_i,$$

where $Z(a|a_j)$ is defined, in Lemma 3.5 in the appendix, as Firm j's probability of ending up in a state **below** a if it searches with a_j as the cut-off. We emphasize on the most important case, in which both firms' search costs are low.

Proposition 3.10. *For $0 < c_1 < c_2 < \frac{1}{2}\lambda(1-e^{-\lambda T})$, there must exist a unique equilibrium (a_1^*, a_2^*), in which $a_1^* > a_2^* \geq 0$. Specifically,*

1. *if $I\left(0|F^{-1}\left(1-\sqrt{\frac{2c_1}{\lambda(1-e^{-\lambda T})}}\right), c_2\right) > 0$, the unique equilibrium is a pair of cut-off rules (a_1^*, a_2^*), $a_1^* > a_2^* > 0$, that satisfy*

$$\lambda \int_{a_i^*}^{\bar{a}} \left[Z(a|a_j^*, T) - Z(a_i^*|a_j^*, T)\right] dF(a) = c_i;$$

2. *if $I\left(0|F^{-1}\left(1-\sqrt{\frac{2c_1}{\lambda(1-e^{-\lambda T})}}\right), c_2\right) \leq 0$, the unique equilibrium is a pair of cut-off rules $\left(F^{-1}\left(1-\sqrt{\frac{2c_1}{\lambda(1-e^{-\lambda T})}}\right), 0\right)$.*

Proof. See Appendix 3.A.5. □

The existence of equilibrium is proved by using Brouwer's fixed point theorem. As expected, a cost (valuation) advantage would drive a firm to search more actively than its opponent. The following statement shows that while an increase in cost advantage of the firm in advantage would make the firm more active in searching and its opponent less active, a further cost disadvantage of the firm in disadvantage would make both firms less active in searching.

Proposition 3.11. *For $0 < c_1 < c_2 < \frac{1}{2}\lambda(1-e^{-\lambda T})$ and $I\left(0|F^{-1}\left(1-\sqrt{\frac{2c}{\lambda(1-e^{-\lambda T})}}\right), c_2\right) > 0$, in which case there is a unique equilibrium (a_1^*, a_2^*), $a_1^*, a_2^* > 0$,*

1. *for fixed c_2, $\frac{\partial a_1^*}{\partial c_1} < 0$ and $\frac{\partial a_2^*}{\partial c_1} > 0$;*
2. *for fixed c_1, $\frac{\partial a_1^*}{\partial c_2} < 0$ and $\frac{\partial a_2^*}{\partial c_2} < 0$.*

Proof. See Appendix 3.A.5. □

The intuition is simple. When the cost of the firm in advantage decreases, this firm would be more willing to search, while the opponent firm would be discouraged because the marginal increase in the probability of winning from continuing searching in any state is reduced, and

$c_2 \ \backslash \ c_1$	Region 1	Region 2	Region 3
Region 1	(a_1^*, a_2^*)	$\left(F^{-1}\left(1-\sqrt{\frac{2c_1}{\lambda(1-e^{-\lambda T})}}\right), 0\right)$	$(0, -1)$
Region 2	/	$(0, 0)$	$(0, -1)$
Region 3	/	/	$(0, -1), (-1, 0)$

Table 3.1: Equilibria in all non-marginal cases.

therefore the opponent firm would lower its cut-off. When the cost of the firm at a disadvantage increases, the firm would be less willing to search, and the opponent firm would consider it less necessary to search actively because the probability of winning in any state has increased.

A comparison between the equilibrium strategies in this model and that of the benchmark model can be made.

Corollary 3.5. *For* $0 < c_1 < c_2 < \frac{1}{2}\lambda(1-e^{-\lambda T})$ *and* $I\left(0|F^{-1}\left(1-\sqrt{\frac{2c}{\lambda(1-e^{-\lambda T})}}\right), c_2\right) > 0$, *in which case there is a unique equilibrium* (a_1^*, a_2^*), $a_1^* > a_2^* > 0$, a_1^* *and* a_2^* *satisfy*

1. $a_1^* < a^*(c, T)$ *for the corresponding* $c = c_2 > c_1$*;*

2. $a_2^* < a^*(c, T)$ *for the corresponding* $c = c_1 < c_2$*.*

Based on the benchmark model, a cost reduction for Firm 1 will result in both firms searching with cut-offs below the original one; a cost increase for Firm 2 will certainly result in Form 2 searching with a cut-off below the original one.

The equilibrium for the other non-marginal cases (conditional on $c_1 < c_2$), together with the above case, are stated in Table 3.1 without proof. The regions in Table 3.1 are the same as in Figure 3.1. The row (column) number indicates in which region c_1 (c_2) lies, and each element in each cell represents a corresponding equilibrium. For example, the element in the cell at the second row and the second column means that for $c_1, c_2 \in (\frac{1}{2}\lambda(1-e^{-\lambda T}), \frac{1}{2}\lambda(1+e^{-\lambda T}))$, there is a unique equilibrium, in which both firms search with cut-off 0. This shows that Firm 1 is more active in searching than is Firm 2.

Remark. *Similar results can be found in a model with the same search cost but with different arrival rates of innovations for the two firms.*

3.9 Concluding Remarks

In this chapter, we studied the long-run effects of head starts in innovation contests in which each firm decides when to stop a privately observed repeated sampling process before a preset deadline. Unlike an advantage in innovation cost or innovation ability, which encourages a firm to search more actively for innovations and discourages its opponent, a head start has non-monotone effects. The head starter is discouraged from searching if the head start is large, and its strategy remains the same if the head start is small. The latecomer is discouraged from searching if the head start is large but is encouraged to search more actively if it is in the middle range. Our main finding is that, if the head start is in the middle range, in the long run, the head starter is doomed to lose the competition with a payoff of zero and the latecomer will take the entire surplus for the competing firms. As a consequence, our model can exhibit either the "preemption effect" or the "replacement effect", depending on the value of the head start.

Our results have implications on antitrust problems. Market regulators have concerns that the existence of market dominating firms, such as Google, may hinder competitions, and they take measures to curb the monopoly power of these companies. For instance, the European Union voted to split Google into smaller companies.[20] Our results imply that in many cases the positions of the dominant firms are precarious. In the long run, they will be knocked off their perch. These firms' current high positions, in fact, may promote competitions in the long run because they encourage their rivals to exert efforts to innovate and reach high targets. Curbing the power of the current dominating firms may benefit the society and these firms' rivals in the short run, but in the long run it hurts the society because it discourages innovation. However, the the dominating firms' positions are excessively high, which deters new firms from entering the the market, a market regulator could take some actions.

The results have also implications on R&D policies. When selecting an R&D policy, policy makers have to consider both the nature of the R&D projects and the market structure. If the projects are on radical innovations, subsidizing innovation costs effectively increases competition when the market is blank (no advanced substitutive technology exists in the market). However, when there is a current market dominating firm with an existing advanced technology, a subsidy may not be effective. The dominating firm has no incentive to innovate, and the latecomer, even if it is subsidized, will not innovate more actively.

In our model we have only one head starter and one latecomer. The model can be extended to

[20] "Google break-up plan emerges from Brussels," Financial Times, November 21, 2014.

include more than two firms, and similar results still hold. One extension is to study the designing problem in our framework. For example, one question is how to set the deadline. If the designer is impatient, she may want to directly take the head starter's initial innovation without holding a contest; if she is patient, she may set a long deadline in order to obtain a better innovation. Some other extensions include: to consider a model with a stochastic number of firms; to consider a model with cumulative scores with or without regret instead of a model with repeated sampling.

3.A Appendix

3.A.1 Preliminaries

To justify Assumption 3.1, we show in the following that, for any given strategy played by a firm's opponent, there is a constant cut-off rule being the firm's best response. If the cut-off value is above zero, it is actually the unique best response strategy, ignoring elements associated with zero probability events. We argue only for the case that both firms' initial states are 0. The arguments for the other cases are similar and thus are omitted.

Suppose $a_1^I = a_2^I = 0$. For a given strategy played by Firm j, we say at time t

$$\underline{a}_i^t := \inf\{\tilde{a} \in A | \text{Firm } i \text{ weakly prefers stopping to continuing searching in state } \tilde{a}\}$$

is Firm i's **lower optimal cut-off** and

$$\bar{a}_i^t := \inf\{\tilde{a} \in A | \text{Firm } i \text{ strictly prefers stopping to continuing searching in state } \tilde{a}\}$$

is Firm i's **upper optimal cut-off** .

Lemma 3.4. *Suppose $a_1^I = a_2^I = 0$. For any fixed strategy played by Firm j, Firm i's best response belongs to one of the three cases.*

i. Not to search: $\bar{a}_i^t = \underline{a}_i^t = -1$ for all $t \in [0, T]$,

ii. Search with a constant cut-off rule $\hat{a}_i \geq 0$: $\bar{a}_i^t = \underline{a}_i^t = \hat{a}_i \geq 0$ for all $t \in [0, T]$.

iii. Both not to search and search until being in a state above 0: $\bar{a}_i^t = 0$ and $\underline{a}_i^t = -1$ for all $t \in [0, T)$.

Proof of Lemma 3.4. Fix a strategy of Firm j. Let $P(a)$ denote the probability of Firm j ending up in a state **below** a at time T. $P(a)$ is either constant in a or strictly increasing in a. It is a constant if and only if Firm j does not search.[21] If this is the case, Firm i's best response

[21] More generally, it is constant if and only if the opponent firm conducts search with a measure 0 over $[0, T]$.

is to continue searching with a fixed cut-off $\hat{a}_i^t = \bar{a}^{t_i} = \underline{a}^{t_i} = 0$ for all t. In the following, we study the case in which $P(a)$ is strictly increasing in a.

Step 1. We argue that, given a fixed strategy played by Firm j, Firm i's best response is a (potentially history-dependent) cut-off rule. Suppose at time t Firm i is in a state $\tilde{a} \in [0,1]$. If it is strictly marginally profitable to stop (continue) searching at t, then it is also strictly marginally profitable to continue searching if it is in a state higher (lower) than $\tilde{a}$. Let the upper and lower optimal cut-offs at time t be $\bar{a}_i^t$ and $\underline{a}_i^t$, respectively, as defined previously.

Step 2. We show that $\{\bar{a}_i^t\}_{t=0}^T$ and $\{\underline{a}_i^t\}_{t=0}^T$ should be history-independent. We use a discrete version to approximate the continuous version. Take any $\tilde{t} \in [0,T)$. Let $\{t_l\}_{l=0}^k$, where $t_l - t_{l-1} = \frac{T-\tilde{t}}{k} =: \delta$ for $l = 1, ..., k$, be a partition of the interval $[\tilde{t}, T]$. Suppose Firm i can only make decisions at $\{t_l\}_{l=0}^k$ in the interval $[\tilde{t}, T]$. Let $\{\bar{a}^{t_l}\}_{l=0}^{k-1}$ and $\{\underline{a}^{t_l}\}_{l=0}^{k-1}$ be the corresponding upper and lower optimal cut-offs, respectively, and $G^\delta(a)$ be Firm i's probability of discovering **no** innovation with a value **above** a in an interval δ.

At t_{k-1}, for Firm i in a state a, if it stops searching, the expected payoff is $P(a)$; if it continues searching, the expected payoff is

$$\begin{aligned} &G^\delta(a)P(a) + \int_a^1 P(\tilde{a})dG^\delta(\tilde{a}) - \delta c_i \\ =&P(a) + \int_a^1 [P(\tilde{a}) - P(a)]dG^\delta(\tilde{a}) - \delta c_i. \end{aligned}$$

The firm strictly prefers continuing searching if and only if searching in the last period strictly increases its expected payoff,

$$e^\delta(a) := \int_a^1 [P(\tilde{a}) - P(a)]dG^\delta(\tilde{a}) - \delta c_i > 0.$$

$e^\delta(a)$ strictly decreases in a and $e^\delta(1) \le 0$. Because $e^\delta(0)$ can be either negative or positive, we have to discuss several cases.

Case 1. If $e^\delta(0) < 0$, Firm i is strictly better off stopping searching in any state $a \in [0,1]$. Thus, $\bar{a}^{t_{k-1}} = \underline{a}^{t_{k-1}} = -1$.

Case 2. If $e^\delta(0) = 0$, Firm i is indifferent between stopping searching and continuing searching with 0 as the cut-off, if it is in state 0; strictly prefers stopping searching, if it is in any state above 0. Then $\bar{a}^{t_{k-1}} = 0$ and $\underline{a}^{t_{k-1}} = -1$.

Case 3. If $e^\delta(0) > 0 \ge \lim_{a\to 0} e^\delta(a)$, Firm i is strictly better off continuing searching in state 0, but stopping searching once it is in a state above 0. Thus, $\bar{a}^{t_{k-1}} = \underline{a}^{t_{k-1}} = 0$.

Case 4. If $\lim_{a\to 0} e^\delta(a) > 0$, then Firm i's is strictly better off stopping searching if it is in a state above $\hat{a}^{t_{k-1}}$ and continuing searching if it is in a state below $\hat{a}^{t_{k-1}}$, where the optimal cut-off

$\hat{a}^{t_{k-1}} > 0$ is the unique value of a that satisfies,

$$\int_a^1 [P(\tilde{a}) - P(a)]dG^\delta(\tilde{a}) - \delta c = 0.$$

Thus, in this case $\bar{a}^{t_{k-1}} = \underline{a}^{t_{k-1}} = \hat{a}^{t_{k-1}}$.

Hence, the continuation payoff at $t_{k-1} \geq 0$ for Firm i in a state $a \in [0, 1]$ is

$$\omega(a) = \begin{cases} P(a) + \int_a^1 [P(\tilde{a}) - P(a)]dG^\delta(\tilde{a}) - \delta c & for\ a < \underline{a}^{t_{k-1}} \\ P(a) & for\ a \geq \underline{a}^{t_{k-1}}. \end{cases}$$

Then, we look at the time point t_{k-2}. In the following, we argue that $\bar{a}^{t_{k-2}} = \bar{a}^{t_{k-1}}$. The argument for $\underline{a}^{t_{k-2}} = \underline{a}^{t_{k-1}}$ is very similar and thus is omitted.

First, we show that $\bar{a}^{t_{k-2}} \leq \bar{a}^{t_{k-1}}$. Suppose $\bar{a}^{t_{k-2}} > \bar{a}^{t_{k-1}}$. Suppose Firm i is in state $\bar{a}^{t_{k-2}}$ at time t_{k-2}. Suppose Firm i searches between t_{k-2} and t_{k-1}. If it does not discover any innovation with a value higher than $\bar{a}^{t_{k-2}}$, then at the end of this period it stops searching and takes $\bar{a}^{t_{k-2}}$. However, $\bar{a}^{t_{k-2}} > \bar{a}^{t_{k-1}}$ implies

$$\begin{aligned} 0 &= \int_{\bar{a}^{t_{k-2}}}^1 [P(\tilde{a}) - P(\bar{a}^{t_{k-2}})]dG^\delta(\tilde{a}) - \delta c_i \\ &< \int_{\bar{a}^{t_{k-1}}}^1 [P(\tilde{a}) - P(\bar{a}^{t_{k-1}})]dG^\delta(\tilde{a}) - \delta c \leq 0. \end{aligned}$$

The search cost is not compensated by the increase in the probability of winning from searching between t_{k-2} and t_{k-1}, and thus the firm strictly prefers stopping searching to continuing searching at time t_{k-2}, which contradicts the assumption that $\bar{a}^{t_{k-2}}$ is the upper optimal cut-off. Hence, it must be the case that $\bar{a}^{t_{k-2}} \leq \bar{a}^{t_{k-1}}$.

Next, we show that $\bar{a}^{t_{k-2}} = \bar{a}^{t_{k-1}}$.

In *Case 1*, it is straightforward that Firm i strictly prefers stopping searching at t_{k-2}, since it is for sure not going to search between t_{k-1} and t_k. Hence, Firm i stops searching before t_{k-1}, and $\bar{a}^{t_{k-2}} = \bar{a}^{t_{k-1}} = \underline{a}^{t_{k-2}} = \underline{a}^{t_{k-1}} = -1$.

For $\bar{a}^{t_{k-1}} \geq 0$, we prove by contradiction that $\bar{a}^{t_{k-2}} < \bar{a}^{t_{k-1}}$ is not possible. Suppose the inequality holds. If Firm i stops searching at t_{k-2}, it would choose to continue searching at t_{k-1}, and its expected continuation payoff at t_{k-2} is $\omega(\bar{a}^{t_{k-2}})$. If the firm continues searching, its expected continuation payoff is

$$\omega(\bar{a}^{t_{k-2}}) + \int_{\bar{a}^{t_{k-2}}}^1 [\omega(a) - \omega(\bar{a}^{t_{k-2}})]dG^\delta(a) - \delta c. \tag{3.6}$$

In *Cases 2 and 3*, $\bar{a}^{t_{k-1}} = 0$ implies $\bar{a}^{t_{k-2}} = -1$. Then,

$$\begin{aligned}
&\int_{\bar{a}^{t_{k-2}}}^{1} [\omega(a) - \omega(\bar{a}^{t_{k-2}})] dG^{\delta}(a) - \delta c_i \\
=&\int_{-1}^{1} [P(a)] dG^{\delta}(a) - \delta c_i \\
=&e^{\delta}(-1) \\
\geq&0
\end{aligned}$$

which means that Firm i in state 0 is weakly better off continuing searching between t_{k-2} and t_{k-1}, which implies that $\bar{a}^{t_{k-2}} \geq 0$, resulting in a contradiction.

