Essays in Contest Theory

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Economics

By

Jorge Francisco Catepillan Tessi

EVANSTON, ILLINOIS

December 2018
Contest theory is an area of game theory that studies environments in which agents make sunk investments in order to get a prize. These investments could be money, effort, time, etc. Contest theory is used to study a wide range of applications, like political contests, research and development, advertisement campaigns, rent-seeking, among others. In this dissertation, I study different features of three different contests.

Chapter 1 is coauthored with Ron Siegel. In this Chapter, we study how information affects the ex-ante expected payoff of two players in an all-pay auction. This auction is considered a winner-takes-all contest, as only one player gets the full prize. By information we mean how private information is distributed. Understanding how information affects the ex-ante utility of players allows us to understand the incentives players have to participate in this type of contests. We find there are two characteristics of a distribution that has a high impact in the ex-ante utility of a player. First, how strong a player can be, and second, how informative the distribution is. We find that when the other player has a distribution with high types, it is better to have a non-informative distribution than one
with weights in high types. Also, we find that in a setting where one bidder is informed and the other is not, the distribution that maximizes ex-ante payoff for the informed player is one that has weights in at most two points, which are always the highest and lowest possible from the ones available to choose from. Therefore, such a distribution is not necessarily one that has the highest expectation, nor the lowest risk. The intuition behind this result is that in some cases there is information to gain from a distribution that has a high variance, because being ex-ante uninformed discourages aggressive bids by the other player, which offers the chance of being a low type when in fact the player is a high type bidder which results in a bigger payoff. Surprisingly, we find that such information can lead to very big gains, as in some cases the preferred distribution is very different from the one that maximizes the likelihood of winning.

In Chapter 2, which is a joint work with José Espín, we study a war of attrition with allocation externalities. A war of attrition is a game in which players have to decide when to quit, and their utility depends on how many players have quit before them. It has been used to study situations in which agents are willing to bear losses in anticipation of profitability following other players exiting. In that setting, we study a game with perfect information in continuous time, allowing for allocation externalities. This means that the players’ valuation of the prize depends on the identity of the player that doesn’t get a prize. As in the classical game, we find multiplicity of equilibria and conditions in order to get them. We find that most of the properties of the classic war of attrition still hold, but some equilibria, neglected by previous literature, also arise. We discuss how to use the results of our research when the marginal cost of time is not constant and/or the value of the prize is time dependent.
Finally, Chapter 3 is a joint work with Jorge Lemus. In this Chapter, we study sequential contests allowing for effort accumulation, period-specific prizes and cost of effort, and a contest success function that is not necessarily homogeneous in effort. In this setting, we find novel qualitative implications that are absent under the standard homogeneity assumption. First, we provide a condition that determines the equilibrium trade-off between competition and cost savings. We then show that rent dissipation and aggregate effort crucially depend on the shape of the contest success function and the relation between prizes and cost of effort among contests.
Acknowledgements

I am very grateful to my advisor, Jeff Ely, and my committee members, Wojciech Olszewski and Yingni Guo, for their helpful advice, guidance, and support throughout this project.

I thank my coauthors, Ron Siegel, José Espín and Jorge Lemus, for their insights and fruitful discussions. Ron and I are jointly responsible for Chapter 2, José and I are responsible for Chapter 2, and Jorge and I are responsible for Chapter 3.

Most importantly, I would especially like to thank my family. My wife, Carolina, has been extremely supportive of me through this long journey and helped me any way she could to get me to this point. My mother, María José, and siblings gave me continuous support, encouragement, and understanding. This endeavor would not have been possible without them on my side.
# Table of Contents

ABSTRACT 3

Acknowledgements 6

List of Tables 9

List of Figures 10

Chapter 1. Ex-ante Expected Utility in an All-Pay Auction 11
   1.1. Introduction 12
   1.2. Model 15
   1.3. Equilibrium 16
   1.4. Simple Case 20
   1.5. Optimal \( p_h \) 24
   1.6. Zero Payoff 25
   1.7. One side private information 29
   1.8. Future Work 31

Chapter 2. War of Attrition with Information Externalities 33
   2.1. Introduction 34
   2.2. Setting 40
   2.3. Equilibrium of the Stage Game 42
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.4.</td>
<td>Symmetric Valuations</td>
<td>54</td>
</tr>
<tr>
<td>2.5.</td>
<td>Independent Valuations</td>
<td>55</td>
</tr>
<tr>
<td>2.6.</td>
<td>Time-varying valuations</td>
<td>58</td>
</tr>
<tr>
<td>2.7.</td>
<td>Discussion</td>
<td>60</td>
</tr>
<tr>
<td>Chapter 3</td>
<td>Dynamic Tullock Contest</td>
<td>62</td>
</tr>
<tr>
<td>3.1.</td>
<td>Introduction</td>
<td>63</td>
</tr>
<tr>
<td>3.2.</td>
<td>Literature review</td>
<td>66</td>
</tr>
<tr>
<td>3.3.</td>
<td>Model</td>
<td>68</td>
</tr>
<tr>
<td>3.4.</td>
<td>Rent Dissipation and Aggregate Effort across Contests</td>
<td>79</td>
</tr>
<tr>
<td>3.5.</td>
<td>Extension: Head Starts with Variable Carryover rate</td>
<td>85</td>
</tr>
<tr>
<td>3.6.</td>
<td>Conclusion</td>
<td>87</td>
</tr>
</tbody>
</table>

References | 90 |

Appendix A. Ex-ante Expected Utility in an All-Pay Auction | 96 |
| A.1. | Appendix I | 96 |
| A.2. | Appendix II | 110 |

Appendix B. War of Attrition with Information Externalities | 116 |

Appendix C. Dynamic Tullock Contest | 120 |
List of Tables

2.1 Example: Payoffs 42
## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Equilibrium for different values of $v_h$, $p_h$ and $v$</td>
<td>21</td>
</tr>
<tr>
<td>1.2</td>
<td>Shape of $U_2$ as a function of $p_h$ for different values of $v$, $v_h$ and $v_l$.</td>
<td>23</td>
</tr>
<tr>
<td>1.3</td>
<td>Areas of $(v_h, v_l)$ in which different $U_2$ emerge</td>
<td>24</td>
</tr>
<tr>
<td>1.4</td>
<td>Optimal $p_h^*$</td>
<td>25</td>
</tr>
<tr>
<td>3.1</td>
<td>Best response function, for a static Tullock contest with parameter $\theta$.</td>
<td>71</td>
</tr>
<tr>
<td>3.2</td>
<td>Regions for different cases of second period payoffs.</td>
<td>73</td>
</tr>
<tr>
<td>3.3</td>
<td>Aggregate effort for the two contests as a function of $\delta$.</td>
<td>87</td>
</tr>
</tbody>
</table>
CHAPTER 1

Ex-ante Expected Utility in an All-Pay Auction
1.1. Introduction

An all-pay auction is an auction where the highest bidder gets the object, but everyone pays their own bid. Even though this is not a common auction, this game theoretical model has been used to study a wide range of economic applications that involve competition with sunk investments.

In this Chapter, we focus on the two players all-pay auction with incomplete information, and we try to understand how different distributions of types affect the ex-ante utility of players, where a high type implies a high valuation of the object. We do this in two different settings. First, when both players have private information and one player is much stronger than the other; where strong means that it can leave the other player with zero ex-ante expected utility. And second, when there is only one side private information. This allows us to understand what the incentives are to participate in this type of games.

In the first case, we find that a lower probability of being high type is better for the weak player.

In the second case, we find that if you can choose the support and the probabilities of types for the informed player to maximize ex-ante expected utility, the distribution chosen can only take two forms, the highest possible type with probability one, or with weights in both the highest and lowest types available to choose from.

The interesting thing about these results is that in both cases, what we consider to be a good distribution is not necessarily the best distribution ex-ante. In the first case, given that a player is "weak", having a first order dominated distribution is better. In the second case, having a higher variance and lower expected value of the distribution might be the best to have ex-ante.
Our interpretation of these results is the following: In the all-pay auction, much of the surplus is destroyed, mostly because players bid very aggressively. In this context, having a higher probability of a high type bidder has a positive effect, because a high valuation is more probable, and a negative effect, as in equilibrium both players will bid more aggressively. In the first case, if the other player is strong, the negative effect dominates.

For the second result, intuition is that when a player can be a really high type bidder, then that is the best distribution to have. How high depends on the distance between the type of informed and uninformed player, as in the all-pay auction with complete information.

However, if the informed player does not have a high type available to choose from, then the ex-ante utility can only be improved by having a positive probability of being a low type. This is due to a strategic effect. If there is a chance the informed player is a low type bidder, in equilibrium the uninformed player will bid less aggressively. That will benefit the informed player in case of being a high type bidder.

In other words, to maximize the ex-ante utility, the informed player has to be as "good" as possible, or have a chance of being as "bad" as possible.

These results can be applied to any situation in which the all-pay auction has been applied. For example, in R & D races, our results can be interpreted as a prediction that in the presence of a firm that is experimenting with a known technology, the best strategy

---

1Krishna and Morgan (1995) show that the all-pay auction when players’ signals are affiliated and symmetrically distributed generates higher expected revenues for the seller than a first price auction.

2In the all-pay auction with complete information and two players, both bid between zero and the lowest valuation, and the expected utility of the player with the highest valuation is increasing according to the difference of valuations.
for a competitor might be to go with a risky technology, even if the chances of success are lower than other technologies available.

An example of this might be the development of the V2 by the Nazis during World War II. Many experts have discussed the effectiveness of this strategy in order to win the war, and most argue that the high cost of the program and low accuracy of the bombs, and the existence of other technologies that were proved to be more effective (i.e. aircrafts) made it a bad strategic decision.

Freeman Dyson (1979), the renew physicist and mathematician, writes:

"... those of us who were seriously engaged in the war were very grateful to Wernher von Braun. We knew that each V-2 cost as much to produce as a high-performance fighter airplane. We knew that German forces on the fighting fronts were in desperate need of airplanes... From our point of view, the V-2 program was almost as good as if Hitler had adopted a policy of unilateral disarmament." (p. 108)

With our results, we add another dimension to the analyses. When judging how decisions are made, we must consider not only if those decisions make sense in a single agent way, but also how those decisions do well in the context of many players making other decisions. In this particular case, had the Nazis invested in traditional technology, it might have induced a much bigger response from the Allies. And maybe, by committing to this high risk and maybe low payoff strategy, Germany was avoiding that to happen.

---

3Zaloga (2003), on the effectiveness of the V2:

"The cost of the development and manufacture of the V-2 was staggering, estimated by a post-war US study as about $2 billion, or about the same amount as was spent on the Allied atomic bomb program. Yet the entire seven-month V-2 missile campaign delivered less high explosive on all the targeted cities than a single large RAF raid on Germany." (p. 36)

Wernher von Braun was one of the leading engineers of the development of the V2.
1.2. Model

Our model is a classical two player all pay auction with asymmetric information in a discrete type space. There are two players that compete for one prize. Each receives a signal \( s_i \) in \( S_i \subseteq \mathbb{R} \) with \( S_i \) finite and ordered. The joint distribution of \( S \) is described by \( f : S_i \times S_{-i} \rightarrow [0,1] \). We assume that all combination of signals have positive probability of happening.

Knowing the signals, each player bids (exerts effort) and the one with the higher bid(effort) wins the contest. The valuation of winning the auction for player \( i \) is \( V_i : S_i \times S_{-i} \rightarrow \mathbb{R}_+ \). In case we are in a private value model, this is equivalent to saying that the signals are the valuations.

Then, given signals \( s_i, s_{-i} \), and bids \( b_i, b_{-i} \), player \( i \)'s payoff is:

\[
\pi_i(b_i, b_{-i})V_i(s_i, s_{-i}) - b_i
\]

Where \( \pi_i(,,) \) is the probability \( i \) wins given the bids. It is 1 if \( b_i > b_{-i} \), 0 if \( b_i < b_{-i} \); or anything in \([0,1]\) such that \( \pi_i(,,) + \pi_j(,,) = 1 \) if \( b_i = b_{-i} \).

We will assume a monotinicity condition on the probabilities and payoffs:

**Assumption 1.** Monotonicity \((M)\), for \( i = 1, 2 \), \( f(s_{-i}|s_i)V_i(s_i, s_{-i}) \) is increasing in \( s_i \) for every \( s_{-i} \);
1.3. Equilibrium

As with the all pay auction with complete information, Siegel (2013) shows that this game does not have an equilibrium in pure strategies. Therefore, we have to focus on mixed strategies.

A strategy for player $i$ is going to be the cumulative distribution function of the bids player $i$ makes given her type $s_i$; $G_i(s_i, \cdot)$. Call, $BR_i(s_i, G_j)$ the set of best responses of player $i$ to strategy $G_j$ of player $j$ when $i$’s type is $s_i$. Then an equilibrium is going to be a pair of strategies $(G_1, G_2)$ such that $G_i$ assigns measure one to $BR_i(s_i, G_j)$

**Theorem 2.** (Siegel 2013) Under monotonicity, there is a unique equilibrium $(G^*_1, G^*_2)$ of this game, such that:

- There is no bid at which both players have an atom.
- There is no positive bid at which some player has an atom.
- If a positive bid is not a best response for some player or any of her signals, then no weakly higher bid is a best response for any signal.
- Each player has at least one signal for which 0 or bids arbitrary close to 0 are best response.
- Each best response set is an interval or a point, and for two consecutive signals, the upper bound of the best response of the low one is equal to the lower bound of the best response of the high one.