For *Case 4*, in which $\bar{a}^{t_{k-1}} > 0$, we have in (3.6)

$$\begin{aligned}
&\int_{\bar{a}^{t_{k-2}}}^{1} [\omega(a) - \omega(\bar{a}^{t_{k-2}})] dG^{\delta}(a) \\
=&\int_{\bar{a}^{t_{k-2}}}^{\bar{a}^{t_{k-1}}} \left[\left(P(a) + \int_{a}^{1} [P(\tilde{a}) - P(a)] dG^{\delta}(\tilde{a})\right) - \left(P(\bar{a}^{t_{k-2}}) + \int_{\bar{a}^{t_{k-2}}}^{1} [P(\tilde{a}) - P(\bar{a}^{t_{k-2}})] dG^{\delta}(\tilde{a})\right)\right] dG^{\delta}(a) \\
&+ \int_{\bar{a}^{t_{k-1}}}^{1} \left[P(a) - \left(P(\bar{a}^{t_{k-2}}) + \int_{\bar{a}^{t_{k-2}}}^{1} [P(\tilde{a}) - P(\bar{a}^{t_{k-2}})] dG^{\delta}(\tilde{a})\right)\right] dG^{\delta}(a) \\
=&\int_{\bar{a}^{t_{k-2}}}^{1} \left[P(a) - P(\bar{a}^{t_{k-2}})\right] dG^{\delta}(a) + \int_{\bar{a}^{t_{k-2}}}^{\bar{a}^{t_{k-1}}} \left[\int_{a}^{1} [P(\tilde{a}) - P(a)] dG^{\delta}(\tilde{a})\right] dG^{\delta}(a) \\
&- \int_{\bar{a}^{t_{k-2}}}^{1} \left[\int_{\bar{a}^{t_{k-2}}}^{1} [P(\tilde{a}) - P(\bar{a}^{t_{k-2}})] dG^{\delta}(\tilde{a})\right] dG^{\delta}(a) \\
=&G^{\delta}(\bar{a}^{t_{k-2}}) \int_{\bar{a}^{t_{k-2}}}^{1} \left[P(a) - P(\bar{a}^{t_{k-2}})\right] dG^{\delta}(a) + \int_{\bar{a}^{t_{k-2}}}^{\bar{a}^{t_{k-1}}} \left[\int_{a}^{1} [P(\tilde{a}) - P(a)] dG^{\delta}(\tilde{a})\right] dG^{\delta}(a) \\
>&0.
\end{aligned}$$

Hence, at t_{k-2} Firm i would strictly prefer continuing searching, which again contradicts the assumption that $\bar{a}^{t_{k-2}}$ is the upper optimal cut-off. Consequently, $\bar{a}^{t_{k-2}} = \bar{a}^{t_{k-1}}$.

By backward induction from t_{k-1} to t_0, we have $\bar{a}^{t_0} = \bar{a}^{t_{k-1}}$. Taking the limit we get

$$\bar{a}^{t} = \lim_{\delta \to 0} \bar{a}^{T-\delta} =: \bar{a} \quad \text{for all } t \in [0, T).$$

Similarly,

$$\underline{a}^{t} = \lim_{\delta \to 0} \underline{a}^{T-\delta} =: \underline{a} \quad \text{for all } t \in [0, T).$$

In addition, $\bar{a} \neq \underline{a}$ when and only when $\bar{a} = 0$ and $\underline{a} = -1$.

As a consequence, Firm i's best response is not to search, if $\bar{a} = \underline{a} = -1$; to continue searching if it is in a state below $\bar{a}$ and to stop searching once the firm in a state above $\bar{a}$, if $\bar{a} = \underline{a} \geq 0$. $\square$

In brief, the above property is proved by backward induction. Take Case [ii] for example. If at the last moment a firm is indifferent between continuing and stopping searching when it is in a certain state, which means the increase in the probability of winning from continuing searching equals the cost of searching, and therefore there is no gain from searching. Immediately before the last moment the firm should also be indifferent between continuing searching and not given the same state. This is because, if the firm reaches a higher state from continuing searching, it weakly prefers not to search at the last moment, and thus the increase in the probability of winning from continuing searching at this moment equals the cost of searching as well. By induction, the firm should be indifferent between continuing and stopping searching in the same state from the very beginning.

In Case [iii], Firm i generally has uncountably many best response strategies. By Assumption 3.1, we rule out most strategies and consider only two natural strategies: not to search at all and to search with 0 as the cut-off.

Lemma 3.5. *Suppose a firm's initial state is* 0*, and she searches with a cut-off* $\hat{a} \geq 0$*. Then, the firm's probability of ending up in a state* ***below*** $a \in [0,1]$ *at time* T *is*

$$Z(a|\hat{a},T) = \begin{cases} 0 & if\ a = 0 \\ e^{-\lambda T[1-F(a)]} & if\ 0 < a \leq \hat{a} \\ e^{-\lambda T[1-F(\hat{a})]} + \left[1 - e^{-\lambda T[1-F(\hat{a})]}\right] \frac{F(a)-F(\hat{a})}{1-F(\hat{a})} & if\ a > \hat{a}. \end{cases}$$

$1 - e^{-\lambda T[1-F(\hat{a})]}$ is the probability that the firm stops searching before time T, and $\frac{F(a)-F(\hat{a})}{1-F(\hat{a})}$ is the conditional probability that the innovation above the threshold the firm discovers is in between $\hat{a}$ and a.

Proof of Lemma 3.5. For $a = 0$, it is clear that $Z(a|\hat{a},T) = 0$.

For $0 < a \leq \hat{a}$,

$$Z(a|\hat{a},T) = \sum_{l=0}^{\infty} \frac{e^{-\lambda T}(\lambda T)^l}{l!} F^l(a) = e^{-\lambda T[1-F(a)]}.$$

For $a > \hat{a}$, we approximate it by a discrete time model. Let $\{t_l\}_{l=0}^{k}$, where $0 = t_0 < t_1 < ... < t_k = T$, be a partition of the interval $[0,T]$, and let $\delta_l = t_l - t_{l-1}$ for $l = 1,2,...,k$. Define π as

$$\parallel \pi \parallel = \max_{1 \leq l \leq k} |\delta_l|.$$

Then,

$$\begin{aligned}
Z(a|\hat{a},T) =& Z(\hat{a}|\hat{a},T) + \lim_{\|\pi\|\to 0} \sum_{l=1}^{k} Z(a|\hat{a},t_{l-1}) \left[\sum_{n=1}^{\infty} \frac{e^{-\lambda\delta_l}(\lambda\delta_l)^n}{n!} [F^n(a) - F^n(\hat{a})] \right] \\
=& Z(\hat{a}|\hat{a},T) + \lim_{\|\pi\|\to 0} \sum_{l=1}^{k} e^{-\lambda t_{l-1}[1-F(\hat{a})]} \lambda e^{-\lambda\delta_l} \big([F(a) - F(\hat{a})] + O(\delta_l) \big) \delta_l \\
=& Z(\hat{a}|\hat{a},T) + \int_0^T \lambda e^{-\lambda t[1-F(\hat{a})]} [F(a) - F(\hat{a})] dt \\
=& Z(\hat{a}|\hat{a},T) + \left[1 - e^{-\lambda T[1-F(\hat{a})]} \right] \frac{F(a) - F(\hat{a})}{1 - F(\hat{a})},
\end{aligned}$$

where the second term on the right hand side of each equality is the firm's probability of ending up in a state between $\hat{a}$ and a. The term $Z(\hat{a}|\hat{a},t_n)$ used here is a convenient approximation when δ_l is small. The second equality is derived from the fact that

$$\sum_{n=2}^{\infty} \frac{(\lambda\delta_l)^n}{n!} [F^n(a) - F^n(a^*)] < \frac{\lambda^2\delta_l^2}{2(1-\lambda\delta_l)} = o(\delta_l).$$

□

Lemma 3.6. *Given $a > a'$, $Z(a|\tilde{a},T) - Z(a'|\tilde{a},T)$*

1. *is constant in $\tilde{a}$ for $\tilde{a} \geq a$;*
2. *strictly decreases in $\tilde{a}$ for $\tilde{a} \in (a', a)$;*
3. *strictly increases in $\tilde{a}$ for $\tilde{a} \leq a'$.*

This single-peaked property says that the probability of ending up in a state between a' and a is maximized if a firm chooses strategy a'.

Proof of Lemma 3.6. First, we show that $\frac{1-e^{-\lambda Tx}}{x}$ strictly decreases in x over $(0, 1]$ as follows. Define $s := \lambda T$ and take x_1, x_2, $0 < x_1 < x_2 \leq 1$, we have

$$\frac{1 - e^{-sx_1}}{x_1} > \frac{1 - e^{-sx_2}}{x_2},$$

implied by

$$\begin{aligned}
&\frac{\partial (1 - e^{-sx_1})x_2 - (1 - e^{-sx_2})x_1}{\partial s} = x_1x_2(e^{-sx_1} - e^{-sx_2}) \geq 0 \quad (= 0 \text{ iff } s = 0) \text{ and} \\
&(1 - e^{-sx_1})x_2 - (1 - e^{-sx_2})x_1 = 0 \quad \text{for} \quad s = 0.
\end{aligned}$$

Next, define $x := 1 - F(a)$, $x' := 1 - F(a')$, and $\tilde{x} := 1 - F(\tilde{a})$. We have

$$Z(a|\tilde{a},T) - Z(a'|\tilde{a},T) = \begin{cases} e^{-\lambda T x} - e^{-\lambda T x'} & \text{for } \tilde{a} \geq a \\ (1 - e^{-\lambda T x'}) - (1 - e^{-\lambda T \tilde{x}})\frac{x}{\tilde{x}} & \text{for } \tilde{a} \in (a', a) \\ (1 - e^{-\lambda T \tilde{x}})\frac{x' - x}{\tilde{x}} & \text{for } \tilde{a} \leq a'. \end{cases}$$

It is independent of $\tilde{a}$ for $\tilde{a} \geq a$, strictly increasing in $\tilde{x}$ and thus strictly decreasing in $\tilde{a}$ for $\tilde{a} \leq a'$, and strictly decreasing in $\tilde{x}$ and thus strictly increasing in $\tilde{a}$ for $\tilde{a} \leq a'$. □

Lemma 3.7. *Suppose Firm j with initial state 0 plays a strategy $\hat{a}_j$. Then, the **instantaneous gain** on payoff from searching for Firm i in a state a is*

$$\lambda \int_{a_i}^{1} [Z(a|\hat{a}_j,T) - Z(a_i|\hat{a}_j,T)] dF(a) - c.$$

Proof of Lemma 3.7. For convenience, denote $H(a)$ as $Z(a|\hat{a}_j,T)$ for short. The instantaneous gain from searching for Firm i in a state a_i is

$$\begin{aligned}
&\lim_{\delta \to 0} \frac{\left(e^{-\lambda\delta} H(a_i) + \lambda\delta e^{-\lambda\delta}\left[\int_{a_i}^{1} H(a)dF(a) + F(a_i)H(a_i)\right] + o(\delta) - \delta c\right) - H(a_i)}{\delta} \\
=&\lim_{\delta \to 0} \frac{-(1 - e^{-\lambda\delta})H(a_i) + \lambda\delta e^{-\lambda\delta}\left[\int_{a_i}^{1} H(a)dF(a) + F(a_i)H(a_i)\right] + o(\delta) - \delta c}{\delta} \\
=&\lim_{\delta \to 0} \frac{-\lambda\delta e^{-\lambda\delta} H(a_i) + \lambda\delta e^{-\lambda\delta}\left[\int_{a_i}^{1} H(a)dF(a) + F(a_i)H(a_i)\right] + o(\delta) - \delta c}{\delta} \\
=& - \lambda H(a_i) + \lambda \left[\int_{a_i}^{1} H(a)dF(a) + F(a_i)H(a_i)\right] - c \\
=&\lambda \int_{a_i}^{1} \left[Z(a|\hat{a}_j,T) - Z(a_i|\hat{a}_j,T)\right] dF(a) - c.
\end{aligned}$$

□

3.A.2 Proofs for the Benchmark Case

Proof of Theorem 3.1. We prove the theorem case by case.

Case[i]. When Firm i does not search, Firm j's best response is to search with cut-off 0. For $\frac{1}{2}\lambda(1 + e^{-\lambda T}) \leq c$, when Firm j searches with any cut-off $a_j \geq 0$, Firm i's best response is not to search, since the instantaneous gain from searching for Firm i in state 0 is

$$\begin{aligned}
&\lambda \int_0^1 Z(a|a_j, T) dF(a) - c \\
\leq &\lambda \int_0^1 Z(a|0, T) dF(a) - c \\
= &\lambda \int_0^1 [e^{-\lambda T} + (1 - e^{-\lambda T}) F(a)] dF(a) - c \\
= &\lambda \left[e^{-\lambda T} + \frac{1}{2}(1 - e^{-\lambda T}) \right] - c \\
= &\frac{1}{2}\lambda(1 + e^{-\lambda T}) - c \\
\leq &0 \ \ (= 0 \text{ iff } \frac{1}{2}\lambda(1 + e^{-\lambda T}) = c),
\end{aligned}$$

where the first inequality follows from Lemma 3.6. Hence, there are two pure strategy equilibria, in each of which one firm does not search and the other firm searches with 0 as the cut-off, and if $\frac{1}{2}\lambda(1 + e^{-\lambda T}) = c$ there is also an equilibrium in which both firms search with 0 as the cut-off.

Case [ii]. First, we show that any strategy with a cut-off value higher than zero is a dominated strategy. When Firm j does not search, Firm i prefers searching with 0 as the cut-off to any other strategy. Suppose Firm j searches with $\hat{a}_j \geq 0$ as the cut-off. The instantaneous gain from searching for Firm i in a state $a_i > 0$ is

$$\begin{aligned}
&\lambda \int_{a_i}^1 [Z(a|\hat{a}_j, T) - Z(a_i|\hat{a}_j, T)] dF(a) - c \\
\leq &\lambda \int_{a_i}^1 [Z(a|a_i, T) - Z(a_i|a_i, T)] dF(a) - c \\
= &\frac{1}{2}\lambda(1 - e^{-\lambda T})[1 - F(a_i)]^2 - c \\
< &0,
\end{aligned}$$

where the first inequality follows from Lemma 3.6. Hence, once Firm i is in a state above 0, it has no incentive to continue searching any more. In this case, Firm i prefers either not to conduct any search or to search with 0 as the cut-off to any strategy with a cut-off value higher than zero.

Second, we show that the prescribed strategy profile is the unique equilibrium. It is sufficient to show that searching with 0 as the cut-off is the best response to searching with 0 as the cut-off. Suppose Firm j searches with 0 as the cut-off, the instantaneous gain from searching for Firm i

in state $a = 0$ is

$$\begin{aligned}&\lambda \int_0^1 Z(a|0,T)dF(a) - c\\=&\frac{1}{2}\lambda(1+e^{-\lambda T}) - c\\>&0.\end{aligned}$$

That is, Firm i is strictly better off continuing searching if it is in state 0, and strictly better off stopping searching once it is in a state above 0. Hence, the prescribed strategy profile is the unique equilibrium.

Case [iii]. First, we prove that among the strategy profiles in which each firm searches with a cut-off higher than 0, the prescribed symmetric strategy profile is the unique equilibrium. Suppose a pair of cut-off rules (a_1^*, a_2^*), in which $a_1^*, a_2^* > 0$, is an equilibrium, then Firm i in state a_i^* is indifferent between continuing searching and not. That is, by Lemma 3.7, we have

$$\lambda \int_{a_i^*}^1 \left[Z(a|a_j^*, T) - Z(a_i^*|a_j^*, T)\right] dF(a) - c = 0. \tag{3.7}$$

Suppose $a_1^* \neq a_2^*$. W.l.o.g., we assume $a_1^* < a_2^*$. Then,

$$\begin{aligned}c =&\lambda \int_{a_1^*}^1 [Z(a|a_2^*, T) - Z(a_1^*|a_2^*, T)]dF(a)\\>&\lambda \int_{a_2^*}^1 [Z(a|a_2^*, T) - Z(a_2^*|a_2^*, T)]dF(a)\\>&\lambda \int_{a_2^*}^1 [Z(a|a_1^*, T) - Z(a_2^*|a_1^*, T)]dF(a) = c\end{aligned}$$

resulting in a contradiction. Hence, it must be the case that $a_1^* = a_2^*$.

Next, we show the existence of equilibrium by deriving the unique equilibrium cut-off value $a^* := a_1^* = a_2^*$ explicitly. Applying Lemma 3.5 to (3.7), we have

$$\begin{aligned}&\lambda \int_{a^*}^1 \left[1 - e^{-\lambda T[1-F(a^*)]}\right] \frac{F(a) - F(a^*)}{1 - F(a^*)} dF(a) = c\\\Leftrightarrow&\frac{1}{2}[1 - F(a^*)]\left[1 - e^{-\lambda T[1-F(a^*)]}\right] = \frac{c}{\lambda}.\end{aligned} \tag{3.8}$$

The existence of a solution is ensured by the intermediate value theorem: when $F(a^*) = 1$, the term on the left hand side of (3.8) equals to 0, smaller than $\frac{c}{\lambda}$; when $F(a^*) = 0$, it equals to $\frac{1-e^{-\lambda T}}{2}$, larger than or equals to $\frac{c}{\lambda}$. The uniqueness of the solution is insured by that the term on the left hand side of the above equality is strictly decreasing in a^*.

Second, we show that there is no equilibrium in which one firm searches with 0 as the cut-off. Suppose Firm j searches with 0 as the cut-off. The instantaneous gain from searching for Firm i in a state $a_i > 0$ is

$$\begin{aligned}
&\lambda \int_{a_i}^{1} [Z(a|0,T) - Z(a_i|0,T)]dF(a) - c \\
=&\lambda \int_{a_i}^{1} (1 - e^{-\lambda T})[F(a) - F(a_i)]dF(a) - c \\
=&\frac{1}{2}\lambda(1 - e^{-\lambda T})[1 - F(a_i)]^2 - c, \qquad (3.9)
\end{aligned}$$

which is positive when $a_i = 0$ and negative when $a_i = 1$. By the intermediate value theorem, there must be a value $\hat{a}_i > 0$ such that (3.9) equals 0 when $a_i = \hat{a}_i$. Hence, Firm i's best response is to search with $\hat{a}_i$ as the cut-off. However, if Firm i searches with $\hat{a}_i$ as the cut-off, it is not Firm j's best response to search with 0 as the cut-off, because

$$\begin{aligned}
0 =& \int_{\hat{a}_j}^{1} [Z(a|0,T) - Z(\hat{a}_j|0,T)]dF(a) - c \\
<& \int_{\hat{a}_j}^{1} [Z(a|\hat{a}_j,T) - Z(\hat{a}_j|\hat{a}_j,T)]dF(a) - c \\
<& \int_{0}^{1} [Z(a|\hat{a}_j,T) - Z(0|\hat{a}_j,T)]dF(a) - c,
\end{aligned}$$

which means that Firm j strictly prefers continuing searching when it is in a state slightly above 0.

Last, we show that there is no equilibrium in which one firm does not search. Suppose Firm j does not search. Firm i's best response is to search with 0 as the cut-off. However, Firm j then strictly prefers searching when it is in state 0, since the instantaneous gain from searching for the firm in state 0 is again

$$\lambda \int_{0}^{1} Z(a|a_i,T)dF(a) - c > 0.$$

□

Proof of Lemma 3.2. The expected total cost of a firm who searches with a cut-off $a \geq 0$ is

$$\begin{aligned}
&c\left[\int_{0}^{T} \frac{\partial(1 - Z(a|a,t))}{\partial t} t dt + TZ(a|a,T)\right] \\
=&(1 - e^{-\lambda T[1-F(a)]})\frac{c}{\lambda[1 - F(a)]}, \qquad (3.10)
\end{aligned}$$

which strictly increases in a. In Regions 2 and 3, in equilibrium, the probability of winning for each firm is

$$\begin{aligned}&\frac{1}{2}[1 - Z^2(0|a^*(c,T),T)]\\=&\frac{1}{2}(1 - e^{-2\lambda T}),\end{aligned} \tag{3.11}$$

and thus the expected payoff to each firm is the difference between the expected probability of winning (3.11) and the expected search cost (3.10), setting a to be $a^*(a,T)$:

$$\frac{1}{2}(1 - e^{-2\lambda T}) - (1 - e^{-\lambda T[1-F(a^*(c,T))]})\frac{c}{\lambda[1 - F(a^*(c,T))]}. \tag{3.12}$$

The limit of (3.12) as T approaches infinity is 0.

□

3.A.3 Proofs for the Head Start Case

First, we state two crucial lemmas for the whole section.

Lemma 3.8.

1. *For $a_1^I > a^*(c,T)$, not to search is Firm 1's strictly dominant strategy.*
2. *For $a_1^I = a^*(c,T)$, not to search is Firm 1's weakly dominant strategy. If Firm 2 searches with cut-off a_1^I, Firm 1 is indifferent between not to search and search with a_1^I as the cut-off; Otherwise, Firm 1 strictly prefers not to search.*

Proof of Lemma 3.8. Suppose Firm 2 does not search, Firm 1's best response is not to search. Suppose Firm 2 searches with cut-off $a_2 \geq a_1^I$. If Firm 1 searches with cut-off $a_1 \geq a_1^I$, following from Lemma 3.6, the instantaneous gain from searching for Firm 1 in any state $a_1 \geq a_1^I \geq a^*(c,T)$ is

$$\begin{aligned}&\lambda\int_{a_1}^{1}[Z(a|a_2,T) - Z(a_1|a_2,T)]dF(a) - c\\\leq&\lambda\int_{a^*(c,T)}^{1}[Z(a|a^*(c,T),T) - Z(a^*(c,T)|a^*(c,T),T)]dF(a) - c,\end{aligned} \tag{3.13}$$

where equality holds if and only if $a_1 = a_2 = a^*(c,T)$. The right hand side of inequality (3.13) is less than or equal to 0 (it equals to 0 iff $c \geq \frac{1}{2}\lambda[1 - e^{-\lambda T}]$). Hence, the desired results follow. □

Lemma 3.9.

1. *For $a_1^I > F^{-1}(1 - \frac{c}{\lambda})$, not to search is Firm 2's strictly dominant strategy.*

2. *For* $a_1^I = F^{-1}(1 - \frac{c}{\lambda})$, *not to search is Firm* 2*'s weakly dominant strategy. If Firm* 1 *does not search, Firm* 2 *is indifferent between not to search and search with any* $\hat{a}_2 \in [a_2^I, a_1^I]$ *as the cut-off. If Firm* 1 *searches, Firm* 2*'s strictly prefers not to search.*

Proof of Lemma 3.9. If Firm 1 does not search, the instantaneous gain from searching for Firm 2 in a state $a_2 \leq a_1^I$ is

$$\lim_{\delta \to 0} \frac{\lambda \delta e^{-\lambda \delta}[1 - F(a_1^I)] + o(\delta) - c\delta}{\delta} = \lambda[1 - F(a_1^I)] - c \begin{cases} < 0 & \text{in Case [1]} \\ = 0 & \text{in Case [2].} \end{cases}$$

If Firm 1 searches, Firm 2's instantaneous gain is even lower. Hence, the desired results follow. □

Proof of Theorem 3.2. [1],[2], and [3] directly follow from Lemmas 3.8 and 3.9. We only need to prove [4] and [5] in the following.

[4]. Following Lemma 3.8, if Firm 2 searches with cut-off a_1^I, Firm 1 has two best responses: not to search and search with cut-off a_1^I. If Firm 1 searches with cut-off a_1^I, the instantaneous gain from searching for Firm 2 is

$$\begin{aligned}
&\lambda \int_{a_1^I}^1 Z(a|a_1^I, T) dF(a) - c \\
=&\lambda \int_{a_1^I}^1 \left[e^{-\lambda T[1-F(a_1^I)]} + (1 - e^{-\lambda T[1-F(a_1^I)]}) \frac{F(a) - F(a_1^I)}{1 - F(a_1^I)} \right] dF(a) - c \\
=&\frac{1}{2}\lambda(1 + e^{-\lambda T[1-F(a_1^I)]})[1 - F(a_1^I)] - c \\
>&\lambda[1 - F(a_1^I)] - c \\
>&0
\end{aligned}$$

if it is in a state $a_2 < a_1^I = a^*(c, T)$; it is

$$\lambda \int_{a_2}^1 [Z(a|a_1^I, T) - Z(a_2|a_1^I, T)] dF(a) - c < 0$$

if it is in a state $a_2 > a_1^I = a^*(c, T)$. Hence, the two prescribed strategy profiles are equilibria.

[5]. First, there is no equilibrium in which either firm does not search. If Firm 2 does not search, Firm 1's best response is not to search. However, if Firm 1 does not search, Firm 2's best response is to search with cut-off a_1^I rather than not to search. If Firm 2 searches with cut-off a_1^I, then not to search is not Firm 1's best response, because the instantaneous gain from searching

for Firm 1 in state a_1^I is

$$\begin{aligned}&\lambda\int_{a_1^I}^1[Z(a|a_1^I,T)-Z(a_1^I|a_1^I,T)]dF(a)-c\\=&\frac{1}{2}\lambda(1-e^{-\lambda T[1-F(a_1^I)]})[1-F(a_1^I)]-c\\>&0,\end{aligned}$$

where inequality holds because $a_1^I < a^*(c,T)$.

Next, we argue that there is no equilibrium in which either firm searches with cut-off a_1^I. Suppose Firm i searches with cut-off a_1^I. Firm j's best response is to search with a cut-off $\hat{a}_j \in [a_1^I, a^*(c,T))$. This is because the instantaneous gain from searching for Firm j in a state $a' \geq a_1^I$ is

$$\lambda\int_{a'}^1 p_2[Z(a|a_1^I,T)-Z(a'|a_1^I,T)]dF(a)-c. \tag{3.14}$$

(3.14) is larger than 0 when $a' = a_1^I$. It is less than 0 if $a' = a^*(c,T)$, because by Lemma 3.6 we have

$$\begin{aligned}&\lambda\int_{a^*}^1 p_2[Z(a|a_1^I,T)-Z(a^*|a_1^I,T)]dF(a)-c\\<&\lambda\int_{a^*(c,T)}^1[Z(a|a^*(c,T),T)-Z(a^*|a^*(c,T),T)]dF(a)-c\\=&0.\end{aligned}$$

Then, the intermediate value theorem and the strict monotonicity yield the unique cut-off value of $\hat{a}_j \in (a_1^I, a^*(c,T))$.

However, if Firm j searches with cut-off $\hat{a}_j \in [a_1^I, a^*(c,T))$, Firm i's best response is to search with a cut-off value $\hat{a}_i \in (\hat{a}_j, a^*(c,T))$ rather than a_1^I, because the instantaneous gain from searching for Firm i in a state $\tilde{a}$ is

$$\lambda\int_{\tilde{a}}^1[Z(a|\hat{a}_1,T)-Z(\tilde{a}|\hat{a}_1,T)]dF(a)-c\begin{cases}<0 & \text{for } \tilde{a}=a^*(c,T)\\>0 & \text{for } \tilde{a}=\hat{a}_1\end{cases}$$

and it is monotone w.r.t. $\tilde{a}$. This results in contradiction. Hence, there is no equilibrium in which either firm searches with a_1^I as the cut-off.

Last, we only need to consider the case in which each firm searches with a cut-off higher than a_1^I. Following the same argument as in the proof of Theorem 3.1, we have $(a^*(c,T), a^*(c,T))$ being the unique equilibrium.

□

Proof of Proposition 3.2. We apply Theorem 3.2 here for the analysis. We only need to show the case for $a_1^I \in (a^*(c,T), F^{-1}(1-\frac{c}{\lambda}))$. In this case, Firm 1 does not to search and Firm 2 searches with a_1^I as the cut-off. Now, take the limit $a_1^I \to a^*(c,T)$ from the right hand side of $a^*(c,T)$. In the limit, where Firm 2 searches with $a^*(c,T)$ as the cut-off, Firm 1 weakly prefers not to search. If $a_1^I = a^*(c,T)$, Firm 1 is actually indifferent between searching and not. Hence, a head start in the limit makes Firm 1 weakly better off. Firm 1's payoff when it does not search is $e^{-\lambda T[1-F(a_1^I)]}$, the probability of Firm 2 ending up in a state below a_1^I, is strictly increasing in a_1^I. Hence, a higher value of the head start makes Firm 1 even better off. □

Proof of Proposition 3.3. [1]. For T being small, $a^*(c,T) = 0$. $D^M(0, a_1^I) = 0$, and the partial derivative of $D^M(T, a_1^I)$ w.r.t. T when T is small is

$$\frac{\partial D^M(T, a_1^I)}{\partial T} = \lambda(1-a_1^I)e^{-\lambda T(1-a_1^I)}[1-\frac{c}{\lambda(1-a_1^I)}] - \lambda e^{-2\lambda T} + ce^{-\lambda T},$$

which equals to $-\lambda a_1^I < 0$ at the limit of $T = 0$.

[2]. Follows from Propositions 3.1 and 3.2. □

Proof of Proposition 3.4. $D^M(T, a_1^I)$ is strictly decreasing in a_1^I, and it goes to the opposite of (3.12), which is less than 0, as a_1^I goes to $F^{-1}(1-\frac{c}{\lambda})$, and

$$(1-e^{-\lambda T[1-F(a^*(c,T))]}) - \frac{1}{2}(1-e^{-2\lambda T}) \tag{3.15}$$

as a_1^I goes to $a^*(c,T)$. Hence, if (3.15) is positive, Case 1 yields from the intermediate value theorem; Case 2 holds if (3.15) is negative. □

3.A.4 Proofs for the Extended Models

Proof of Proposition 3.7. We argue that, to determine a subgame perfect equilibrium, we only need to consider two kinds of strategies profiles:

a Firm 1 retains its initial innovation and does not search, and Firm 2 searches with a_1^I as the cut-off;

b Firm 1 discards its initial innovation and searches with a_2^I as the cut-off, and Firm 2 retains its initial innovation.

First, suppose $c < \frac{\lambda}{2}(1+e^{-\lambda T})$. If Firm 1 retains the initial innovation, it will have no incentive to search, and Firm 2 is indifferent between discarding the initial innovation and not. In either case, Firm 2 searches with a_1^I as the cut-off. Given that Firm 1 has discarded its initial innovation, Firm 2 has no incentive to discard its initial innovation as shown in Proposition 3.2.

Second, suppose $c > \frac{\lambda}{2}(1 + e^{-\lambda T})$. In the subgame in which both firms discard their initial innovation, there are two equilibria, in each of which one firm searches with 0 as the cut-off and the other firm does not search. Hence, to determine a subgame perfect equilibrium, we have to consider another two strategy profiles, in addition to [a] and [b]:

c Firm 1 discards its initial innovation and searches with 0 as the cut-off, and Firm 2 discards its initial innovation and does not search.

d Firm 1 discards its initial innovation and does not search, and Firm 2 discards its initial innovation and searches with 0 as the cut-off.

However, we can easily rule out [c] and [d] from the candidates for equilibria. In [c], Firm 2 obtains a payoff of 0. It can deviate by retaining its initial innovation so as to obtain a positive payoff. Similarly, in [d], Firm 1 can deviate by retaining its initial innovation to obtain a positive payoff rather than 0.

Last, it remains to compare Firm 1's payoff in [a] and [b]. In [a], Firm 1's payoff is

$$e^{-\lambda T[1-F(a_1^I)]}. \tag{3.16}$$

In [b], it is

$$(1 - e^{-\lambda T[1-F(a_2^I)]})(1 - \frac{c}{\lambda[1 - F(a_2^I)]}). \tag{3.17}$$

The difference between these two payoffs, (3.17) and (3.16), is increasing in T, and it equals -1 when $T = 0$ and goes to $1 - \frac{c}{\lambda[1-F(a_2^I)]} > 0$ as T approaches infinity. Hence, the desired result is implied by the intermediate value theorem. □

Proof of Proposition 3.8. The backward induction is similar to the proof of Proposition 3.7, and thus is omitted. □

Proof of Proposition 3.9. The equilibrium for the subgame starting from time t_0 derives from Theorem 3.2. Suppose at time t_0, Firm i is in a state a_i^0, where $\max\{a_1^0, a_2^0\} \geq a_1^I$. Assume $a_i^0 > a_j^0$. If $a_i^0 > F^{-1}(1 - \frac{c}{\lambda})$, then Firm i obtains a continuation payoff of 1, and Firm j obtains 0. If $a_i^0 \in (a^*(c, T - t_0), F^{-1}(1 - \frac{c}{\lambda}))$, then Firm i obtains a continuation payoff of $e^{-\lambda(T-t_0)[1-F(a_i^0)]}$, and Firm j obtains $(1 - e^{-\lambda(T-t_0)[1-F(a_i^0)]})(1 - \frac{c}{\lambda[1-F^{-1}(a_i^0)]})$.

To prove this result, we first show that not to search before t_0 is Firm 2's best response regardless of Firm 1's action before time t_0. It is equivalent to showing that not to search before

t_0 is Firm 2's best response if Firm 2 knows that Firm 1 is definitely going to be in any state $a_1^0 \geq a_L^*$ at time t_0.

As we have shown before, for any $a_1^0 \geq (>) F^{-1}(1-\frac{c}{\lambda})$, Firm 2 (strictly) prefers not to conduct searching before time t_0.

If $a_1^0 \in [a_L^*, F^{-1}(1-\frac{c}{\lambda}))$, Firm 2's unique best response before time t_0 is not to search. The instantaneous gain from searching at any time point before t_0 for Firm 2 in a state below a_1^0 is

$$
\begin{aligned}
&\lambda\big[[1-F(F^{-1}(1-\frac{c}{\lambda}))] + \int_{a_1^0}^{F^{-1}(1-\frac{c}{\lambda})} e^{-\lambda(T-t_0)[1-F(a)]} dF(a) \\
&\qquad - [1-F(a_1^0)](1-e^{-\lambda(T-t_0)[1-F(a_i^0)]})(1-\frac{c}{\lambda[1-F^{-1}(a_i^0)]})\big] - c \\
=&\lambda\big[\int_{a_1^0}^{F^{-1}(1-\frac{c}{\lambda})} e^{-\lambda(T-t_0)[1-F(a)]} dF(a) \\
&\qquad - [1-F(a_1^0)](1-e^{-\lambda(T-t_0)[1-F(a_i^0)]})(1-\frac{c}{\lambda[1-F^{-1}(a_i^0)]})\big],
\end{aligned}
$$

which is strictly negative when $T-t_0$ is sufficiently large, and thus conducting a search before time t_0 actually makes Firm 2 strictly worse off in this case.

Next, we show that Firm 1's best response before time t_0 is to search with $F^{-1}(1-\frac{c}{\lambda})$ as the cut-off, if Firm 2 does not search before t_0. To see this, look at the instantaneous gain from searching for Firm 1 in a state below $F^{-1}(1-\frac{c}{\lambda})$:

$$
\begin{aligned}
&\lambda\big[[1-F(F^{-1}(1-\frac{c}{\lambda}))] \\
&\qquad + \int_{a_1^I}^{F^{-1}(1-\frac{c}{\lambda})} e^{-\lambda(T-t_0)[1-F(a)]} dF(a) - [1-F(a_1^I)]e^{-\lambda(T-t_0)[1-F(a_1^I)]}\big] - c \\
=&\int_{a_1^I}^{F^{-1}(1-\frac{c}{\lambda})} e^{-\lambda(T-t_0)[1-F(a)]} dF(a) - [1-F(a_1^I)]e^{-\lambda(T-t_0)[1-F(a_1^I)]} \\
>&\int_{\tilde{a}}^{F^{-1}(1-\frac{c}{\lambda})} e^{-\lambda(T-t_0)[1-F(a)]} dF(a) - [1-F(\tilde{a})]e^{-\lambda(T-t_0)[1-F(a_1^I)]} \\
>&(1-\frac{c}{\lambda}-F(\tilde{a}))e^{-\lambda(T-t_0)[1-F(\tilde{a})]} - [1-F(\tilde{a})]e^{-\lambda(T-t_0)[1-F(a_1^I)]} \\
=&e^{-\lambda(T-t_0)[1-F(a_1^I)]}[1-\frac{c}{\lambda}-F(\tilde{a})]\left(e^{\lambda(T-t_0)[F(\tilde{a})-F(a_1^I)]} - \frac{1-F(\tilde{a})}{1-\frac{c}{\lambda}-F(\tilde{a})}\right),
\end{aligned}
$$

where $\tilde{a}$ is any value in $(a_1^I, F^{-1}(1-\frac{c}{\lambda}))$. The term on the right hand side of the last equality is strictly positive if $T-t_0$ is sufficiently large. Hence, the desired result yields.

□

3.A.5 Proofs for the Case with Asymmetric Costs

Proposition 3.12. *If* $0 < c_1 < c_2 < \frac{1}{2}\lambda(1 - e^{-\lambda T})$ *there exists a pure strategy equilibrium* (a_1^*, a_2^*) *with* $a_1^*, a_2^* \geq 0$.

Proof of Proposition 3.12. We prove the existence of equilibrium by applying Brouwer's fixed point theorem. First, same as in the previous proofs, if Firm j searches with a cut-off $\hat{a}_j \geq 0$, the instantaneous gain from searching for Firm i in state 0 is

$$\lambda \int_0^1 Z(a|1, T) - c_i > 0,$$

and thus Firm i is better off continuing searching if it is in state 0.