With those properties of the equilibrium, Siegel (2013) gives an algorithm to compute it. It works in the following way.
1.3.1. Algorithm to compute the Equilibrium

In order to make things easier, enumerate $S_i$ from the highest element, to the lowest, $S_i = \{s_i^{N_i}, ..., s_i^1\}$, where $N_i$ is the cardinality of $S_i$.

Given the properties of the eq. shown in Siegel (2013) we know that for each type, the best response set is going to be an interval. Because this intervals are glued to each other for one player, then we are going to have sub intervals where players face known types. In other words, in that sub intervals, players know which type of the other player will bid lower, higher, or is going to mix. Then, the algorithm helps to find all this sub intervals.

The algorithm to compute the equilibrium is the following.

- Step 1: Interval 1.

Define

\begin{equation}
  t_i^1 = s_i^1
\end{equation}

\begin{equation}
  t^1 = (t_1^1, t_2^1)
\end{equation}

We are going to get the density functions that make both players indifferent in case both have their respective highest type.

That density is

\begin{equation}
  \pi_i^1 = \frac{1}{v_j(\alpha_i) f_j(\alpha_j)}
\end{equation}
This because if player $i$ bids something that beats all types of player $j$ lower than $t_{ij}^1$, or something a little bigger, in order to be indifferent player $j$ must be bidding with a rate that increases probability in the same what cost of player $i$ increase.

Now, define:

$$d_1 = \min\{\frac{1}{\pi_{i}^1}, \frac{1}{\pi_{j}^1}\}$$  \hspace{1cm} (1.4)$$

$$p_{i}^1 = 1 - d_1 \cdot \pi_{i}^1$$  \hspace{1cm} (1.5)$$

Finally, if for some $i$, $p_{i}^1 = 0$ and there is no other element in $S_i$ that is lower than $t_{i}^1$, stop here. If that is the case, the equilibrium is such that:

(1) For player $j$, bid zero for every type $s_{j} < s_{i}^1$

(2) For player $j$, bid zero with probability $p_{j}^1$, and according to density $\pi_{j}^1$ in $[0, d_1]$

(3) For player $i$, bid with density $\pi_{i}^1$ in $[0, d_1]$.

If we are not in that case, then:

$$t_{i}^2 = \begin{cases} 
  t_{i}^1 & \text{if } p_{i}^1 \neq 0 \\
  s_{i}^2 & \text{o.w.}
\end{cases}$$  \hspace{1cm} (1.6)$$

And go to the next step.

- Step $k$: Interval $k$ In this interval, the only types that are mixing are $t_{i}^k$ and $t_{j}^k$. Call $t^k = (t_{i}^k, t_{j}^k)$. 
Then, because in equilibrium the types that are mixing here is known to player, the density each one will be using is:

\[ \pi_i^k = \frac{1}{V_j(t^k)J_i(t^k)} \]

Define

\[ d_k = \min\{\frac{1-p_i^{k-1}}{\pi_1^k}, \frac{1-p_i^{k-1}}{\pi_2^k}\} \]  
\[ p_i^k = 1 - p_i^{k-1} - d_k \cdot \pi_i^k \]

If for player \( i \), \( p_i^k \) is zero, and there is no type lower than \( t_i^k \) available. Finish the loop. Otherwise

If we are not in that case, then define:

\[ t_i^{k+1} = \begin{cases} t_i^k & \text{if } p_i^1 \neq 0 \\ \bar{s}_i & \text{o.w.} \end{cases} \]

Where \( \bar{s}_i \) is the highest type of \( i \) in \( S_i \) that is lower than \( t_i^k \).

Go to step \( k + 1 \).

- Eventually, you are going to end at some step \( \hat{k} \). To compute the equilibrium first define:

\[ I_{\hat{k}} = (0, d_{\hat{k}}] \]
\[ I_k = (\sum_{l=k+1}^{\hat{k}} d_l, \sum_{l=k}^{\hat{k}} d_l], \forall k < \hat{k} \]
The equilibrium strategy for player $i$ is going to be

$$G_i(s_i, x) = \begin{cases} 
0 & \text{if } x < 0 \\
1 & \text{if } x = 0, s_i \in S_i, s_i < t^k_i \\
1 - \sum_{l \leq k, s_i = t^l_i} \pi^k_i & \text{if } x = 0, s_i \in S_i, s_i \geq t^k_i \\
1 + x \cdot \pi^k_i - \sum_{l > k, s_i = t^l_i} \pi^l_i & \text{if } x \in I_k, s_i \in S_i, s_i \geq t^k_i 
\end{cases}$$

(1.13)

What the previous expression is saying is that the distributions of bids given $i$’s type can have two different forms:

(1) When $i$’s type is too low, she that will never bid against $j$, then the equilibrium bid for player $i$ is to bid zero if that is her type.

(2) When her type is big enough, then on each interval, she will bid with the density function that makes $j$ indifferent of bidding in that interval.

1.4. Simple Case

Before getting into a more complicated setting, we study a simple case in which player 1 has only one signal, and player 2 has signal $H$ with probability $p_h$ and $L$ with probability $p_l$. To make things easier, we assume $V_1(s, H) = V_1(s, L) = v$ and that $V_2(s, H) = v_h$, and $V_2(s, L) = v_l$ with $v_l < v_h$ and $f(s, L) = p_l, f(s, H) = p_h$.

Depending on how $p_h v$ and $v_h$ compare, we can have two types of equilibria, illustrated in Figure 1.1. There, each line represents the interval for a Player, $g_i$ is the equilibrium
Case 1, \( p_h v \geq v_h \)

\[
g_1 = \frac{1}{v_h} \]

\[
g_2 = \frac{1}{(p_h v)} \]

Only type \( v_h \) mixes

Length = \( v_h \)

Interval

Case 2 and 3, \( p_h v \leq v_h \)

\[
g_1 = \frac{1}{v_l} \]

\[
g_2 = \frac{1}{(p_l v)} \]

Type 2 = \( v_l \)

Length 1 = \( d^* \)

Length 2 = \( p_h v \)

Figure 1.1. Equilibrium for different values of \( v_h, p_h \) and \( v \)

density function in that interval, and the length is either \( v_h \) or \( d^* + p_h v \) with \( d^* = \min \left[ v_l \left( 1 - \frac{p_h v}{v_h} \right) , p_l v \right] \).