Next, let us define for each Firm j a critical value

$$\alpha_j = \sup\{a_j \in [0, 1] \mid I(0|\alpha_j, c_i) = \lambda \int_0^1 [Z(a|\alpha_j) - Z(0|\alpha_j)]dF(a) - c_i > 0\}.$$

Suppose there is a $\alpha_j \in (0, 1)$ such that

$$I(0|\alpha_j, c_i) = \lambda \int_0^1 [Z(a|\alpha_j) - Z(0|\alpha_j)]dF(a) - c_i = 0.$$

For any $\hat{a}_j \in [0, \alpha_j]$,

$$I(0|\hat{a}_j, c_i) \geq 0 \text{ and}$$
$$I(1|\hat{a}_j, c_i) < 0.$$

By the intermediate value theorem and the strict monotonicity of $Q(a|\hat{a}_j, c_i)$ in a, there must exist a unique $\tilde{a}_i \in [0, 1)$ such that

$$I(\tilde{a}_i|\hat{a}_j, c_i) = 0.$$

That is, if Firm j searches with cut-off $\hat{a}_j$, Firm i's best response is to search with cut-off $\tilde{a}_i$.

For any $\hat{a}_j \in (\alpha_j, 1]$, if the set is not empty,

$$I(0|\hat{a}_j, c_i) < 0.$$

That is, Firm i's best response is to search with cut-off 0.

Then, we could define two best response functions $BR_i : [0,1] \to [0,1]$ where

$$BR_i(\hat{a}_j) := \begin{cases} 0 & \text{for } \hat{a}_j \in (\alpha_j, 1] \text{ if it is not empty} \\ \tilde{a}_i & \text{where } I(\tilde{a}_i|\hat{a}_j, c_i) = 0 \text{ for } \hat{a}_j \in [0, \alpha_j]. \end{cases}$$

It is also easy to verify that BR_i is a continuous function over $[0,1]$. Hence, we have a continuous self map $BR : [0,1]^2 \to [0,1]^2$ where

$$BR = (BR_1, BR_2)$$

on a compact set, and by Brouwer's fixed point theorem, there must exist of a pure strategy equilibrium in which each Firm searches with a cut-off higher than or equal to 0. □

Proof of Proposition 3.10. First, using the same arguments as in the proof of Proposition 3.1, we claim that if there exists an equilibrium it must be the case that each firm searches with a cut-off higher than or equal to 0 with one strictly positive value for one firm.

Next, we show that there can be no equilibrium in which Firm 2 searches with a cut-off $\hat{a}_2 > 0$ and Firm 1 searches with cut-off 0. Such a strategy profile $(0, \hat{a}_2)$ is an equilibrium if and only if

$$\lambda \int_0^1 [Z(a|\hat{a}_2, T) - Z(0|\hat{a}_2, T)] dF(a) - c_1 \le 0, \quad \text{and}$$
$$\lambda \int_{\hat{a}_2}^1 [Z(a|0, T) - Z(\hat{a}_2|0, T)] dF(a) - c_2 = 0.$$

However,

$$\begin{aligned} 0 &= \lambda \int_{\hat{a}_2}^1 [Z(a|0, T) - Z(\hat{a}_2|0, T)] dF(a) - c_2 \\ &< \lambda \int_{\hat{a}_2}^1 [Z(a|\hat{a}_2, T) - Z(\hat{a}_2|\hat{a}_2, T)] dF(a) - c_2 \\ &< \lambda \int_0^1 [Z(a|\hat{a}_2, T) - Z(0|\hat{a}_2, T)] dF(a) - c_1 \le 0, \end{aligned}$$

resulting in a contradiction.

Next, we derive the necessary and sufficient conditions for the existence of an equilirbium in which Firm 2 searches with a cut-off 0 and Firm 1 searches with a cut-off strictly higher than 0.

A pair of cut-off rules $(\hat{a}_1, 0)$, $\hat{a}_1 > 0$, is an equilibrium if and only if

$$\lambda \int_{\hat{a}_1}^{1} [Z(a|0,T) - Z(\hat{a}_1|0,T)]dF(a) - c_1 = 0 \quad \text{and} \tag{3.18}$$

$$\lambda \int_{0}^{1} [Z(a|\hat{a}_1,T) - Z(0|\hat{a}_1,T)]dF(a) - c_2 \leq 0, \tag{3.19}$$

where

$$(3.18) \Leftrightarrow \frac{1}{2}\lambda(1-e^{-\lambda T})[1-F(\hat{a}_1)]^2 - c = 0 \Leftrightarrow \hat{a}_j = F^{-1}\left(1 - \sqrt{\frac{2c_1}{\lambda(1-e^{-\lambda T})}}\right). \tag{3.20}$$

Then, (3.20) and (3.19) together imply that $(\hat{a}_i, 0)$ is an equilibrium if and only if

$$I\left(0|F^{-1}\left(1 - \sqrt{\frac{2c_1}{\lambda(1-e^{-\lambda T})}}\right), c_2\right) \leq 0. \tag{3.21}$$

We will see that if (3.21) holds there is no other equilibrium.

When (3.21) does not hold, there is a unique equilibrium, in which each firm searches with a cut-off strictly higher than 0. Because by Proposition 3.12 an equilibrium must exists. Let (a_1^*, a_2^*) be such an equilibrium. We first show that $a_1^* > a_2^*$ must hold by proof by contradiction, and then we show that it must be a unique equilibrium. Such a pair (a_1^*, a_2^*) is an equilibrium if and only if

$$\lambda \int_{a_i^*}^{1} [Z(a|a_j^*,T) - Z(a_i^*|a_j^*,T)]dF(a) = c_i \quad \text{for } i = 1,2 \text{ and } j \neq i. \tag{3.22}$$

Suppose $a_1^* \leq a_2^*$. Applying Lemma 3.6, we have

$$\begin{aligned} c_1 =& \lambda \int_{a_1^*}^{1} [Z(a|a_2^*,T) - Z(a_1^*|a_2^*,T)]dF(a) \\ \geq& \lambda \int_{a_2^*}^{1} [Z(a|a_2^*,T) - Z(a_2^*|a_2^*,T)]dF(a) \\ \geq& \lambda \int_{a_2^*}^{1} [Z(a|a_1^*,T) - Z(a_2^*|a_1^*,T)]dF(a) = c_2, \end{aligned}$$

resulting in a contradiction.

Then, we show the uniqueness of the equilibrium for Cases [1] – [3] by contradiction. For Case [1] we show that the solution to (3.22) is unique, and for Cases [2] and [3] we show that there can be no equilibrium in which each firm searches with a cut-off higher than 0 coexisting with equilibrium $\left(F^{-1}\left(1 - \sqrt{\frac{2c_1}{\lambda(1-e^{-\lambda T})}}\right), 0\right)$. We can prove all of them together. Suppose there are two equilibria (a_1^*, a_2^*) and $(\tilde{a}_1^*, \tilde{a}_2^*)$, where (a_1^*, a_2^*) is a solution to (3.22) and $(\tilde{a}_1^*, \tilde{a}_2^*)$ is either

$\left(F^{-1}\left(1-\sqrt{\frac{2c_1}{\lambda(1-e^{-\lambda T})}}\right),0\right)$ or a solution to (3.22). It is sufficient to show that the following two cases are not possible:

1. $\tilde{a}_1^* > a_1^* > a_2^* > \tilde{a}_2^* \geq 0$ and

2. $a_1^* > \tilde{a}_1^* > a_2^* > \tilde{a}_2^* \geq 0$.

Suppose $\tilde{a}_1^* > a_1^* > a_2^* > \tilde{a}_2^* \geq 0$. Applying Lemma 3.6 we have

$$\begin{aligned}
0 =& \lambda \int_{a_1^*}^1 [Z(a|a_2^*,T) - Z(a_1^*|a_2^*,T)]dF(a) - c_1 \\
<& \lambda \int_{\tilde{a}_1^*}^1 [Z(a|a_2^*,T) - Z(\tilde{a}_1^*|a_2^*,T)]dF(a) - c_1 \\
<& \lambda \int_{\tilde{a}_1^*}^1 [Z(a|\tilde{a}_2^*,T) - Z(\tilde{a}_1^*|\tilde{a}_2^*,T)]dF(a) - c_1 = 0,
\end{aligned}$$

resulting in a contradiction.

Suppose $a_1^* > \tilde{a}_1^* > a_2^* > \tilde{a}_2^* \geq 0$. Applying Lemma 3.6 again, we have

$$\begin{aligned}
0 \geq& \lambda \int_{\tilde{a}_2^*}^1 [Z(a|\tilde{a}_1^*,T) - Z(\tilde{a}_2^*|\tilde{a}_1^*,T)]dF(a) - c_2 \\
>& \lambda \int_{a_2^*}^1 [Z(a|\tilde{a}_1^*,T) - Z(a_2^*|\tilde{a}_1^*,T)]dF(a) - c_2 \\
>& \lambda \int_{a_2^*}^1 [Z(a|a_1^*,T) - Z(a_2^*|a_1^*,T)]dF(a) - c_2 = 0,
\end{aligned}$$

resulting in another contradiction.

□

Proof of Proposition 3.11. For fixed c_2 we have

$$\frac{\partial a_2^*}{\partial a_1^*} = -\frac{\frac{\partial \int_{a_2^*}^1 [Z(a|a_1^*,T)-Z(a_2^*|a_1^*,T)]dF(a)}{\partial a_1^*}}{\frac{\partial \int_{a_2^*}^1 [Z(a|a_1^*,T)-Z(a_2^*|a_1^*,T)]dF(a)}{\partial a_2^*}} = \frac{\int_{a_2^*}^1 \frac{\partial [Z(a|a_1^*,T)-Z(a_2^*|a_1^*,T)]}{\partial a_1^*} dF(a)}{\frac{\partial Z(a_2^*|a_1^*,T)}{\partial a_2^*}} < 0.$$

Then,

$$\begin{aligned}
\frac{\partial a_1^*}{\partial c_1} &= -\frac{-1}{\lambda \int_{a_1^*}^1 \left[\frac{\partial [Z(a|a_2^*,T)-Z(a_1^*|a_2^*,T)]}{\partial a_2^*}\frac{\partial a_2^*}{\partial a_1^*} - \frac{\partial Z(a_1^*|a_2^*,T)}{a_1^*}\right] dF(a)} < 0 \text{ and} \\
\frac{\partial a_2^*}{\partial c_1} &= \frac{\partial a_2^*}{\partial a_1^*}\frac{\partial a_1^*}{\partial c_1} > 0.
\end{aligned}$$

For fixed c_1 we have

$$\frac{\partial a_1^*}{\partial a_2^*} = -\frac{\frac{\partial \int_{a_1^*}^1 [Z(a|a_2^*,T) - Z(a_1^*|a_2^*,T)] dF(a)}{\partial a_2^*}}{\frac{\partial \int_{a_1^*}^1 [Z(a|a_2^*,T) - Z(a_1^*|a_2^*,T)] dF(a)}{\partial a_1^*}} = \frac{\int_{a_1^*}^1 \frac{\partial [Z(a|a_2^*,T) - Z(a_1^*|a_2^*,T)]}{\partial a_2^*} dF(a)}{\frac{\partial Z(a_1^*|a_2^*,T)}{\partial a_1^*}} > 0.$$

Then,

$$\frac{\partial a_2^*}{\partial c_2} = -\frac{-1}{\lambda \int_{a_2^*}^1 \left[\frac{\partial [Z(a|a_1^*,T) - Z(a_2^*|a_1^*,T)]}{\partial a_1^*} \frac{\partial a_1^*}{\partial a_2^*} - \frac{\partial Z(a_2^*|a_1^*,T)}{a_2^*} \right] dF(a)} < 0 \text{ and}$$

$$\frac{\partial a_1^*}{\partial c_1} = \frac{\partial a_1^*}{\partial a_2^*} \frac{\partial a_2^*}{\partial c_2} < 0.$$

□

Chapter 4

Optimal Prize Structures in Elimination Contests

4.1 Introduction

Many real-life interactions among different agents such as elections, the implementation of innovations, promotion tournaments and sports can be well described and analyzed through contest models. Many contests involve multiple stages where the number of agents compete at successive stages until the winner is finally determined. The most prominent model sharing this feature is the sequential elimination contest, which is commonly known from playoff rounds in sports competitions. This model is also a good description of many corporate tournaments where employees from low hierarchical levels compete for promotions to higher hierarchical levels. This contest structure is also found in politics, where candidates compete in localized contests and where the winners subsequently often compete against each other at higher levels. R&D can also be described by such tournaments, where the firm with the most efficient technology wins the market[1].

In this chapter, we look for the optimal prize structure of a sequential elimination contest that maximizes the profit of the designer. Existing literature mainly considers the objective of maximizing agents' efforts in the tournament (usually either the average level or the effort in the final round). The prize pool is usually assumed to be a fixed amount of money, which can be distributed to the agents according to their performance. We think that this is a suitable assumption for one-stage simultaneous tournaments, but in multi-stage tournaments, the designer's valuation of various stages may differ. Thus, we allow for an arbitrary form of the designer's objective. We assume that agents' efforts at different stages produce the output according to some output function, which defines how exactly the designer values various combinations of efforts at different stages. In our model, the prize pool is not fixed, as the designer can pay the contestants

[1] An example of multi-stage R&D contest is Pre-commercial Procurement used in EU. Here, firms participate in a R&D multi-stage contest where competing projects are evaluated phase-by-phase (e.g., proposal, prototype, testing) and competitors are eliminated sequentially (European Comission, 2007)

any amount of money. The designer's problem is to choose the prize structure that maximizes the designer' profit, which is the difference between the value of the produced output and the value of the prizes distributed to the agents. This problem can be seen in many real-life applications. For example, the owners of a firm care about their profit, not the average level of effort applied by the workers. Moreover, they can value the efforts of workers at higher positions more than the efforts of those at lower positions(a higher position means a later stage in our model). As we show, in this case, the optimal prize structure would drastically differ from the one that is optimal when the average level of efforts is maximized. Hence, the results obtained in the literature for multi-stage elimination under the objective of total (or average) effort maximization are not valid in the general case. For example, in his classical paper, Rosen (1986) shows that the optimal prize structure is linear with a bigger prize gap in the final round. Our results show that this is not true in general. The prize structure can be not only concave or convex but also non-monotone.

We find that depending on the output function various prize structures might be optimal. The prize at the first stage is always negative, and thus, it is essentially an entry fee. The equilibrium level of effort is efficient and by appropriate choice of entry fee, the designer is able to extract the full surplus. The structure of other prizes can differ substantially depending on the output function. We consider an example of a sport tournament to illustrate the result when the prize structure is increasing. We also provide an example in which the output function depends only on one parameter, where for various values of this parameter five types of prize structures might be optimal: (1) increasing concave prize structure, (2) increasing linear prize structure, 3) Increasing convex prize structure, 4) winner-take-all structure, and (5) decreasing prize structure with the big prize to the winner. The last case with negative prize differences is the most interesting of all cases. It turns out that if the designer values efforts at each stage much more than the efforts at the previous stage, the difference between the prizes at these stages could be negative. Thus, the participants who survive longer may receive smaller prizes than those eliminated at the earlier stages. We refer to this structure as a *trap structure.* Since non-monotonicity of prize structures with respect to stages is new in the contests literature, we study this issue in detail and provide sufficient conditions for both monotonicity and non-monotonicity of prize structures.

Decreasing prize structures are not observed in real-life direct contests very often. One possible explanation is that they require unlimited liability and substantial payments *from* participants. It may not be feasible to make such contracts in real-life contests. However, in indirect real-life contests, that is situations that are not directly tournaments but can be thought as tournaments, "prizes" can be decreasing. The following story may be an example of such situation. Students who are interested in careers in academia participate in a real-life contest called “getting tenured”. This contest can be seen as a multi-stage elimination tournament where the stages are: 1) Undergraduate, 2) Master, 3) PhD, 4) Junior faculty and 5) Tenured faculty. One may agree that the prize when an average student is eliminated at stage 1)—gets an average job for an undergraduate—may be higher than the prize he gets when he is eliminated at stage 4) as it could be difficult for an untenured researcher to get a good job in industry[2].

We also investigate the possible optimal types of prize structure shapes. We find that any shape with increasing prizes would be optimal for some specific designer's preferences over stages. This result says that apriori the designer cannot favor any particular class of prize structures, for example convex, concave, or linear. Specific valuations of different stages by the designer should be taken into account.

Then, we consider the case of limited liability when the designer is not allowed to allocate negative prizes. Hence, the obtained optimal prize structure is not feasible in this case. We add one more degree of freedom for the designer, in which she can optimize. Not only is optimization with respect to the prize structure available, but also the choice of a contest success function (CSF) is possible. It means that at each stage, she can specify the probability of an agent moving to the next stage depending on realized efforts. The important result here is that it is still possible to implement efficiency and extract full surplus from agents even with limited liability by choosing CSF in a special way from the class of Tullock functions. With this additional degree of freedom, it is possible to avoid non-monotone structures. The optimal prize structure would be always increasing.

This chapter is structured as follows. In the next section, we provide a literature review. Then, we describe our main model and solve for the optimal prize structure. Next, we investigate some properties of the solution and provide an example to illustrate the solution. Then we show that any increasing shape of prize structure may be optimal for a certain designer's valuation of stages. Finally, we consider the limited liability case and allow the designer to choose a CSF. The last section concludes and summarizes the chapter.

4.2 Literature Review

The literature about tournaments, for example, that concerning lottery contests, R&D, patents, or innovation implementation, is surveyed by Konrad et al. (2009) in detail. The classic work that describes simple tournaments is Lazear and Rosen (1981). They consider the case of the simple

[2] I would like to thank an anonymous referee for this example.

simultaneous one-stage tournament and compare rank-order wage schemes with wages based on individual output and find that for risk-neutral agents, both wage schemes allocate resources in an efficient way. The playoff or elimination tournaments considered in this chapter were analyzed for the first time in Rosen (1986), to which this chapter is largely related. While Rosen (1986) describes a double-elimination tournament with multiple stages, we allow for a multi-elimination tournament; in other respects, the contest architecture is the same in both studies. Further, different from this chapter, in Rosen (1986) there is a fixed prize pool. The designer's objective is simply to maximize the same constant level of effort of all agents through the tournament. He has found that the optimal incentive scheme is such that the difference between prizes at all stages is the same. That is, the optimal prize growth is linear. It is true for all stages except the last one. Since the contest will not continue, the prize spread between the first and the second places should be significantly larger compared with the prize spreads between the final and the semi-final, the semi-final and the quarter-final, and so on. It is also never optimal to pay losers in the first round because it lowers their incentives to put forth an effort. The result of this work is a particular case of our study in which the output function is simply a sum of all efforts. In other cases, the optimal prize structure seriously differs from Rosen's optimal structure. We show that the optimal prize structure strongly depends on the output function and that only in a separable linear case it coincides with Rosen's optimal structure.

Skaperdas (1996) gives an axiomatic characterization of contest success functions. Under a set of reasonable axioms the probability that an individual i wins a contest has a form $\frac{f(x_i)}{\sum_j f(x_j)}$, where x_j is the effort of individual j. We use such functions in our model. Clark and Riis (1998) generalize Skaperdas results for the asymmetric setting.

Further, there are several papers that study the efficiency in contests. Chung (1996) discusses a rent-seeking model where productive efforts increase the single rent for which agents compete. He considers a winner-take-all contest with linear costs. For this setup, he shows that the equilibrium efforts are always greater than socially optimal ones. However, this is not the case in this chapter because efforts in our model do not increase prizes directly and all agents are assumed to be risk-neutral. Hence, in our setting the designer is always able to implement the efficient level of effort and extract the whole rent from the contestants.

There are several papers where optimal prize structures in simultaneous one-stage contests are considered. Moldovanu and Sela (2001) show that for convex cost functions it is optimal to give positive prizes not only to the winner, although for concave and linear cost functions, the winner-take-all structure is optimal. In a similar framework, Shaffer (2006) compares payoffs and efforts arising from exogenously given prizes with those from effort-dependent prizes. Cohen and Sela (2007) characterize the optimal effort-dependent prize structure in the one-stage all-pay auction setup. Depending on the designer's objective, they find that the optimal reward may decrease or increase with the players' efforts.

Schweinzer and Segev (2012) analyze simultaneous Tullock contests. They give necessary and sufficient conditions under which there is an equilibrium under the winner-take-all structure and show that if it exists, then it is unique. If it does not exist, they construct a prize structure with several prizes, under which an equilibrium exists. Though we consider a multi-stage tournament instead of a simultaneous tournament, the condition for the existence of a symmetric pure strategy equilibrium is the same.

Though the architectural structure of the contest and the number of stages are fixed in this chapter, there are several papers that consider the question of optimal tournament design. Fu and Lu (2012) show that for a fixed prize pool and a linear cost function, under the objective of maximization of the total effort, the optimal structure is such that at each state only one contestant is eliminated until the final, and the single winner takes over the entire prize. However, in this chapter, we show that for a classical elimination tournament, the winner-take-all structure is almost never optimal unless the cost function is linear. Gradstein and Konrad (1999) allow the designer to choose the number of stages in a contest and the way the agents are matched. They show that the optimal number of stages crucially depends on the particular Tullock function.

The role of punishment (negative prize) for losers in a one-stage tournament was discussed in Moldovanu et al. (2012) and Thomas and Wang (2013). The multi-stage contests considered in this chapter allow allocating negative prizes not only to the losers but also to the contestants who have been successful in every stage but one. Thus, in the optimum, we obtain not only negative prizes for the losers at the first stage, but actually decreasing and non-monotone prize schemes.

Kolmar and Sisak (2014) analyze the public good provision by heterogeneous players. The contest prizes are financed from taxation. This is similar to our model because, at each stage in our model, the prizes are paid from the revenue, generated by the agents. Kolmar and Sisak show how multiple prizes can be used to achieve efficiency.

Further, there are several papers, such as Gürtler and Kräkel (2010), Parreiras and Rubinchik (2010), Ryvkin (2007) that consider tournament settings with heterogeneous agents. We discuss complications which arise in our model if we allow for heterogeneity of agents.

4.3 Main Model

The tournament begins with m^N players and proceeds sequentially through N stages. At each stage, all participants who survived up to the current stage are randomly selected into groups. They compete within groups, and only one winner from each group moves to the next stage. This situation is well-known for $m = 2$, which is the case of football playoffs and tennis tournaments. Winners move to the next round and losers are eliminated from the subsequent play. In the next round other groups are randomly drawn, and again, half of the participants are excluded from further competition. In a general case there are m^{N+1-n} agents at the stage n who are distributed to $m^{N+1-n}/m = m^{N-n}$ groups. Each agent competes in a group of m contestants. The top prize W_{N+1} is awarded to the winner of the final match, who has won N matches overall. Other finalists, that is losers at the final stage, are awarded the second place and get the prize W_N for having won $N-1$ stages and losing the last one. In earlier stages, all participants eliminated at the same stage get equal prizes. We denote W_n as the prize for the losers at stage n. We emphasize that each agent receives exactly one prize only at the stage where he is eliminated.

In our main model equally talented players are considered. The probability of agent i moving to the next round $P(x_{i,n}, \mathbf{x}_{-i,n})$ is the function of an agent's level of effort $x_{i,n}$ at that stage and a vector of effort levels of competitors $\mathbf{x}_{-i,n}$ in the same group. It is assumed to be symmetric, increasing in $x_{i,n}$ and decreasing in each component of $\mathbf{x}_{-i,n}$. In this section the probability of winning a match at some particular stage is assumed to be the following Tullock function:

$$P(x_{i,n}, \mathbf{x}_{-i,n}) = \frac{x_{i,n}^a}{x_{i,n}^a + \sum\limits_{j \neq i} x_{j,n}^a},$$

where $a > 0$ is a constant parameter and the sum is taken across contestants in the group with agent i. In real life this means that there are some observable characteristics that are connected with levels of effort and show whose effort is higher, but not perfectly.

Here we come to the crucial part of our model, namely, the output function, which determines, how the designer values agents' efforts at different stages. Denote $\mathbf{x}_n = (x_{i,n}, \mathbf{x}_{-i,n})$ as a vector of agents' efforts at stage n. Let $\Pi(\mathbf{x}_1, ..., \mathbf{x}_N)$ be some concave, increasing in each component, continuously differentiable output function of efforts that is symmetric with respect to different efforts at the same stage. We emphasize that the output function is not necessarily separable with respect to effort levels at different stages. In the existing literature, only specific examples are considered. For example, the designer's objective is to maximize either the level of effort at the final or the average level of effort through the whole tournament. In this chapter, we characterize the optimal prize structure for a general form of an output function.

When the prize structure is announced, agents choose their efforts by considering the announced prize structure. If the prize structure $\Omega = \{W_1, ..., W_N, W_{N+1}\}$ is announced, the profit of the designer corresponds to the difference between the output produced by the agents and the total amount of prizes.

$$\Pi(\mathbf{x}_1(\Omega), ..., \mathbf{x}_N(\Omega)) - (\sum_n (m^{N+1-n} - m^{N-n})W_n + W_{N+1}). \tag{4.1}$$

At each stage, every player who has survived up to this point chooses his effort. However, applying the level of effort x is accompanied by the costs equal to $C(x) = x^\gamma$, $\gamma > 0$. A player's decision of how much effort to spend in any match depends on a cost-benefit analysis. Higher effort increases the probability of winning this match and moving to the next stage but involves higher costs.

We assume that the efforts of contestants are socially desirable. In orther words, we assume that there exists a unique positive vector of levels of efforts at all stages, $\mathbf{x}^e$, that maximizes the social surplus, that is, $\mathbf{x}^e = \arg\max_{\mathbf{x}}\{\Pi(\mathbf{x}_1, ..., \mathbf{x}_N) - \sum_n \sum_j C(x_{j,n})\}$, which can be found from the first-order condition:

$$\frac{\partial \Pi(\mathbf{x}_1^e, ..., \mathbf{x}_N^e)}{\partial x_{i,n}} = C'(x_n^e). \tag{4.2}$$

Suppose that some prize structure Ω is announced. Denote V_n as the value of participation in the tournament for every player at stage n. Since all players are assumed to be symmetric, this value of participation is equal for all players. It consists of two components. The first one is the prize, which is earned if the match is lost, and the player is eliminated. This event occurs with the probability $1 - P(x_{i,n}, \mathbf{x}_{-i,n})$. The other is the value of moving to the next stage if the match is won. The probability of this event is $P(x_{i,n}, \mathbf{x}_{-i,n})$. Either way, he also incurs costs of effort $C(x_{i,n})$. Instead of participation in the contest, agents can choose not to participate at all, which brings them reservation utility 0[3]. We assume unlimited liability in this section, that is, the designer is able to allocate negative prizes for agents, provided that their participation constraint holds. Agents are assumed to be risk-neutral[4]. The solution concept is a sub-game

[3] Instead of 0 it could also be some positive number u_R. All results are largely the same, with the only change being a parallel shift of the whole prize structure upwards. The shape of the prize structure and the differences between prizes remain unchanged.

[4] For risk-averse agents, we can use the similar techniques by using a concave function $U(W)$ instead of W and

perfect equilibrium. We also restrict our attention to symmetric equilibria. Therefore, we can write the agent's problem recursively as follows:

$$V_{i,n} = \max_{x_{i,n}\geq 0}(1 - \frac{x_{i,n}^a}{x_{i,n}^a + \sum\limits_{j\neq i} x_{j,n}^a})W_n + \frac{x_{i,n}^a}{x_{i,n}^a + \sum\limits_{j\neq i} x_{j,n}^a}V_{i,n+1} - C(x_{i,n}). \quad (4.3)$$

4.4 Solution and Main Results

4.4.1 Solution to agents' problem

We start with the solution to the agents' problem (4.3). Existence of symmetric equilibria with a positive level of effort depends on the relation between γ and a. The solution is described in the following lemma. Index i is skipped because we consider a symmetric solution. We use ΔW_n to denote the prize spread between prizes at stages n and $n+1$, that is $\Delta W_n = W_{n+1} - W_n$.

Lemma 4.1. *Suppose that* $\sum\limits_{j=n}^{N} \varkappa^{j-n}\Delta W_j \geq 0$ *for all* $1 \leq n \leq N$.

If $a > \frac{m}{m-1}\gamma$ *there is no symmetric equilibrium in pure strategies.*

If $a \leq \frac{m}{m-1}\gamma$ *there is a unique symmetric equilibrium in pure strategies such that the level of effort* x_n^* *at stage* n, *given the prize structure* Ω, *does not depend on the prizes at all earlier stages and increases with an increase in the prize difference at that stage and at all later stages with decreasing weights:*

$$\frac{\gamma m^2}{a(m-1)}C(x_n^*) = \Delta W_n + \varkappa\Delta W_{n+1} + \varkappa^2\Delta W_{n+2} + ... + \varkappa^{N-n}\Delta W_N, \quad (4.4)$$

where $\varkappa = \frac{(\gamma-a)m+a}{\gamma m^2}$,
and $x_n^* = 0$ *if* $\Delta W_n + \varkappa\Delta W_{n+1} + \varkappa^2\Delta W_{n+2} + ... + \varkappa^{N-n}\Delta W_N < 0$.

Proof. See Appendix 4.A. □

If the condition on prize differences ($\sum\limits_{j=n}^{N} \varkappa^{j-n}\Delta W_j \geq 0$ for each n) does not hold, the levels of efforts will be zero at all stages, where $\sum\limits_{j=n}^{N} \varkappa^{j-n}\Delta W_j < 0$. However, as we show below, when the designer chooses the optimal structure, this is never the case and the efficient level of efforts is implemented.

Lemma 4.1 can be used to find the optimal prize scheme. From this moment we consider only the case $a \leq \frac{m}{m-1}\gamma$, when the symmetric equilibrium in pure strategies exists.

$\Delta U(W_n) = U(W_{n+1}) - U(W_n)$ instead of $\Delta W_n = W_{n+1} - W_n$ in what follows. However, the efficiency is lost and the characterization of the optimal prize structure would be more complex.

4.4.2 Designer's problem

The designer knows how agents choose efforts (from Lemma 4.1). Thus, she maximizes the profit, given by expression (4.3), subject to the agents' participation constraint. Thus, the designer's problem can be written by using prize spreads in the following way:

$$\begin{aligned} \Pi(\mathbf{x}_1(\Omega), ..., \mathbf{x}_N(\Omega)) - m^N W_1 \quad & -(\sum_n m^{N-n} \Delta W_n) \Longrightarrow \max_{\Omega}, \\ s.t. V_1 > 0. & \end{aligned} \tag{4.5}$$

Solving this problem, we can formulate the following proposition:

Proposition 4.1. *The optimal prize structure implements the efficient level of efforts in the equilibrium at all stages, i.e.,* $x_n^* = x_n^e$ *for each* n.

Proof. See Appendix 4.A. □

The result is intuitive since there is no private information and the agents are risk neutral. One can also notice that if the designer does not value efforts at some stage at all, then the equilibrium level of efforts in that stage would be equal to zero.

If we take the equilibrium levels of efforts from Proposition 1 and rewrite equations (2.12), we can formulate the following result:

Proposition 4.2. *The optimal prize structure satisfies the following:*

$$\begin{aligned} W_1 &= -\frac{\gamma \varkappa m^2}{a(m-1)} C(x_1^*) q, \\ \Delta W_n &= \frac{\gamma m^2}{a(m-1)} (C(x_n^*) - \varkappa C(x_{n+1}^*)), \quad n \neq N, \\ \Delta W_N &= \frac{\gamma m^2}{a(m-1)} C(x_N^*), \end{aligned} \tag{4.6}$$

where $\{x_n^*\}$ *satisfies:*

$$\frac{\partial \Pi(\mathbf{x}_1^*, ..., \mathbf{x}_N^*)}{\partial x_{i,n}} = C'(x_n^*).$$

Proof. See Appendix 4.A. □

This result suggests that the prize structure must be constructed as follows. First, the prize at the first stage is always non-positive and depends on the efficient level of efforts at this stage. Thus, we can think about it as the entry fee. It is used in order to extract the full rent from the agents[5]. Next, the optimal prize spread at some particular stage depends on the efficient level of effort at that stage and the next stage. The higher the desirable level of effort at some stage is, the higher the optimal difference in prizes between this and the next stage is. The effect of the efficient level of effort at the next stage is opposite. A higher efficient level of effort at the next stage implies smaller prize difference at the current stage. If the difference between efficient levels of effort is not too large or simply negative (i.e. efficient level of efforts is decreasing), the prize structure is increasing because $C(x_n^*) - \varkappa C(x_{n+1}^*) > 0$. However, if the efficient level of efforts at the next stage is much higher, then prize spread would be necessarily negative. Finally, the difference in prizes for the winner and other finalists is necessary positive. Hence, we get the following corollary.

Corollary 4.1. *1. The optimal prize structure is monotone and increasing with respect to stages if and only if* $C(x_n^e) \geq \varkappa C(x_{n+1}^e)$ *for* $n = 1, ..., N-1$.

2. If at some stage the efficient level of effort is much lower than that at the next stage, namely $C(x_n^e) < \varkappa C(x_{n+1}^e)$, *the optimal prize at the next stage must be lower than the current prize:* $W_{n+1} < W_n$.

Since $\varkappa = \frac{(\gamma-a)m+a}{\gamma m^2} < \frac{\gamma m}{\gamma m^2} = \frac{1}{m}$ in order to satisfy $C(x_n^e) < \varkappa C(x_{n+1}^e)$, the efficient level of effort at stage $n+1$ must be much higher. The intuition behind this result is that negative prize difference at the current stage, which makes the next prize lower, will also put agents in the situation where they strongly do not want to lose at the next stage. Otherwise, they would simply prefer taking the prize of the current round and not going further. The prize for the winner would be so big that every agent would benefit from going further and being closer to it, even if the prize structure decreases at some point. Indeed,

$$V_n - V_{n-1} = (1-\varkappa) \sum_{k=n-1}^{N} \varkappa^{k-n+1} \Delta W_k = \frac{(1-\varkappa)\gamma m^2}{a(m-1)} C(x_{n-1}^*) \geq 0. \tag{4.7}$$

We can show also that the prize for the winner is always the biggest prize. Indeed, the sum of later prize spreads starting from any stage is positive:

[5] If reservation utility is greater than zero $u_R > 0$, then the expression for the prize at the first stage in Proposition!4.2 must be changed to $W_1 = u_R - \sum_n \varkappa^n \Delta W_n$. Essentially, this is a parallel shift of the optimal prize structure.

$$\begin{aligned}\sum_{k=n}^{N}\Delta W_k &= \sum_{k=n}^{N-1}\frac{\gamma m^2}{a(m-1)}(C(x_n^*)-\varkappa C(x_{n+1}^*))+\frac{\gamma m^2}{a(m-1)}C(x_N^*) = \\ &= \sum_{k=n+1}^{N}\frac{\gamma m^2}{a(m-1)}(1-\varkappa)C(x_k^*)+\frac{\gamma m^2}{a(m-1)}C(x_n^*)\geq 0.\end{aligned}$$

4.4.3 Example of a sport tournament

Consider pairwise elimination tournament with 2^N players that proceeds sequentially through N stages. For example, if $N=4$ we have 16 players who play firstly eighth-finals, then quarter-finals, then semi-finals, and the final at the last stage. According to this scheme the playoff rounds of many sports are conducted. Assume that there is a representative fan, who wants to buy a ticket for each match at each stage and has the following utility function:

$$U(\overline{x}_n, n) = \overline{x}_n\sqrt{n} - p_n, \tag{4.8}$$

where $\overline{x}_n$ is the average level of effort in a match and p_n is the price at the stage n. We multiply efforts at stage n on a square root of n to reflect the importance of later stages. The cost function is $C(x)=x^2$. CSF has the following form: $P(x_{i,n},\mathbf{x}_{-i,n})=\frac{x_{i,n}}{x_{i,n}+\sum\limits_{j\neq i}x_{j,n}}$ The contest designer extracts full surplus from a fan and charges $p_n = x_n\sqrt{n}$. Thus, the output function $\Pi(\mathbf{x}_1,...,\mathbf{x}_N) = \sum_{n=1}^{N}2^{N-n}\overline{x_n}\sqrt{n}$. Proposition 4.1 implies that the equilibrium level of effort coincides with the efficient one $x_n^e = \arg\max_{x_n\geq 0}(x_n\sqrt{n}-2x_n^2) = 0.25\sqrt{n}$. Thus, the equilibrium level of effort increases in later stages with a speed of $\sqrt{n}$. Applying Proposition 4.2 and considering that $\gamma=2$, $a=1$, $m=2$, and $\varkappa=3/8$ we have the following optimal prize structure:

$$\begin{aligned} W_1 &= -\frac{3}{16}, \\ \Delta W_n &= \frac{5n-3}{16}, \quad n\neq N, \\ \Delta W_N &= N/2. \end{aligned}$$

Going back from differences to levels we have $W_1 = -\frac{3}{16}$, $W_2 = -\frac{1}{16}$, $W_3 = \frac{3}{8},...,W_n = \frac{(5n-11)n}{32},...,W_{N+1} = \frac{(5N+5)N}{32}$. Thus, the prizes at the first and the second stages are negative, and then, it grows with an increasing rate of order n^2. In this example of a sport tournament, the prize structure is strictly increasing everywhere. However, we can construct a simple example to illustrate that it is not always true.