1.4.1. Ex-ante utility

Given those strategies, we can show that the ex-ante utility of player 2 is going to be:

\[
U_2 = \begin{cases} 
0 & \text{if } v_h \leq p_h v \\
\left( 1 - \frac{v}{v_h} \right) (v_h - v_l) & \text{if } p_h \leq \min \left\{ \frac{v_h}{v} , \frac{v_h - v_l}{v_h - v_l} \right\} \\
p_l (v_l - p_h \frac{v_h}{v} - p_l v) + p_h (v_h - v) & \text{o.w.} 
\end{cases}
\]

The second term is a parabola, and the third term is always increasing in \( p_h \). That can be easily verified. After studying this function, we can say that there are 3 areas of interest:

(1) \( v \geq v_h \). If that is the case, then \( U_2 \) is either zero, or like the second term.
(2) $v < v_h$ and $(v_l + v_h) \leq 2v$. If that is the case, then $U_2$ can only take the form of the second and third term, but with the kink of the second term happening before it changes to the third term.

(3) $v < v_h$ and $(v_l + v_h) > 2v$. Here, we are as before, but the kink of the second term does not occur before the change.

Figure 1.2 shows how the utility is affected by $p_h$ in each of this cases, and Figure 1.3 shows how different combinations of $(v_h, v_l)$ leave you in each of the cases. Note that we have divided Case 2 in two, for reasons we will explain later.

This two figures are key in understanding our results. We can see that for low values of $v_h$ relative to $v$, Player 2’s ex-ante utility can only be affected by $p_h$ because of informational advantages. That is, there is no way Player 2 can be "better" than Player 1. Then, for many cases having a less informative distribution is the only thing that can increase her payoff.

For high values of $v_h$ relative to $v$, Player 2 can become stronger by putting more weight in her distribution on $v_h$. Thus, there is no informational gain that can overcome that. It is always better to put more weight in the high type.

However, things become more interesting when $v_h$ is greater than $v$ but not for much. In that case, both strategies are good ("informative advantage" and being "strong"). At least locally. Then for some cases it is better to use the informational advantage, and for others just be as "strong" as possible.
\begin{align*}
(a) & \quad v \geq v_h
\end{align*}

\begin{align*}
(b) & \quad v < v_h \text{ and } (v_l + v_h) \leq 2v, \text{ for } v \text{ close to } v_h
\end{align*}

\begin{align*}
(c) & \quad v < v_h \text{ and } (v_l + v_h) \leq 2v, \text{ for } v \text{ far from } v_h
\end{align*}

\begin{align*}
(d) & \quad v < v_h \text{ and } (v_l + v_h) > 2v
\end{align*}

Figure 1.2. Shape of $U_2$ as a function of $p_h$ for different values of $v$, $v_h$ and $v_l$. 
Figure 1.3. Areas of \((v_h, v_l)\) in which different \(U_2\) emerge

1.5. Optimal \(p_h\)

From our previous analysis, we can see that \(p_h^*\) is going to be either 1 or \(\frac{v_h}{2v}\), which is the kink of the parabola. The area in which either of this two values is the optimal is represented in Figure 1.4.

As we explained before, the key is that in order to choose, for small values of \(v_h\) it only picks whatever maximizes the "informational advantage". When \(v_h\) is big, it is optimal to pick a distribution with the highest possible weight in \(v_h\). On the other cases, the optimal \(p_h\) comes from comparing the max informational advantage she can get to the utility of a distribution with all weight on \(p_h\).

Because the last utility depends on \(v_h - v\), for close values of \(v_h\) and \(v\), it is always better to have some uncertainty to maximize ex-ante utility.
If the contest doesn’t have private information, at least one of the players has an expected payoff of zero before entering. In this section, we want to study when a player gets zero expected payoff in the presence of private information, and how her type distribution affects that. What we do is to get necessary and sufficient condition for a contestant to have zero ex-ante expected payoff, and then study how that conditions change with the distribution of types.

The first result is a consequence of the Monotonicity assumption:

**Corollary 3.** In equilibrium, the expected payoff of a player given its signal is non decreasing on its signal.

**Proof.** Take two consecutive signals \( s_i < s'_i \). Suppose both signals induce a positive probability of a positive effort. Then, the supremum of the best response set of the lower one, and the infimum of the best response set of the high signal are equal. Call that point
Then, the expected utility in equilibrium given the small signal is

\[ \mathbb{E}[P|s_i] = \sum_{s_{-i} \in S_{-i}} f(s_{-i}|s_i)V_i(s_i, s_{-i})G_{-i}(x, s_{-i}) - x \]

\[ \leq \sum_{s_{-i} \in S_{-i}} f(s_{-i}|s'_i)V_i(s'_i, s_{-i})G_{-i}(x, s_{-i}) - x \]

\[ = \mathbb{E}[P|s'_i] \]

We used our assumption of monotonicity.

If none of the signals induce a positive probability of a positive bidding, the statement holds trivially (both have expected payoff of zero). Same if only the higher signal induces positive effort.

\[ \square \]

This corollary is very useful, because it indirectly states that a player will have an ex-ante expected payoff of zero if and only if her payoff is zero when she gets her highest possible signal. With this, is easier to get conditions for a zero expected payoff.

**Lemma 4.** Player \( i \) will have zero expected payoff if and only if:

\[ \mathbb{E} \left[ \frac{V_i(\tilde{s}_i, S_{-i})}{V_{-i}(S_{-i}, \tilde{s}_i)} \right] \leq \mathbb{P}[S_i = \tilde{s}_i] \]

where \( \tilde{s}_i \) is the highest signal for \( i \).

**Proof.** Using the previous corollaries, the only thing we need to show is that player \( i \) has zero expected payoff conditional on getting the highest signal. In such case, her expected payoff is
Where $T$ is the sup of the union of all best response sets. This is the payoff when player $i$ decides to exert effort $T$. If she does it, then she wins with probability one, and gets the expected payoff we put there.

Note that if that is the case, then bidding zero is in the best response set for a player $i$. However, because of Theorem 4.1, that implies that for all signals lower than $\bar{s}_i$, player $i$ bids zero.

In order for that to happen, we have to end the algorithm because there are no more types of $-i$ to keep going. In other words, that implies that all types of $-i$ make a positive bid with positive probability.

The condition for that to happen is that:

\[
\sum_{s_{-i} \in S_{-i}} f(s_{-i}|\bar{s}_i)V_i(\bar{s}_i, s_{-i}) \leq 1
\]

The equation can be re-written as:

\[
\frac{1}{f(\bar{s}_i)} \sum_{s_{-i} \in S_{-i}} \frac{f(s_{-i})V_i(\bar{s}_i, s_{-i})}{V_{-i}(s_{-i}, \bar{s}_i)} \leq 1
\]

Or:

\[
\mathbb{E}\left[\frac{V_i(\bar{s}_i, S_{-i})}{V_{-i}(S_{-i}, \bar{s}_i)}\right] \leq \mathbb{P}[S_i = \bar{s}_i]
\]
Fix the distribution of player \(-i\). Assuming \(V_i(s_i, S_{-i})\) is strictly increasing in \(s_i\), the distribution of \(S_i\) affects the condition in two ways.