4.4.4 Example with different optimal structures

In this section, we provide a simple but rich enough example to illustrate different optimal structures.

Suppose that the output function has the following form:

$$\Pi(.) = \sum_n (\lambda^n \sum_i x_{i,n}), \quad \lambda > 0.$$

Costs are quadratic:

$$C(x) = x^2$$

Thus, the designer cares about the average level of efforts at different stages. The parameter λ determines the weights she attaches to different stages.

Then, we can apply Proposition 4.2 and get the following optimal prize structure:

$$\begin{aligned} \Delta W_n &= \frac{m^2}{m-1}\frac{\lambda^{2n}}{2}(1-\varkappa\lambda^2), n \neq N, \\ \Delta W_N &= \frac{m^2}{m-1}\frac{\lambda^{2N}}{2}. \end{aligned}$$

Now we consider several cases for the parameter λ :

1. $\lambda < 1$ (Figure 4.1). In this case, the designer values the later stages less than the earlier stages. The optimal prize structure is increasing ($1 - \varkappa\lambda^2 > 0$) and concave ($\lambda^{2n}$ decreases with larger values of n, and hence, ΔW_n falls)[6].

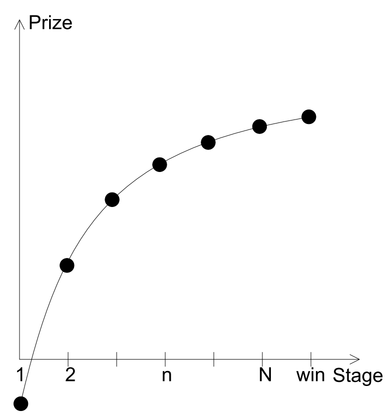


Figure 4.1: Concave prize shape

[6]At the last stage, the prize structure is not necessary concave.

2. $\lambda = 1$ (Figure 4.2). All stages are equally important. The optimal prize structure is linearly increasing with a jump in the final ($\Delta W_n = const > 0, n \neq N$). This result is the case of Rosen (1986), where the designer maximizes the same average level of effort during the tournament.

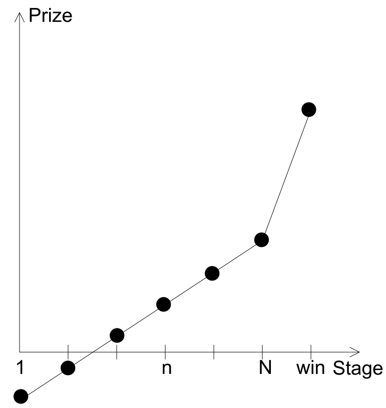


Figure 4.2: Linear prize shape

3. $1 < \lambda < \sqrt{\frac{1}{\varkappa}}$ (Figure 4.3). The designer values the later stages more, but not drastically. The optimal prize structure is increasing ($1 - \varkappa\lambda^2 > 0$) and convex ($\lambda^{2n}$ increases with larger values on n, and hence, ΔW_n increases).

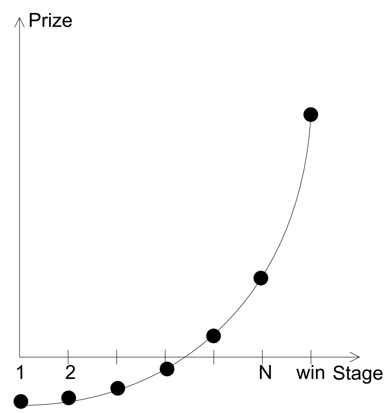


Figure 4.3: Convex prize shape

4. $\lambda = \sqrt{\frac{1}{\varkappa}}$ (Figure 4.4). The optimal prize structure is winner-take-all (all prize spreads equal to zero $\Delta W_n = 0$, except the first and the last stages). Further, there are several papers where the winner-take-all structure turns out to be optimal in other settings (Krishna and Morgan, 1998, Moldovanu and Sela, 2001).

5. $\lambda > \sqrt{\frac{1}{\varkappa}}$ (Figure 4.5). The designer values the later stages drastically more than the earlier ones. As $1 - \varkappa\lambda^2 < 0$, the prize spread ΔW_n should be negative for all intermediate stages

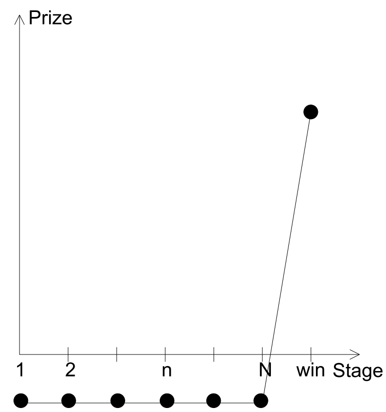


Figure 4.4: Winner-take-all

of the tournament. Therefore, the optimal prize structure is decreasing, with a large final prize being awarded to the winner. This is an example of a "trap structure". When the designer values each subsequent stage much more than the previous one, her valuation of the final is so high that she tries to make the gap between the prize for the winner and prizes for the other finalists as high as possible. Thus, using negative prize differences - and, hence, negative prizes - the designer puts agents in a situation where they are punished more if they go closer to the final and lose there. At the later stages, stakes become extremely large, which enforces very high levels of efforts, as is needed by the principal. Though the prizes become more negative and agents who survive longer obtain smaller prizes at all stages except the final, the value from surviving until the later stages increases because the agent gets closer to the final prize.

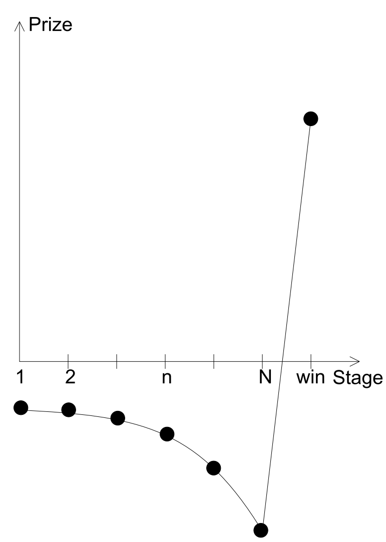


Figure 4.5: Trap

4.4.5 (Non-)Monotonicity of prize structures

In the previous discussion, we have shown that optimal prize structures may vary a lot. In other words, the shape of a prize structure may not be only convex or concave but even non-monotone or decreasing. Here, we consider the case of a separable output function and directly address the question of monotonicity, not in terms of the efficient level of efforts but in terms of the output function.

We assume here that the output function is separable with respect to different stages, that is, $\Pi(\mathbf{x}_1(\Omega), ..., \mathbf{x}_N(\Omega)) = \sum\limits_n \Pi_n(\mathbf{x}_n(\Omega))$.[7] To begin with, suppose that $\frac{\partial \Pi_n(x,...,x)}{\partial x_n} \geq \frac{\partial \Pi_{n+1}(x,...,x)}{\partial x_{n+1}}$ for any x, which means that efforts at earlier stages are more important for the designer than at later stages. Then, the equilibrium level of efforts falls during the contest and $\frac{\gamma m^2}{a(m-1)}(C(x_n^*) - \varkappa C(x_{n+1}^*)) \geq 0$[8]. Hence, all prize spreads are non-negative and the prize structure is increasing. However, in real-life situations, this assumption usually does not hold. Thus, we need to consider a more plausible case where $\frac{\partial \Pi_n(\mathbf{x}_n^*)}{\partial x_n} < \frac{\partial \Pi_{n+1}(\mathbf{x}_{n+1}^*)}{\partial x_{n+1}}$, that is, efforts at later stages are more important than at the earlier ones. This is a natural assumption in many real-life situations such as corporate tournaments and sports tournaments. For example, in application to a firm, this would mean that the activities of workers at higher levels of corporate hierarchies are more important than the activities of those at lower ones. In a sport tournament, the assumption means that the performance of contestants in last rounds is valued more than in the early ones. This case is not only the most reasonable but also the most interesting one: the prize structure is not necessarily monotone here.

The main result here is that if the valuations of effort do not increase too much from each stage to the next stage, and, simultaneously, the output function is concave enough, then the optimal prize structure is always non-decreasing. The inverted conditions together serve as sufficient conditions for "trap" structures. The exact statement is the following[9]:

Proposition 4.3. *1. If $\frac{\partial \Pi_{n+1}(x,...,x)}{\partial x_{n+1}} \leq \frac{1}{\varkappa}\frac{\partial \Pi_n(x,...,x)}{\partial x_n}$ and $x\frac{\partial \Pi_n(x,...,x)}{\partial x_n}$ is decreasing for all x and n, then the optimal prize structure is increasing at all stages. If the inequality is strict and $x\frac{\partial \Pi_n(x,...,x)}{\partial x_n}$ is strictly decreasing, then the optimal prize structure is strictly increasing.*

2. If at some stage n, the opposite holds, that is, $\frac{\partial \Pi_{n+1}(x,...,x)}{\partial x_{n+1}} \geq \frac{1}{\varkappa}\frac{\partial \Pi_n(x,...,x)}{\partial x_n}$ and $x\frac{\partial \Pi_n(x,...,x)}{\partial x_n}$ is

[7] In many real-life applications, this is a reasonable assumption. For example, it is natural to assume that for sports events the revenues from selling tickets on semi-final matches do not depend on the teams' efforts in quarter-finals.

[8] As $\frac{\partial \Pi_n(.)}{\partial x_{i,n}} = C'(x_n^*)$, a decrease of the derivative of the output function would lead to a decrease of the equilibrium level of efforts.

[9] For simplicity of notations we skip index i because all agents are treated in the same way.

increasing for all x, then the optimal prize is decreasing at stage n. If the inequality is strict and $x\frac{\partial \Pi_n(x,...,x)}{\partial x_n}$ is strictly increasing then the optimal prize is also strictly decreasing at this stage. Hence, the optimal prize structure would be non-monotone[10].

Proof. See Appendix 4.A. □

However, these conditions are only sufficient, not necessary conditions.[11] The most interesting finding is in the second part of this proposition, which implies that if the designer values some stage sufficiently higher than the previous one, the prize at this stage must be lower than the prize at the previous stage. The intuition here is similar to the intuition in the 5th case of the example. Agents react to the prize spreads at all later stages. By decreasing prize at some stage, the designer is able to increase efforts applied at that stage.

4.4.6 Variability of optimal prize structures

As we have seen in an example, many shapes of prize structures can be optimal for specific output functions. Here, we show that the class of prize structures that may be optimal is very large. In fact, *any* increasing shape is optimal for some output function. To show it we start with the following observation:

Claim 4.1. *For any levels of efforts $x_1, ..., x_N$ there exists an output function such that these effort levels are implemented in equilibrium under the optimal prize structure.*

Proof. Consider some arbitrary levels of efforts $x_1, ...x_N$ at different stages. Then we can notice that for these efforts there exists a separable output function $\Pi(\mathbf{x}_1, ..., \mathbf{x}_N) = \sum\limits_n \Pi_n(\mathbf{x}_n)$ such that these efforts would be implemented in equilibrium under the optimal prize structure. For each x_n we can find a concave, increasing in each component, continuously differentiable symmetric function $\Pi_n(\mathbf{x}_n) : \frac{\partial \Pi_n(x_n,...,x_n)}{\partial x_n} = C'(x_n)$. Then, by Proposition 4.1, $x_1, ...x_N$ are implemented in the equilibrium under the optimal prize structure. □

Thus, any levels of efforts may arise as the equilibrium levels of efforts for a particular prize structure.

Now, consider some increasing *shape* of a prize structure, that is, $\{\Delta W_1, ..., \Delta W_N\}$ such that $\Delta W_n \geq 0$. For this prize shape construct effort levels at different stages by the following recursive procedure:

[10] This is because the prize difference in the final is always positive.
[11] See our previous example

$$
\begin{aligned}
x_N &: \quad C(x_N) = \frac{a(m-1)}{\gamma m^2}\Delta W_N, \\
x_{N-1} &: \quad C(x_{N-1}) = \frac{a(m-1)}{\gamma m^2}\Delta W_{N-1} + \varkappa C(x_N), \\
& \quad ... \\
x_n &: \quad C(x_n) = \frac{a(m-1)}{\gamma m^2}\Delta W_n + \varkappa C(x_{n+1}), \\
& \quad ...
\end{aligned}
$$

The construction is valid because $\Delta W_n \geq 0$ implies that x_n are nonnegative. By Claim 4.1 there is an output function such that these levels of efforts are the equilibrium ones under the optimal prize structure. Note also that the constructed efforts and the given prize shape satisfies the system (4.6). Hence, by Proposition **??** it is the optimal shape, that is, $\Delta W_n^* = \Delta W_n$ for any stage n. Thus, we can formulate the following result.

Proposition 4.4. *For any increasing shape of prize structure* $\Delta W_1, ...\Delta W_N \geq 0$, *there exists an output function* $\Pi(\mathbf{x}_1, ..., \mathbf{x}_N)$ *such that this prize shape is optimal.*

There are two important things that we should mention in relation to this result. The first one is that this result is about the optimality of the prize structure shape and not the prize structure itself. That is, we do not state anything concerning prize W_1 for the losers of the first stage. Second, since our recursive procedure has to generate non-negative values of efforts, $\frac{a(m-1)}{\gamma m^2}\Delta W_n + \varkappa C(x_{n+1})$ has to be positive. A sufficient condition for this is positive prize spreads $\Delta W_n \geq 0$. Thus, for some decreasing prize structures, this procedure would not work. Thus, although the set of prize structures that might be optimal is large which includes all increasing prize structures, a non-monotone prize structure can always be suboptimal.

4.5 Limited liability and the optimal CSF

From the previous discussion we know that the optimal prize at the first stage $W_1 = -\frac{\gamma m^2}{a(m-1)}C(x_1^*)$ is always non-positive (and even negative if the designer values the efforts at the first stage). We have already seen that in some cases the optimal prize structure is decreasing at some stages and, hence, non-monotone. In many real-life situations, negative or decreasing prizes are not feasible. In this section, we impose one additional restriction on the optimal prize structure, namely, limited liability. It means that the designer is not allowed to make prizes negative. Thus, the optimal prize structure under unlimited liability is not feasible under limited liability. The main result here suggests that if the sufficient conditions as in Proposition 3 hold, then the optimal prize structure would be close to the optimal one in the case of unlimited liability with only one change at the first stage. That is, $W_1 = 0$ and $\frac{\partial \Pi(.)}{\partial x_{i,n}} = \frac{m}{a(m-1)} \gamma C'(x_n^*)$ and the rest remains the same. However, since the efficient level of effort is not implementable anymore, the profit of the principal reduces in comparison with the case of unlimited liability.

Now, we show how the designer can implement the efficient levels of efforts and extract the full surplus from contestants if she is given one more degree of freedom with respect to the construction of a contest. We assume that the designer not is only free to choose a prize structure but also can implement any CSF. In modeling real-life situations, this can sometimes be a reasonable assumption. If we consider a sport tournament, it is hard to believe that the designer can somehow affect the probability of a win, provided that rules of a game are given. However, if we consider a promotion tournament, the owner of a firm can make decisions about the way of competition between workers, and our assumption is much more plausible here.

In this section, we can drop our previous assumption about the particular form of the cost function. Thus, instead of $C(x) = x^\gamma$, in this part, it can be any cost function $C(x)$ such that equation (4.2) has the unique interior solution.

In the previous sections, CSF was assumed to be the following Tullock function: $P(x_{i,n}, \mathbf{x}_{-i,n}) = \frac{x_{i,n}^a}{x_{i,n}^a + \sum\limits_{j \neq i} x_{j,n}^a}$. We have shown that if degree of convexity of a cost function x^γ is low, that is $a > \frac{m}{m-1}\gamma$, it is not possible to induce positive efforts in a symmetric equilibrium. However, if the designer can choose a contest success function, then, as we show below, for any cost function she is able to implement the efficient level of efforts in the equilibrium. Moreover, this function can be found in the class of Tullock functions: $P(x_{i,n}, \mathbf{x}_{-i,n}) = \frac{f(x_{i,n})}{f(x_{i,n}) + \sum\limits_{j \neq i} f(x_{j,n})}$.

Thus, in this section, we ask the following question: if the designer could choose any $P(x_{i,n}, \mathbf{x}_{-i,n})$, which one is better in the sense of maximizing the profit under limited liability?

Let us suppose that we can find some CSF that satisfies the following condition for every $e > 0$:

$$\frac{P'_{x_{i,n}}(e, \mathbf{e})}{P(e, \mathbf{e})} = \frac{C'(e)}{C(e)}. \tag{4.9}$$

Then, using the similar arguments as in the previous section, we can write the following system of equations, which defines our optimal prize structure and the equilibrium after choosing a CSF satisfying (4.9):

$$\begin{aligned} \frac{\partial \Pi(.)}{\partial x_{i,n}} &= C'(x_n^*), \\ W_1 &= 0, \\ \Delta W_n &= mC(x_n^*). \end{aligned}$$

We can now see that all efforts at all stages are efficient, and the designer is still able to extract the whole surplus from the agents without allocating negative prizes. Hence, any probability function that satisfies (4.9) would be an optimal CSF. We can also notice that any optimal CSF does not depend on the output function $\Pi(.)$. Thus, irrespective of how the principal values efforts of agents, she should choose the same CSF at all stages.

We have not shown yet the existence of such probabilistic functions that satisfy (4.9). Now, we demonstrate that such a function always exists. That is, for any cost function $C(x)$ we can find a contest success function from the class of Tullock functions $P(x_{i,n}, \mathbf{x}_{-i,n}) = \frac{f(x_{i,n})}{f(x_{i,n})+\sum\limits_{j\neq i} f(x_{j,n})}$, which satisfies sufficient condition (4.9).

Lemma 4.2. *For any function $C(x)$ take $f(x) = C^{\frac{m}{m-1}}(x)$. Then $P(x_{i,n}, \mathbf{x}_{-i,n}) = \frac{f(x_{i,n})}{f(x_{i,n})+\sum\limits_{j\neq i} f(x_{j,n})}$ satisfies the following condition: $\frac{P'_{x_{i,n}}(e,\mathbf{e})}{P(e,\mathbf{e})} = \frac{C'(e)}{C(e)}$.*

Proof. See Appendix 4.A. □

This lemma shows how the optimal CSF is constructed. Thus, the following must be true:

Proposition 4.5. *The following structure of the elimination contest is optimal:*

1. $P(x_{i,n}, \mathbf{x}_{-i,n}) = \frac{f(x_{i,n})}{f(x_{i,n})+\sum\limits_{j\neq i} f(x_{j,n})}$, *where* $f(x) = C^{\frac{m}{m-1}}(x)$
2. $W_1 = 0$, $\Delta W_n = mC(x_n^*)$,
   where $x_n^* : \frac{\partial \Pi(.)}{\partial x_{i,n}} = C'(x_n^*)$.