If the L.H.S. is less than one, then a high probability of being the high type leaves player 1 with an expected payoff of zero. This is similar to the full information case, if player \(-i\) is strong enough (has higher valuations than \(i\)) and player \(i\) distribution is close enough to one with all the weight in \(s_i\), then player \(i\) in equilibrium has an expected payoff of 0.

However, if \(s_i\) increases, it is harder for the L.H.S. to be less than one, and also the probability of being high type must be even bigger in order to have player \(i\) with expected payoff of zero.

This is very intuitive. If given its own signal, player \(-i\) believes that there is a high probability on \(i\) having a high signal, then she will make an effort according to that. Therefore, if she has an advantage (can bid more than player \(i\)), she will make an effort as of player \(i\) is always high type, leaving her with zero expected payoff. However, if there is a small chance on \(i\) being high type, then player \(-i\) expected payoff of using its advantage becomes to costly (because the advantage make sense only if player \(i\) is high). Therefore, she will not use it. This will make player 1 to have positive payoffs in case she is high type.

Also, in the presence of a strong players, a distribution that has a low probability of the high type might be better than a distribution with all the probability there. In terms of participation on the contest, if there was a game in which constants are picked to participate before it, then “bad distributions” contestants have more incentives to participate
than better ones. Therefore, the eq. induces some sort of ex-ante discouragement effect on the potentially good contestants.

We explore more about this on the next section.

1.7. One side private information

On this section we simplify the model. Assume now that we are in a private information environment. This is, the signals are now the valuations. Also, assume signals are independent.

Proposition 5. Suppose there are two players, informed (1) and uninformed (2). This is, player 1 knows her type and the type of the other player, but player 2 just knows her type. Then, fixing player’s 2 type, suppose that player 1(informed) could pick a finite support and distribution on a set $X \subset \mathbb{R}^+\setminus\{0\}$ from where her type is draw. Then, she will pick a set with at most two elements; that will always contain the supremum of $X$, and sometimes, the infimum of $X$.

The idea is that against a given player, player 1 will maximize her utility by being really strong (having a high valuation), or by having some probability of being weak (of having a low valuation). It seems somehow obvious that the supremum of $X$ is going to be involved in her decision. This is associated with having a better position against player 2. However, it might be surprising to find that sometimes the lowest possible value is also involved. Despite the fact that this leaves a chance of ending up being really weak against player two, it also has strategic implications. By having a positive probability of being weak, player 2 accommodates in equilibrium, being less aggressive. By doing that, player 1 ends up with a big surplus in case she is strong, gaining informational rents.
How player 2 accommodates is going to depend on both the probabilities and the distance between the high and low type. A lower low type allows player 1 to have higher probability on the high type, maximizing the informational rents.

We will have a one or two element set depending on how the supremum of $X$ compares to player’s 2 type. It might be the case that it is better to have two types than just one even when the supremum is bigger. However, if the supremum increases, at some point it becomes optimal to have just one type.

It is surprising though that the optimal distribution involves weights in only two types, but apparently is the only thing needed to maximize the informational rents. Whenever the distribution has only one point is because the informational rents are not big enough to compensate having a higher type.

Finally, this private information on valuations can be changed to asymmetric information on cost, by defining the cost of effort as one over the valuation. We also show that everything holds for that case, even though under private information the two models are not necessarily equivalent.

Outline of the proof:

We analyze a problem where the support is fixed, and we want to find the probabilities that maximize the ex-ante expected payoff of player 1. To do that we consider three cases. First, when player’s 2 type is always greater than any element on the support of 1. Second, when it is always less than the lowest element 1’s support. And finally, when it is in between.

For the first case we show a way of simplifying the problem, then that the solution will have no more than 3 active types (actives in the sense that induce positive effort), and
finally that any support with 3 point will have a solution that has positive probability only on 2 points. After that, we solved the 2 points case and conclude for the $N$ case; in which mass is put only on the lowest and highest types of the support.

On the second case, we show that it is always optimal to have all the mass on the highest type on the support.

Finally, the third case is a little more complicated, it involves a mixture of the previous cases. We divided the set of possible probabilities in two. One subset was similar to the first case, and the other to the second. With some small modification of what was previously done, we manage to prove that the solution would not have more than 2 point with positive mass. After that, we solved the case for 2 points, which ended up being either everything on the highest, or some probability on the lowest and some on the highest in the same way of the first case. However for some combinations the solution for two was not viable. To conclude, we use the analysis of 2 points to solve the $N$ case.

All the solutions had something in common, they where increasing on the highest type, and weakly decreasing on the lowest one. Therefore this allows us to conclude the proposition.

\subsection*{1.8. Future Work}

Proposition 5.2 works for finite supports. Naturally, a question that arises is if the results can be extended to any distribution in a compact set. A way to do that is by showing that when a discrete distribution of types converges to a continuous distribution, the ex ante expected utility of participating in the game also converges. That should probably involve showing that the resulting equilibrium strategies are also converging,
and that in the case of one side incomplete information, with the informed player having a continuous distributions, there is a unique equilibrium. Most of these has been shown for symmetric players, in a private model. Amman and Leining (1996) show the uniqueness of equilibrium in the symmetric and continuous case, and Contat (2013) shows that in that context, if some discrete distribution are converging (uniformly) to the continuous distribution, then the equilibrium strategies converge in probability. It seems that the proofs used by the previous authors could be replicated for our particular case, and maybe extended for asymmetric situation. With that result we could extend our result to the general case.
CHAPTER 2

War of Attrition with Information Externalities
2.1. Introduction

This paper analyzes War of Attrition (WoA) games when players have asymmetric valuations for prizes and may suffer from allocation externalities.

The War of Attrition is a game in which players compete for a prize by staying in the game until some player quits. There is an instantaneous cost to stay, but it is only paid if the game hasn’t ended yet.

Many papers in both the theoretical and empirical literature use WoA but they usually make strong assumptions regarding the preferences of agents. Moreover, they usually focus on one particular equilibrium without characterizing the full set of equilibria. Usually, because the game studied involves rounds of elimination until only one player is left, there would be a $n$ player $n$ prize game, where at each stage a WoA is played and one player is eliminated.

Because we are interested in understanding all the properties of the game, we focus on only one stage of the previously described game, namely the game in which $n$ players play, but only one player is eliminated. To differentiate from the previous literature, we allow for allocation externalities, that is, the utility a player gets for staying in the game might depend on the identity of the player that is eliminated.

We focus on this game because in the general case in which all players are eliminated one by one, our setting would allow for a player to have preferences on the sequence of eliminated players.

In this first attempt to allow for allocation externalities, we study the simplest WoA possible. That is, without private information, and with deterministic valuations and marginal costs that are not time dependent.
By solving the game equilibria, we are able to show that in equilibrium we either see the game solved at zero, or at any time greater than zero with a positive probability. In the latter case, we show that all analyses are based on the hazard rate of the equilibrium strategies, that are directly determined by a system of linear equations and inequalities that take into account any small interval of size, \( dt \), at which players are active (with a probability of quitting in that time) and at which players are not.

From there we can build three types of equilibrium. First, one in which all players have a chance of quitting at any time. Second, an equilibrium in which a set of players stay forever and the rest play with a positive chance of quitting at any time. And third, an equilibrium in which the "active" players change over time.

We are able to pin down the exact conditions in which each type of equilibrium arises and show that in the classical WoA (no externalities) both the second and third type of equilibrium can be observed, something that previous literature has neglected.

For the general game, we show conditions on the primitives for the first equilibrium to exist. We show that there is a parallel between some equilibria of the second and third type, and for games with three players, we are able to reduce the third type of eq. to only strategies that are not active until some time \( t^* \), but behave like the first type afterwards.

Later, we discuss how to extend these results to both marginal costs and valuations that are time dependent and how to use the results in applied settings.

2.1.1. Applications of WoA games

The original game of WoA was proposed by Maynard Smith (1974). Bishop and Cannings (1976) showed that this game has a unique Evolutionary Stable Equilibrium (ESE). This
game consisted on two identical players fighting for a prize they value and both contestants will pay a cost equal to the length of the time taken to resolve the contest. The contest is resolved when one of the players quit. This is a model with symmetric preferences, continuous time and complete information. The first generalization of this simple model is found in Haigh and Cannings (1989).\textsuperscript{1} They generalize the model to allow for many players and many prizes, but with symmetric preferences. They provide a solution for the case where preferences are symmetric and prizes are not ordered in increasing value.

There is vast literature in WoA games with some variations from the main assumptions made in the original game. Kapur (1995) proposes a model of WoA in discrete time intervals. Vehn et al. (2018) and Bulow and Klemperer (1999) propose models of WoA with asymmetric information.\textsuperscript{2} Bulow and Klemperer (1999) generalized the Maynard Smith (1974) article by having players paying the waiting cost even after they have quit. They allowed for heterogeneous players preferences but they do not allow for heterogeneous prizes. In particular all players assign a value of zero for the first $K$ prizes and a positive value to the remaining prizes.

Park and Smith (2008) provide a general characterization of stopping games where they characterize wars of attrition as being games in which having more predecessors helps. In our setting we focus in a more particular game, but allow for the identity of the predecessor to matter.

\textsuperscript{1} Hendricks et al. (1988) generalized the original game by allowing for a more general payoff and cost functions that change over time.
\textsuperscript{2} Murto and Valimaki (2011) present a model of exit with learning. The model can be considered of one of WoA with private valuations that change over time. Hopenhayn and Squintani (2011) present a model of WoA with private valuation and changing values over time.
Strictly speaking a WoA is a type of all-pay auction. In the case with $N+1$ players and $N$ identical objects it is identical to a second-price all-pay auction, since the game ends after one player quits. In the case with $N+1$ players and $N$ different objects the game at each stage is also a second-price all-pay auction.

For the same reason it is related to the contest literature since a contest is a first-price all-pay auction (APA). Siegel (2009) and Siegel (2010) present a generalized model of asymmetric contests. A WoA is equivalent to a second price APA while a contest is equivalent to a first price all pay auction. While he characterizes the unique equilibrium in contests we characterize all possible equilibria in WoA using two simple properties. In particular we show how an equilibrium where all players play a pure strategy always exists and it is a particular case of a broader class of equilibria. Taylor (1995) propose an (APA) to explain research tournaments. In research, inputs as unobservables and outcomes unverifiable, thus companies can promise a monetary prize to the manufacturer that produces the best prototype. This form of competition among manufacturers resembles an APA. (See Fudenberg (1983) for a model of leapfrogging and competition in patent races.)

Our characterization of preferences resembles that of Jehiel et al. (1996). The preferences of our stage-game are equivalent to theirs. They propose an optimal auction that is efficient in the presence of externalities while we only study that situation in a War of Attrition.

WoA games have been used in many applications like worker strikes, new technology standards, lobbying, political campaigns, R&D, exit in oligopolies, Coase conjecture,
etc. Yet, as we show below there are other applications for the WoA once we allow for externalities (e.g. identity-dependent preferences).

Fudenberg and Tirole (1986) present an exit game with duopolistic competition where firms profitability is private information. The solution of the model is a pure strategy equilibrium in the time of exit. Takahashi (2015) generalizes Fudenberg and Tirole (1986) model to allow for oligopolistic competition and estimates an exit game. However, this model still retains the strong assumptions of symmetry (identical productivity across firms) and interdependence (the effect of exit in profits is identity-independent).

Alesina and Drazen (1991) propose a WoA model to explain the delays for macroeconomic stabilization observed in the data. Different socioeconomic groups try to wait out other sectors. The game ends when one sector gives in and accepts a disproportionate burden on the adjustment. Pithcford and Wright (2011) apply a WoA model to debt restructuring. Here the sovereign debtor plays a passive role while the creditors compete to each other in a WoA. Waiting to settle is costly, but those who settle when there are fewer creditors will get better conditions.

Porter (1995) and Hendricks and Porter (1996) estimate a WoA in the wildcat tracks for exploratory oil drilling. There are information externalities that generate a free-rider problem: drilling cost (millions of dollars) is incurred by the driller but information regarding the existence of oil in the area becomes common knowledge. The value of waiting is the value of an option in drilling: you drill if it there is enough oil. The value of drilling is lower than waiting, because you incur the cost even when there is no oil. Thus the WoA. Firms have 5 years to initiate exploration; otherwise the rights revert back to the government. It is a WoA with finite horizon and discrete time. (Hendricks et al.,
1988) solve the case of a WoA in continuous time and complete information with finite horizon. In this case the hazard rate is not constant, rather the probability that each firm drills in the last period is high. Empirically the hazard rate is U-shaped. In this case a WoA is equivalent to a two-armed bandit problem with N players.

Becker (1983) presents a model of lobbyists competing to get political influence with some government agent. Baye et al. (1993) propose an all-pay auction (APA) to explain the lobbying process. They show how a politician can choose the set of "finalist" to maximize rent extraction. Hirsh and Shotts (2015) present a model of contests applied to policy developers (or bureaucrats) competing to get their projects approved by a decision maker.

Cabral (2004) studies a model of delayed entry. When firms decide to enter a market they have to pay 1$ per period during $K$ periods after which entry will occur. The model resembles a WoA with finite horizon.