Proof. See Appendix 4.A. □

This is our main result here. We can compare it with Proposition 4.2. If the designer is able to choose the CSF, then the optimal prize structure is always non-decreasing and satisfies the limited liability requirement. Thus, there is no need to implement "trap" and other non-monotone structures. The intuition behind this result is that the optimal CSF makes agents indifferent between participation at each stage and choosing zero level of effort. Hence, at each stage, $V_n = W_n$. If we compare this to the results from the previous section we see that, in fact, the whole dynamic structure of the tournament is broken because the value of participation at each stage equals the prize at that stage. Therefore, for contestants, this tournament is equivalent to participation in a sequence of independent one-stage tournaments. Thus, the prize difference must be non-negative at each stage because this prize difference is equivalent to the prize in the particular one-stage tournament.

If designer's valuation of effort is increasing with later stages, then the optimal prize structure would be convex because prize spreads, $\Delta W_n = mC(x_n^*)$, would be increasing. In our opinion, this explains why one can observe convex prize structures very often in real-life tournaments.

4.6 Discussion

In this chapter, we considered elimination contests and studied how the optimal prize structure depends on the objective of the designer. Efforts of agents at different stages of a tournament generate output for the principal according to some output function. Depending on this function we characterized the optimal prize structure in the tournament that gives the highest profit for the designer. We showed that the optimal prize structure is also efficient. Sometimes, the optimal prize structure is non-monotone if the designer's valuation of efforts at some stage is much higher than that at the previous stage. To illustrate the variability of optimal structures we gave a simple example where different prize structures were optimal, depending on one parameter. Some of these structures have already been characterized as optimal in the existing literature; however, this is not the case for non-monotone and decreasing prize schemes. For example, under the "trap" structure, prizes for agents are smaller at the later stages than at the earlier ones. If the prize at the first stage is negative, it means that all prizes at all later stages would also be negative, except the prize for the winner of the tournament. Further, we provided necessary and sufficient conditions for the optimality of non-monotone structures.

In addition, we considered the case of limited liability, where the designer is not able to offer negative prizes but is free to choose a contest success function. We showed that though prizes cannot be negative, the optimal choice of CSF enables us to implement efficient levels of efforts and extract the full surplus with only positive prizes. The optimal CSF does not depend on the output function. Thus, irrespective of how the principal values efforts at different stages, it is optimal to choose the same CSF.

The assumptions about the constant rate of elimination, equal cost functions and contest success function at all stages make it possible to get an analytical solution for the optimal prize structure in the form as in Proposition 4.2. It would be interesting to consider a model with heterogeneous agents. Unfortunately, due to increased complexity of the model, we are unable to obtain the analytical results for this case. The problem is that if agents are ex-ante different, the continuation payoff depends on the current opponent's type and on all other competitors' types. Hence, Proposition 4.1 does not hold anymore and the equilibrium levels of efforts are different from the efficient ones. However, the logic of the analysis suggests that non-monotone prize structures would be optimal also for a setup with heterogeneous agents. Further, there is no reason to think that relaxing assumptions or considering heterogeneous agents would help avoid non-monotone and decreasing prize structures in the optimum.

4.A Appendix

Proof of Lemma 4.1. First, assume that the solution to the agent's problem is interior. We prove that the conditions for that are exactly those as specified in the statement of this lemma. Differentiating (2.1) with respect to $x_{i,n}$ for $1 \leq n \leq N$, we get the first-order condition:

F.O.C.:

$$\frac{a x_{i,n}^{a-1} \sum\limits_{j \neq i} x_{j,n}^{a}}{\left(x_{i,n}^{a} + \sum\limits_{j \neq i} x_{j,n}^{a} \right)^{2}} (V_{i,n+1} - W_n) = C'(x_{i,n}).$$

In the symmetric equilibrium we skip index i later and thus have:

$$V_{n+1} - W_n = \frac{m^2}{m-1} \frac{x_n^*}{a} C'\left(x_n^*\right) = \frac{m^2}{m-1} \frac{\gamma}{a} C\left(x_n^*\right). \tag{4.10}$$

Taking $C(x_n^*)$ from the previous equation and substituting it into the value function we get the following difference equation for V_n:

$$V_n = \frac{m-1}{m} W_n + \frac{1}{m} V_{n+1} - \frac{m-1}{m^2} \frac{V_{n+1} - W_n}{\gamma / a}.$$

Writing it recursively we get

$$V_{n+1} - W_n = \Delta W_n + \varkappa \Delta W_{n+1} + \varkappa^2 \Delta W_{n+2} + ... + \varkappa^{N-n+1} \Delta W_N, \tag{4.11}$$

where $\varkappa = \frac{(\gamma - a)m + a}{\gamma m^2}$.

Then,

$$\frac{m^2}{m-1}\frac{\gamma}{a}C\left(x_n^*\right) = V_{n+1} - W_n = \Delta W_n + \varkappa\Delta W_{n+1} + \varkappa^2\Delta W_{n+2} + ... + \varkappa^{N-n+1}\Delta W_N.$$

Obviously, the level of efforts does not depend on prizes at the previous stages and is greater than zero if $\Delta W_n + \varkappa\Delta W_{n+1} + \varkappa^2\Delta W_{n+2} + ... + \varkappa^{N-n+1}\Delta W_N > 0$.

Now, we show, when the interior solution is an equilibrium, that is:

$$x_n^* \in Arg\max_{x_n\geq 0}\{(1 - \frac{x_n^a}{x_n^a + (m-1)(x_n^*)^a})W_n + \frac{x_n^a}{x_n^a + (m-1)(x_n^*)^a}V_{n+1} - C(x_n)\}.$$

Consider the case $a = \frac{m}{m-1}\gamma$. Then, $\varkappa = 0$, $V_n = W_n$, $V_{n+1} - W_n = mC(x_n^*)$. Hence, we need to show that

$$x_n^* \in Arg\max_{x_n\geq 0}\{\frac{x_n^a}{x_n^a + (m-1)(x_n^*)^a}mC(x_n^*) - C(x_n)\}.$$

Denote

$$Q_n(x_n) := \frac{x_n^a}{x_n^a + (m-1)(x_n^*)^a}mC(x_n^*) - C(x_n).$$

Notice that $Q_n(x_n^*) = 0$. Hence, we need to show that $Q_n(x_n) \leq 0$ for all $x_n > 0$. Denote $x_n^a = f(x_n)$. Then we need to show that

$$mf(x_n)C(x_n^*) - f(x_n)C(x_n) - mC(x_n)f(x_n^*) + f(x_n^*)C(x_n) \leq 0.$$

If $a = \frac{m}{m-1}\gamma$ then $f(x_n) = C^{\frac{m}{m-1}}(x)$. Hence, the last inequality can be rewritten as

$$(mC(x_n^*) - C(x_n))C^{\frac{1}{m-1}}(x_n) - (m-1)C(x_n^*)C^{\frac{1}{m-1}}(x_n^*) \leq 0.$$

The derivative of the left-hand side is equal to the following expression:

$$\frac{m}{m-1}C'(x_n)C^{\frac{1}{m-1}}(x_n)\frac{C(x_n^*) - C(x_n)}{C(x_n)}.$$

For $x_n < x_n^*$, this expression is positive. For $x_n > x_n^*$, it is negative. Hence, $Q_n(x_n)$ attains maximum at $x_n = x_n^*$, which guarantees that the interior stationary point is a global maximizer. However, it is not a unique maximizer. Since $V_n = W_n$, applying x_n^* gives the same payoff as applying zero level of effort at each stage.

Next, in the case $a < \frac{m}{m-1}\gamma$ we have $\varkappa > 0$. By the similar arguments, the interior solution to F.O.C. would be a unique global maximizer.

Now, we show that if $a > \frac{m}{m-1}\gamma$, there is no symmetric equilibrium in pure strategies. If it exists, then continuation values satisfy (4.11). Since we supposed that $\sum\limits_{j=n}^{N} \varkappa^{j-n}\Delta W_j \geq 0$, equality (4.11) implies that $V_{n+1} - W_n \geq \Delta W_n$. On the other hand, since $a > \frac{m}{m-1}\gamma$ we have $\varkappa < 0$. Hence, the same equality (1.13) implies $V_{n+1} < W_{n+1}$. Hence, we obtain a contradiction which means that there is no symmetric equilibrium in pure strategies. □

Proof of Proposition 4.1. From the proof of Lemma 4.1, we know that

$$V_1 = W_1 + \sum_n \varkappa^n \Delta W_n.$$

Hence, we can rewrite the designer's problem (4.5) as

$$\Pi(\mathbf{x}_1(\Omega), ..., \mathbf{x}_N(\Omega)) - m^N W_1 - (\sum_n m^{N-n}\Delta W_n) \Longrightarrow \max_{\Omega},$$

$$s.t. W_1 + \sum_n \varkappa^n \Delta W_n \geq 0.$$

We can notice that the participation constraint must be binding because otherwise, it would be possible to decrease W_1 and increase the profit. Considering this, we can substitute $W_1 = -\sum\limits_n \varkappa^n \Delta W_n$ into the profit function:

$$\Pi(\mathbf{x_1}(\Omega), ..., \mathbf{x_N}(\Omega)) + m^N \sum_n \varkappa^n \Delta W_n - (\sum_n m^{N-n}\Delta W_n) \Longrightarrow \max_{\Omega}$$

First, we note that from Lemma 4.1 for any $1 \leq n \leq N$ the following is true:

$$\frac{\partial x_n^*}{\partial \Delta W_n} = \frac{a(m-1)}{m^2 \gamma C'(x_n^*)}.$$

Next, we take first-order conditions:

$$\begin{aligned} \frac{\partial}{\partial \Delta W_n} : m^{N+1-n} \quad & \frac{\partial \Pi(\mathbf{x}_1(\Omega), ..., \mathbf{x}_N(\Omega))}{\partial x_{i,n}} \frac{\partial x_n^*}{\partial \Delta W_n} + \\ + m^{N+1-(n-1)} \quad & \frac{\partial \Pi(\mathbf{x}_1(\Omega), ..., \mathbf{x}_N(\Omega))}{\partial x_{i,n-1}} \frac{\partial x_{n-1}^*}{\partial \Delta W_n} + \\ + ... + m^N \quad & \frac{\partial \Pi(\mathbf{x}_1(\Omega), ..., \mathbf{x}_N(\Omega))}{\partial x_{i,1}} \frac{\partial x_1^*}{\partial \Delta W_n} = m^{N-n} - m^N \varkappa^n, \end{aligned}$$

$$\frac{\partial}{\partial \Delta W_{n-1}} \quad : \quad m^{N+1-(n-1)} \frac{\partial \Pi(\mathbf{x}_1(\Omega), ..., \mathbf{x}_N(\Omega))}{\partial x_{i,n-1}} \frac{\partial x^*_{n-1}}{\partial \Delta W_{n-1}} + ... +$$
$$+ \qquad m^N \frac{\partial \Pi(\mathbf{x}_1(\Omega), ..., \mathbf{x}_N(\Omega))}{\partial x_{i,1}} \frac{\partial x^*_1}{\partial \Delta W_{n-1}} = m^{N-(n-1)} - m^N \varkappa^{n-1}.$$

As we have already noticed,

$$\frac{\partial x^*_{n-1}}{\partial \Delta W_{n-1}} = \frac{a(m-1)}{m^2 \gamma C'(x^*_{n-1})}.$$

Using Lemma 4.1 we can get the following:

$$\frac{\partial x^*_{n-1}}{\partial \Delta W_n} = \varkappa \frac{a(m-1)}{m^2 \gamma C'(x^*_{n-1})} = \varkappa \frac{\partial x^*_{n-1}}{\partial \Delta W_{n-1}}.$$

Substituting the last expression into the F.O.C. we obtain

$$m^{N+1-n} \frac{\partial \Pi(\mathbf{x}_1(\Omega), ..., \mathbf{x}_N(\Omega))}{\partial x_{i,n}} \frac{a(m-1)}{m^2 \gamma C'(x^*_n)} + \varkappa (m^{N-(n-1)} - m^N \varkappa^{n-1}) =$$
$$= \quad m^{N-n} - m^N \varkappa^n.$$

Hence,

$$\frac{\partial \Pi(\mathbf{x}_1(\Omega), ..., \mathbf{x}_N(\Omega))}{\partial x_{i,n}} = C'(x^*_n).$$

This holds for every $n \neq 1$.

For $n = 1$ the following holds:

$$m^N \frac{\partial \Pi(\mathbf{x}_1(\Omega), ..., \mathbf{x}_N(\Omega))}{\partial x_{i,1}} \frac{\partial x^*_1}{\partial \Delta W_1} = m^{N-1} - \varkappa m^N.$$

Hence,

$$\frac{\partial \Pi(\mathbf{x}_1(\Omega), ..., \mathbf{x}_N(\Omega))}{\partial x_{i,1}} = \frac{m^{N-1} - \varkappa m^N}{m^N} * \frac{m^2 \gamma C'(x^*_1)}{a(m-1)} = C'(x^*_1).$$

The only thing we need to explain is why the F.O.C. gives us the optimum. The equilibrium under the proposed prize structure coincides with the social optimum, that is $\mathbf{x}^* = \mathbf{x}^e$, and simultaneously, the designer is able to extract full surplus from agents. Hence, the obtained prize structure is optimal.

□

Proof of Proposition 4.2. The expressions for the equilibrium effort levels follow from Proposition 4.1.

Then, for any $1 \leq n \leq N-1$, we can express ΔW_n by using the expression for two consequent effort levels from Lemma 4.1:

$$\Delta W_n = \frac{\gamma m^2}{a(m-1)}(C(x_n^*) - \varkappa C(x_{n+1}^*)),\ n \neq N$$

$$\Delta W_N = \frac{\gamma m^2}{a(m-1)} C(x_N^*)$$

For W_1 we have the following from the previous equations and the proof of Proposition 4.1:

$$\begin{aligned} W_1 &= -\sum_n \varkappa^n \Delta W_n = \\ &= -\frac{\gamma m^2}{a(m-1)}\left[\varkappa^N C(x_N^*) + \sum_{n=1}^{N-1} \varkappa^n (C(x_n^*) - \varkappa C(x_{n+1}^*))\right] = -\frac{\gamma \varkappa m^2}{a(m-1)} C(x_1^*) \end{aligned}$$

□

Proof of Proposition 4.3. From Proposition 4.2 for $n \neq N$ and a separable output function

$$\Delta W_n = \frac{\gamma m^2}{a(m-1)}(C(x_n^*) - \varkappa C(x_{n+1}^*)),\ n \neq N,$$

$$\frac{\partial \Pi_n(.)}{\partial x_n} = C'(x_n^*).$$

The second equality can be equivalently rewritten as

$$x_n^* \frac{\partial \Pi_n(.)}{\partial x_n} = \gamma C(x_n^*).$$

Assume that properties in part (1) hold in a non-strict sense. Then, we have

$$\begin{aligned} C(x_n^*) &= x_n^* \frac{\partial \Pi_n(x_n^*, ..., x_n^*)}{\partial x_n} / \gamma \geq \varkappa x_n^* \frac{\partial \Pi_{n+1}(x_n^*, ..., x_n^*)}{\partial x_{n+1}} / \gamma \geq \\ &\geq \varkappa x_{n+1}^* \frac{\partial \Pi_{n+1}(x_{n+1}^*, ..., x_{n+1}^*)}{\partial x_{n+1}} / \gamma = \varkappa C(x_{n+1}^*). \end{aligned}$$

Thus, $\Delta W_n \geq 0$ for $n \neq N$.

In the final $\Delta W_N = \frac{\gamma m^2}{a(m-1)} C(x_N^*)$, which is always non-negative.

The proof for the case with strict inequalities in part (1) and that for the whole part (2) are similar.

□

Proof of Lemma 4.2.

$$\frac{P'_{x_{i,n}}(e,\mathbf{e})}{P(e,\mathbf{e})} = \frac{f'(e)\sum_{j\neq i} f(e)}{(f(e)+\sum_{j\neq i} f(e))^2} / \left(\frac{f(e)}{f(e)+\sum_{j\neq i} f(e)}\right) =$$
$$= \frac{m-1}{m}\frac{f'(e)}{f(e)} = \frac{m-1}{m}\frac{m}{m-1}\frac{C^{\frac{m}{m-1}-1}(e)}{C^{\frac{m}{m-1}}(e)}C'(e) = \frac{C'(e)}{C(e)}$$

□

Proof of Proposition 4.5. The agent's problem is

$$V_{i,n} = \max_{x_{i,n}}(1 - P(x_{i,n}, \mathbf{x}_{-i,n}))W_n + P(x_{i,n}, \mathbf{x}_{-i,n})V_{i,n+1} - C(x_{i,n}).$$

Let us assume that $V_{i,n+1} - W_n \geq 0$ and the solution is interior and show later that this is true. Then,

$$P'_{x_{i,n}}(x_{i,n}, \mathbf{x}_{-i,n})(V_{i,n+1} - W_n) = C'(x_i).$$

Denote the interior symmetric solution to this equation by x^*_n. Then, it satisfies the following:

$$P'_{x_{i,n}}(x^*_n, \mathbf{x}^*_n)(V_{n+1} - W_n) = C'(x^*_n).$$

From Lemma 4.2 if $P(x_{i,n}, \mathbf{x}_{-i,n}) = \frac{f(x_{i,n})}{f(x_{i,n})+\sum_{j\neq i} f(x_{j,n})}$, where $f(x) = C^{\frac{m}{m-1}}(x)$, then $\frac{P'_{x_{i,n}}(e,\mathbf{e})}{P(e,\mathbf{e})} = \frac{C'(e)}{C(e)}$

Thus,

$$P(x^*_n, \mathbf{x}^*_n)(V_{n+1} - W_n) = C(x^*_n).$$

Then, due to symmetry of the CSF,

$$C(x^*_n) = \frac{1}{m}(V_{n+1} - W_n).$$

Substituting this in the value function we obtain

$$V_n = (1 - \frac{1}{m})W_n + \frac{1}{m}V_{n+1} - \frac{1}{m}(V_{n+1} - W_n) = W_n.$$

Thus, agents' valuations of the participation in each stage would be exactly equal to the prize at that stage. Thus, this solution coincides with the solution obtained in Lemma 4.1 for $a = \frac{m}{m-1}\gamma$. To argue that this interior solution is an equilibrium we could just repeat the argument from the proof of Lemma 4.1 for $a = \frac{m}{m-1}\gamma$.