Montez (2013) proposes a model of a monopolist of durable good where buyers with high valuation engage in a WoA with each other. The seller will reduce the price of the objects only after one buyer has purchase the good, thus the incentive to wait. One could generalized his model to allow for buyers to differ no only in their valuations but also in their discount rate. In this case, and given a fixed sequence of prices, the continuation values sequence for each agent might be different and non-monotonic. Hence, one would need the solution provided here to solve for that game.
2.2. Setting

Definition 6. A Stage WoA Game is defined by $n$ players. Each player is characterized with $\Delta_i = \{\Delta_j^i\}_{j=1}^n$, where each player chooses a quitting time $t_i$ and her utility is:

$$U(t_i, t_{-i}) = \begin{cases} 
-t_i & \text{if } i \text{ is the first to quit} \\
\Delta_i^j - t_j & \text{if } j \text{ is the first to quit} \\
\sum_{j \in K} \Delta_i^j \frac{1}{|K|} - t_i & \text{if } i \text{ ties being the first to quit with the players in } K 
\end{cases}$$

The structure of $\Delta = [\Delta_i^j]_{j=1}^n$, might come from a continuation game. For example, if there are $n$ prizes and players are eliminated one by one, $\Delta$ can be calculated in equilibrium by doing backward induction, to find the SPE game of such game.

Definition 7. Properties of WoA stage-games:

- A stage game is independent if the difference in continuation value of each player is independent of the identity of the player who quits, i.e., $\Delta_i^j = \Delta_i^k$ for all $j,k \neq i$ and for all $i$.
- A stage game is symmetric if $\sum_{j=1,j\neq i}^n \Delta_i^j = \sum_{j=1,j\neq k}^n \Delta_i^j$ for all $i,k$.
- An stage game is ordered when the continuation value of waiting is greater than the continuation value of quitting, i.e., $\Delta_i^j > 0$ for all $i,j$.

The classical WoA would be one that has all three properties.
2.2.1. Examples of WoA

We now illustrate the main contributions of the paper with some simple examples. The examples are conceived as the simplest possible deviations from a symmetric game that can illustrate the main differences between a symmetric and an asymmetric game, and between as asymmetric independent game and a game with externalities. We analyze a stage of the game in which 3 players compete for 3 prizes.

**Example 1**: Classical game. We have a WoA with 3 players. Each player has a valuation of 1 if they exit last, a valuation of $\beta$, with $0 < \beta < 1$, if they exit second and a valuation of 0 if they exit first.

This is a symmetric game without externalities, like the classical games studied in the literature. In the last stage, $\Delta_{ij}^i = 1 - \beta, j \neq i$ for all players. Selecting an equilibrium where everyone mixes without atoms, the expected value of the last stage is $\beta$. This because as we will see later, in that type of equilibrium all surplus is expended.

Then, on the first stage, all players will have $\Delta_{ij} = \beta, j \neq i$

**Example 2**: Take the same game we had before, but now both player 1 and 3 are better when player 2 is the last to leave. That is, if player 1 or 3 is the runner up, but player 2 is the winner, the runner up gets $\beta + \varepsilon$. In any other situation, the runner up gets $\beta$, with $0 < \beta + \varepsilon < 1$. To translate into our notation, in the last stage, $\Delta_{ij}^i = 1 - \beta - \varepsilon$ when $i = 1, 3$ and $j = 2$, $\Delta_{ij}^i = 1 - \beta$ in the other cases, when $i \neq j$. Selecting an equilibrium where everyone mixes and there is no atom, on the first stage, $\Delta_{ii} = \beta + \varepsilon$ when $i \neq j, i = 1, 3; j = 1, 3, \Delta_{i}^i = 0$ and $\Delta_{ii} = \beta$ otherwise.

This example is depicted in 2.1, the matrix on the left shows payoffs for all players depending on the order they leave, and the matrix on the right is describing payoffs of the
Table 2.1. Example: Payoffs

<table>
<thead>
<tr>
<th>Winner</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-</td>
<td>(β + ε, 1, 0)</td>
<td>(β, 0, 1)</td>
</tr>
<tr>
<td>2</td>
<td>(1, β, 0)</td>
<td>-</td>
<td>(0, β, 1)</td>
</tr>
<tr>
<td>3</td>
<td>(1, 0, β)</td>
<td>(0, 1, β + ε)</td>
<td>-</td>
</tr>
</tbody>
</table>

Δ = \begin{bmatrix} 0 & β & β + ε \\ β & 0 & β \\ β + ε & β & 0 \end{bmatrix}

first stage game, assuming everyone mixes without atom in the last stage. Notice that for ε = 0 we are in the first example.

2.3. Equilibrium of the Stage Game

In this section we characterize all equilibria in the stage game\(^3\). For simplicity we restrict attention to ordered games. We relax this assumption in section ??.

In this game, a strategy for a player is a probability distribution over \(\mathbb{R}^+ \cup \{0, +\infty\}\) of the time at which the player exits the stage game.

Also, define \(T\) as the time at which the game ends. This is, given everyone’s strategies, \(T\) is the infimum \(s\) that make \(\Pr[\text{game finishes before } s]\) equal to one. In this game, \(T\) could take any value between 0 and \(+\infty\). Also, for ease of exposition, we define three different set of players:

- Atom (\(A\)): this are the players whose strategy has an atom in \([0, T)\) (a probability greater than zero of an specific exit time). Define \(A = |A|\)
- Mixing (\(M\)): this is the set of players players that are mixing without atoms in \([0, T)\). Define \(M = |M|\).
- Staying (\(S\)): players in this set have a probability greater than zero of not leaving before the game ends. Define \(|S| = S|\).

\(^3\)For ease of exposition we are not considering strategies that include gaps in the probability distribution of exit times. We discuss the existence of such strategies in the appendix.
2.3.1. Equilibrium properties

**Lemma 1.** PURE STRATEGIES. There are at least $n$ pure strategy equilibria, and in all of them, $T = 0$, with at most one player leaving at zero.

Lemma 1 above shows that there always exist one equilibrium in pure strategies. Moreover, there is one such equilibria for each player, so there are at least $n$ equilibria in pure strategies. The results are actually stronger than what lemma 1 suggests, because when one player exits with probability one, there are many strategies followed by the other players that would create an equilibrium. In particular, as long as the strategies of all the other players prescribe that they exit with zero probability at zero, then the player that exits with probability one, has no profitable deviation.

The proof of this Lemma is simple. Suppose there is an equilibrium where the first player to leave does it at some time $t > 0$. In that situation her utility is $-t$, and therefore, she has a profitable deviation by leaving at zero. Therefore, if there is a pure strategy eq., at least one player has to be leaving at zero.

If one player is leaving at zero, and the other players are leaving at some time greater than zero, then nobody has incentives to deviate. Therefore that situation is an equilibrium, and there is at least one for every player.

**Lemma 2.** ATOM. In any equilibria with $T > 0$, $A \leq 1$ and if there is a player in set $A$, it has an atom at zero.