Then,

$$C(x_n^*) = \frac{1}{m}(W_{n+1} - W_n) = \frac{1}{m}\Delta W_n$$

Thus, the level of effort at the particular stage depends only on the prize increase at that stage.

Now, the designer can optimize with respect to the prize structure:

$$\Pi(\mathbf{x}_1(\Omega), ..., \mathbf{x}_N(\Omega)) - m^N W_1 - (\sum_n m^{N-n}\Delta W_n) \Longrightarrow \max_{\Omega},$$

$$s.t. \text{ limited liability: } W_i \geq 0.$$

We must notice here that the participation constraint is automatically satisfied in the case with non-negative prizes because agents always have an opportunity to apply zero level of effort.

Now, we assume that there is no limited liability constraint and $W_1 = 0$. If the solution for this reduced problem satisfies limited liability, we have the solution to the whole problem.

F.O.C. for the reduced problem is

$$m^{N+1-n}\frac{\partial\Pi(\mathbf{x}_1(\Omega), ..., \mathbf{x}_N(\Omega))}{\partial x_{i,n}}\frac{\partial x_n^*}{\partial \Delta W_n} = m^{N-n}.$$

The response of the effort to the change of a prize is

$$\frac{\partial x_n^*}{\partial \Delta W_n} = \frac{1}{m}\frac{1}{C'(x_n^*)}.$$

We substitute the last equation in the F.O.C.

$$m^{N+1-n}\frac{\partial\Pi(\mathbf{x}_1(\Omega), ..., \mathbf{x}_N(\Omega))}{\partial x_{i,n}}\frac{1}{m}\frac{1}{C'(x_n^*)} = m^{N-n}.$$

Hence,

$$\frac{\partial\Pi(\mathbf{x}_1(\Omega), ..., \mathbf{x}_N(\Omega))}{\partial x_{i,n}} = C'(x_n^*).$$

Thus, the equilibrium effort level is efficient. Simultaneously, the designer obtains the whole surplus from the agents: $V_1 = W_1 = 0$. Thus, we have got the solution to the reduced problem, which implements the efficient level of effort and extracts the whole surplus. Since all prize spreads are non-negative, the limited liability restriction is satisfied. Hence, the reduced solution is also the solution to the whole problem and the proposed structure is optimal. □

Bibliography

Akcigit, U. and Q. Liu (2015) "The role of information in innovation and competition," *Journal of the European Economic Association.*

Aoyagi, M. (2010) "Information feedback in a dynamic tournament," *Games and Economic Behavior*, Vol. 70, No. 2, pp. 242–260.

Arrow, K. (1962) *Economic welfare and the allocation of resources for invention*, Princeton: Princeton University Press, pp.609–626.

Azrieli, Yaron and Ritesh Jain (2018) "Symmetric mechanism design," *Journal of Mathematical Economics*, Vol. 74, pp. 108–118.

Baye, M.R. and H.C. Hoppe (2003) "The strategic equivalence of rent-seeking, innovation, and patent-race games," *Games and economic behavior*, Vol. 44, No. 2, pp. 217–226.

Baye, M.R., D. Kovenock, and C.G. De Vries (1996) "The all-pay auction with complete information," *Economic Theory*, Vol. 8, No. 2, pp. 291–305.

Bayus, B.L., S. Jain, and A.G. Rao (2001) "Truth or consequences: An analysis of vaporware and new product announcements," *Journal of Marketing Research*, Vol. 38, No. 1, pp. 3–13.

Bimpikis, K., S. Ehsani, and M. Mostagir (2014) "Designing dynamic contests," *Stanford University, Working paper.*

Bos, O. (2012) "Wars of attrition and all-pay auctions with stochastic competition," *Journal of Mathematical Economics*, Vol. 48, No. 2, pp. 83–91.

Caillaud, Bernard and Philippe Jehiel (1998) "Collusion in auctions with externalities," *The RAND Journal of Economics*, pp. 680–702.

Casas-Arce, P. and F. Martinez-Jerez (2011) "Handicaps in relative performance compensation: An all-pay auction approach," *Mimeo.*

Chakravarty, Surajeet and Todd R Kaplan (2013) "Optimal allocation without transfer payments," *Games and Economic Behavior*, Vol. 77, No. 1, pp. 1–20.

Che, Y.K. and I.L. Gale (1998) "Caps on political lobbying," *American Economic Review*, Vol. 88, No. 3, pp. 643–651.

——— (2003) "Optimal design of research contests," *American Economic Review*, Vol. 93, No. 3, pp. 646–671.

Che, Yeon-Koo and Jinwoo Kim (2006) "Robustly Collusion-Proof Implementation," *Econometrica*, Vol. 74, No. 4, pp. 1063–1107.

——— (2009) "Optimal collusion-proof auctions," *Journal of Economic Theory*, Vol. 144, No. 2, pp. 565–603.

Chung, Tai-Yeong (1996) "Rent-seeking contest when the prize increases with aggregate efforts," *Public Choice*, Vol. 87, No. 1, pp. 55–66.

Clark, Derek J and Christian Riis (1998) "Contest success functions: an extension," *Economic Theory*, Vol. 11, No. 1, pp. 201–204.

Cohen, C. and A. Sela (2007) "Contests with ties," *BE Journal of Theoretical Economics*, Vol. 7, No. 1.

Condorelli, Daniele (2012) "What money canÊijt buy: Efficient mechanism design with costly signals," *Games and Economic Behavior*, Vol. 75, No. 2, pp. 613–624.

CoST (2012) "Openness and accountability in public infrastructure could save $2.5 trillion by 2020," *Press Release*.

Deb, Rahul and Mallesh M Pai (2017) "Discrimination via symmetric auctions," *American Economic Journal: Microeconomics*, Vol. 9, No. 1, pp. 275–314.

Doraszelski, U. (2003) "An R&D race with knowledge accumulation," *RAND Journal of Economics*, pp. 20–42.

Ederer, F. (2010) "Feedback and motivation in dynamic tournaments," *Journal of Economics & Management Strategy*, Vol. 19, No. 3, pp. 733–769.

Erkal, N. and J. Xiao (2015) "Awarding scarce ideas in innovation contests," *Mimeo*.

European Comission (2007) "Pre-commercial Procurement: Driving innovation to ensure sustainable high quality public services in Europe."

——— (2014) "Directive on public procurement."

Fu, Qiang and Jingfeng Lu (2012) "The optimal multi-stage contest," *Economic Theory*, pp. 1–32.

Fudenberg, D., R. Gilbert, J. Stiglitz, and J. Tirole (1983) "Preemption, leapfrogging and competition in patent races," *European Economic Review*, Vol. 22, No. 1, pp. 3–31.

Fullerton, R.L. and R.P. McAfee (1999) "Auctionin entry into tournaments," *Journal of Political Economy*, Vol. 107, No. 3, pp. 573–605.

Gershkov, Alex, Jacob K Goeree, Alexey Kushnir, Benny Moldovanu, and Xianwen Shi (2013) "On the equivalence of Bayesian and dominant strategy implementation," *Econometrica*, Vol. 81, No. 1, pp. 197–220.

Gilbert, R.J. (2006) *Looking for Mr. Schumpeter: Where are we in the competition-innovation debate?*, Cambridge, MA: MIT Press, pp.159–215.

Gilbert, R.J. and D.M. Newbery (1982) "Preemptive patenting and the persistence of monopoly," *American Economic Review*, pp. 514–526.

Gill, D. (2008) "Strategic disclosure of intermediate research results," *Journal of Economics & Management Strategy*, Vol. 17, No. 3, pp. 733–758.

Goltsman, M. and A. Mukherjee (2011) "Interim performance feedback in multistage tournaments: The optimality of partial disclosure," *Journal of Labor Economics*, Vol. 29, No. 2, pp. 229–265.

Gradstein, Mark and Kai A Konrad (1999) "Orchestrating rent seeking contests," *The Economic Journal*, Vol. 109, No. 458, pp. 536–545.

Graham, Daniel A and Robert C Marshall (1987) "Collusive bidder behavior at single-object second-price and English auctions," *Journal of Political economy*, Vol. 95, No. 6, pp. 1217–1239.

Grossman, G.M. and C. Shapiro (1987) "Dynamic R&D competition," *Economic Journal*, Vol. 97, No. 386, pp. 372–387.

Gürtler, Oliver and Matthias Kräkel (2010) "Optimal tournament contracts for heterogeneous workers," *Journal of Economic Behavior & Organization*, Vol. 75, No. 2, pp. 180–191.

Halac, Marina, Navin Kartik, and Qingmin Liu (2017) "Contests for experimentation," *Journal of Political Economy*, Vol. 125, No. 5, pp. 1523–1569.

Harris, C. and J. Vickers (1985a) "Patent races and the persistence of monopoly," *Journal of Industrial Economics*, pp. 461–481.

——— (1985b) "Perfect equilibrium in a model of a race," *Review of Economic Studies*, Vol. 52, No. 2, pp. 193–209.

——— (1987) "Racing with uncertainty," *Review of Economic Studies*, Vol. 54, No. 1, pp. 1–21.

Hillman, A.L. and J.G. Riley (1989) "Politically contestable rents and transfers*," *Economics & Politics*, Vol. 1, No. 1, pp. 17–39.

Judd, K.L. (2003) "Closed-loop equilibrium in a multi-stage innovation race," *Economic Theory*, Vol. 21, No. 2-3, pp. 673–695.

Kaplan, T.R., I. Luski, and D. Wettstein (2003) "Innovative activity and sunk cost," *International Journal of Industrial Organization*, Vol. 21, No. 8, pp. 1111–1133.

Kirkegaard, R. (2012) "Favoritism in asymmetric contests: Head starts and handicaps," *Games and Economic Behavior*, Vol. 76, No. 1, pp. 226–248.

Kolmar, Martin and Dana Sisak (2014) "(In) efficient public-goods provision through contests," *Social Choice and Welfare*, Vol. 43, No. 1, pp. 239–259.

Konrad, K.A. (2002) "Investment in the absence of property rights; the role of incumbency advantages," *European Economic Review*, Vol. 46, No. 8, pp. 1521–1537.

Konrad, K.A. and W. Leininger (2007) "The generalized Stackelberg equilibrium of the all-pay auction with complete information," *Review of Economic Design*, Vol. 11, No. 2, pp. 165–174.

Konrad, Kai A et al. (2009) "Strategy and dynamics in contests," *OUP Catalogue*.

Krähmer, Daniel and Roland Strausz (2015) "Optimal sales contracts with withdrawal rights," *The Review of Economic Studies*, Vol. 82, No. 2, pp. 762–790.

Krishna, V. and J. Morgan (1998) "The winner-take-all principle in small tournaments," *Advances in Applied Microeconomics*, Vol. 7, pp. 61–74.

Lang, M., C. Seel, and P. Strack (2014) “Deadlines in stochastic contests,” *Journal of Mathematical Economics*, Vol. 52, pp. 134–142.

Lazear, Edward P and Sherwin Rosen (1981) “Rank-order tournaments as optimum labor contracts,” *Journal of political Economy*, Vol. 89, No. 5, pp. 841–864.

Lee, T. and L.L. Wilde (1980) “Market structure and innovation: A reformulation,” *Quarterly Journal of Economics*, Vol. 94, No. 2, pp. 429–436.

Leininger, W. (1991) “Patent competition, rent dissipation, and the persistence of monopoly: The role of research budgets,” *Journal of Economic Theory*, Vol. 53, No. 1, pp. 146–172.

Lippman, S.A. and J. McCall (1976) “The economics of job search: A survey,” *Economic inquiry*, Vol. 14, No. 2, pp. 155–189.

Lippman, S.A. and K.F. McCardle (1987) “Dropout behavior in R&D races with learning,” *Rand Journal of Economics*, pp. 287–295.

Loury, G.C. (1979) “Market structure and innovation,” *Quarterly Journal of Economics*, pp. 395–410.

Mailath, George J and Peter Zemsky (1991) “Collusion in second price auctions with heterogeneous bidders,” *Games and Economic Behavior*, Vol. 3, No. 4, pp. 467–486.

Manelli, Alejandro M and Daniel R Vincent (2010) “Bayesian and Dominant-Strategy Implementation in the Independent Private-Values Model,” *Econometrica*, Vol. 78, No. 6, pp. 1905–1938.

Marshall, Robert C and Leslie M Marx (2007) “Bidder collusion,” *Journal of Economic Theory*, Vol. 133, No. 1, pp. 374–402.

Maskin, Eric and JJ Laffont (1979) “A differential approach to expected utility maximizing mechanisms,” *Aggregation and revelation of preferences*.

Maskin, Eric and Jean Tirole (1990) “The principal-agent relationship with an informed principal: The case of private values,” *Econometrica: Journal of the Econometric Society*, pp. 379–409.

——— (1992) “The principal-agent relationship with an informed principal, II: Common values,” *Econometrica: Journal of the Econometric Society*, pp. 1–42.

McAfee, R Preston and John McMillan (1992) “Bidding rings,” *The American Economic Review*, pp. 579–599.

Moldovanu, B. and A. Sela (2001) "The optimal allocation of prizes in contests," *American Economic Review*, pp. 542–558.

Moldovanu, Benny, Aner Sela, and Xianwen Shi (2012) "Carrots and sticks: prizes and punishments in contests," *Economic Inquiry*, Vol. 50, No. 2, pp. 453–462.

Myerson, Roger B (1981) "Optimal auction design," *Mathematics of operations research*, Vol. 6, No. 1, pp. 58–73.

——— (1983) "Mechanism design by an informed principal," *Econometrica: Journal of the Econometric Society*, pp. 1767–1797.

Mylovanov, Tymofiy and Thomas Tröger (2012) "Informed-principal problems in environments with generalized private values," *Theoretical Economics*, Vol. 7, No. 3, pp. 465–488.

——— (2014) "Mechanism design by an informed principal: Private values with transferable utility," *The Review of Economic Studies*, Vol. 81, No. 4, pp. 1668–1707.

OECD (2014) "OECD Foreign Bribery Report-An Analysis of the Crime of Bribery of Foreign Public Officials."

Parreiras, Sérgio O and Anna Rubinchik (2010) "Contests with three or more heterogeneous agents," *Games and Economic Behavior*, Vol. 68, No. 2, pp. 703–715.

PricewaterhouseCoopers and Ecorys (2013) "Identifying and reducing corruption in public procurement in the EU."

Reinganum, J.F. (1981) "Dynamic games of innovation," *Journal of Economic Theory*, Vol. 25, No. 1, pp. 21–41.

——— (1982) "A dynamic game of R&D: Patent protection and competitive behavior," *Econometrica*, pp. 671–688.

——— (1983) "Uncertain innovation and the persistence of monopoly," *American Economic Review*, Vol. 73, No. 4, pp. 741–748.

——— (1985) "Innovation and industry evolution," *Quarterly Journal of Economics*, pp. 81–99.

Rieck, T. (2010) "Information disclosure in innovation contests," *Bonn Econ Discussion Papers*.

Robinson, Marc S (1985) "Collusion and the Choice of Auction," *The RAND Journal of Economics*, pp. 141–145.

Rosen, Sherwin (1986) "Prizes and Incentives in Elimination Tournaments," *The American Economic Review*, pp. 701–715.

Ryvkin, Dmitry (2007) "Tullock contests of weakly heterogeneous players," *Public Choice*, Vol. 132, No. 1, pp. 49–64.

Schöttner, A. (2008) "Fixed-prize tournaments versus first-price auctions in innovation contests," *Economic Theory*, Vol. 35, No. 1, pp. 57–71.

Schumpeter, J.A. (1934) *The theory of economic development*, Cambridge, MA.

——— (1942) *Capitalism, socialism and democracy*, New York.

Schweinzer, Paul and Ella Segev (2012) "The optimal prize structure of symmetric Tullock contests," *Public Choice*, Vol. 153, No. 1, pp. 69–82.

Seel, C. (2014) "The value of information in asymmetric all-pay auctions," *Games and Economic Behavior*, Vol. 86, pp. 330–338.

Seel, C. and P. Strack (2013) "Gambling in contests," *Journal of Economic Theory*, Vol. 148, No. 5, pp. 2033–2048.

Seel, Christian and Philipp Strack (2016) "Continuous time contests with private information," *Mathematics of Operations Research*, Vol. 41, No. 3, pp. 1093–1107.

Seel, C. and C. Wasser (2014) "On optimal head starts in all-pay auctions," *Economics Letters*, Vol. 124, No. 2, pp. 211–214.

Segev, E. and A. Sela (2014) "Sequential all-pay auctions with head starts," *Social Choice and Welfare*, Vol. 43, No. 4, pp. 893–923.

Severinov, Sergei (2008) "An efficient solution to the informed principal problem," *Journal of Economic Theory*, Vol. 141, No. 1, pp. 114–133.

Shaffer, Sherrill (2006) "War, labor tournaments, and contest payoffs," *Economics Letters*, Vol. 92, No. 2, pp. 250–255.

Siegel, R. (2009) "All-pay contests," *Econometrica*, Vol. 77, No. 1, pp. 71–92.

——— (2010) "Asymmetric contests with conditional investments," *American Economic Review*, Vol. 100, No. 5, pp. 2230–2260.

——— (2014) "Asymmetric contests with head starts and nonmonotonic costs," *American Economic Journal: Microeconomics*, Vol. 6, No. 3, pp. 59–105.

Skaperdas, Stergios (1996) "Contest success functions," *Economic theory*, Vol. 7, No. 2, pp. 283–290.

Taylor, C.R. (1995) "Digging for golden carrots: An analysis of research tournaments," *American Economic Review*, pp. 872–890.

Thomas, Jonathan P and Zhewei Wang (2013) "Optimal punishment in contests with endogenous entry," *Journal of Economic Behavior & Organization*, Vol. 91, pp. 34–50.

Tirole, J. (1997) *The theory of industrial organization*, Cambridge, MA: MIT press.

Ulhøi, J.P. (2004) "Open source development: A hybrid in innovation and management theory," *Management Decision*, Vol. 42, No. 9, pp. 1095–1114.

Wirtz, J. (2013) "Feedback and learning in tournaments," *Mimeo*.

WTO (2011) "Agreement on Government Procurement."

Yilankaya, Okan (1999) "A note on the seller's optimal mechanism in bilateral trade with two-sided incomplete information," *Journal of Economic Theory*, Vol. 87, No. 1, pp. 267–271.

Yildirim, H. (2005) "Contests with multiple rounds," *Games and Economic Behavior*, Vol. 51, No. 1, pp. 213–227.