Lemma 2 is an extension of Lemma 1. First, it shows that not only can there be only one player that plays a pure strategy of exiting immediately, but that there can only be one player that exits immediately with a mass of probability. Second, Lemma 2 also
shows that if that player does not exit, then she (and at least some of the other players as we show below) play a mixed strategy for $t > 0$. Third, it shows that such a strategy, exiting with a probability mass at a given point, can only happens at $t = 0$. Notice that in Lemma 1 there was no concern about the timing of the mass of exiting. Since the mass was equal to 1, the player just exits at $t = 0$. However, when the mass is lower than 1, it could be possible that a player play some strategy at $t = 0$, or maybe for some time interval, and then exits with a probability mass. Lemma 1 shows that this cannot be an equilibrium.

The proof is the following. Suppose we are in an equilibrium with $T > 0$, and two players have an atom at some $T > t \geq 0$. Suppose each one leaves at $t$ with probability $p_j$. For player $j$, leaving at $t + \varepsilon$ with $\varepsilon$ very small. Because we can find some interval close to $t$ without atoms, the probability of the game ending in $[t, t + \varepsilon]$ is as close as the probability of the game ending in $t$ as we want it to be. However, because if there is a tie only half of the prize is awarded, by leaving with probability $p_j$ at $t + \varepsilon$ she gets both halves in case player $i$ leaves at $t$, but her cost of doing so increases by as little as we want it to be. Therefore, in equilibrium, two players can not have an atom in the same place.

Assume that some player $j$ has an atom at some $0 < t < T$ of size $p$. Then, the other players are better off by leaving at $t + \varepsilon$ than at $t - \varepsilon$, for $\varepsilon$ small. This because by leaving slightly after $t$, they get $p \cdot V_i^j$ extra but the difference in cost is as low as we want it to be (picking $\varepsilon$ accordingly). Then, in equilibrium, nobody is leaving in $[t - 0.5\varepsilon, t]$. However, if that is the case, then player $j$ is better off by leaving at $t - 0.5\varepsilon$. This argument only works if $t > 0$.
Lemma 3. In any equilibria, in any sub-interval $SI$ of $[0,T)$, if a set of players $P \subset M$ is mixing with positive density in $SI$, and the rest is not mixing in $SI$, then the hazard rate of those players is constant in that sub-interval ($\lambda_i$) and satisfies:

$$\sum_{j \in P, j \neq i} \lambda_j \Delta^i_j = 1, \; \forall i \in P$$

In any equilibrium in mixing strategies, if players have a positive density in a small neighborhood, then they have to be indifferent between leaving at any point on that neighborhood. In particular, they have to be indifferent between leaving at $t$ and $t + dt$. If a player leaves $dt$ later, then she wins more often because the probability of someone leaving increases. The expected payoff of that is going to be the hazard rate at which player $j$ is leaving, times what she gets when that player leaves, times $dt$ plus a term that depends on $dt^2$; for each player that is active in that neighborhood. Also, waiting $dt$ more implies that her cost of staying increases in $dt$ times the probability that nobody leaves in that interval, which is of the order $1 - o(dt^2)$. Dividing by $dt$ and taking the limit of $dt$ to zero, we get exactly the expression of the lemma. We provide a more detailed proof on the appendix.

As you can see, the expression is actually a system of $|P|$ linear equations with $|P|$ unknowns, in which nothing depends on $t$. Therefore, the solution can not depend on $t$.

Corollary 8. In equilibrium, if none of the players of $M$ has a gap in her strategy, then $T = \infty$ and for those players their equilibrium strategy has a constant hazard rate.

\footnote{This because strategies are independent.}
This is a direct implication of Lemma 3. If the same players are always mixing, then the equation holds for all $t$. Therefore, the hazard rate is constant in $[0, T)$. In such case, all strategies have an exponential distribution, with at most one player having an atom at 0. Because the exponential distribution has an infinite support, the only $T$ that could make this happen is $T = \infty$.

**Corollary 9.** *In equilibrium, if $T > 0$, then $T = \infty$.***

As before, suppose $T$ is not $\infty$. Then there must be that for some $\varepsilon$, in $(T - \varepsilon, T)$ the same players are active. Otherwise, there is a contradiction with the definition of $T$. If that is the case, then for those players the hazard rate of their strategy is constant. Therefore, their cumulative distribution is of the form $1 - pe^{-\lambda t}$. In order for that to hold and $T$ satisfy its definition, it must be that all players mixing in $(T - \varepsilon, T)$ have an atom at $T$, which is a contradiction with Lemma 2.

**Lemma 4.** *In any sub-interval $SI$ of $[0, T)$, if a set of players $P$ is mixing with constant hazard rate in $SI$, and the rest is not mixing in $SI$, then for a player $i \notin P$ the value of leaving increases (decreases) in $SI$ if and only if

\[
\sum_{j \in P} \lambda_j \Delta^j_i < (>1), \quad \forall i \in P
\]

This lemma gives us an idea of which player have incentives to leave and which not. The detailed proof is in the appendix, but is very similar to the proof of Lemma 3. If the instant benefit of staying is smaller than the instant cost of staying, the incentives
to leave increase. And the opposite if the instant benefits of staying are greater than the instant costs.

**Lemma 5.** GAPS. In any equilibrium with $T > 0$, in any sub interval of $[0, T)$, there are at least two players mixing.

We would like to prove that in any equilibrium, all players in $\mathcal{M}$ have strategies without gaps. However, we are not able to prove that for the general case. What we can show is that in equilibria with $T > 0$, there is a chance of someone leaving at any time. Later, we discuss how to use this result in order to compute all equilibra.

Suppose there is a subinterval $[a, b] \, ^5$ of $[0, T)$ where nobody is mixing. Assume you can find someone mixing at $b + \varepsilon, \forall \varepsilon > 0$ small. Otherwise, change $b$. Then, a player mixing close to $b$ is better off by mixing in close to $a$, since her cost drops discretely but the probability of winning doesn’t change much.

Now assume only one player is mixing. If that was the case, then she is better of leaving at $0.5(a + b)$ than closer to $b$.

### 2.3.2. Equilibria types

With the previous lemmas, we can identify all equilibria of this game. Depending on the expected finish time, we can have only two type of equilibria.

First, there is a class of equilibria in which the game finishes at $t = 0$. This includes all pure strategy equilibria.

Then, using Lemma 3, we can argue that all other equilibria will have $T = +\infty \, ^6$.

---

5The proof also works for $(a, b)$

6The solution for the CDF will reach 1 only at $t = \infty$
When $T = +\infty$, there are many different equilibria depending on which players are in $\mathcal{A}$, $\mathcal{M}$ and $\mathcal{S}$.

When there are gaps on some players’ strategies, we can use Lemma 3 and 4 to check if equilibrium conditions are satisfied. To do that we have to check on each interval were a player in $\mathcal{M}$ is not mixing the value of leaving for that player is lower or equal than the value of leaving on the interval in which she is leaving. This implies that at the end of a non mixing interval, the value of leaving can not be decreasing.

We discuss later how to find this type of equilibrium.

For the case in which none players have gaps on their strategies, with Lemma 3 and 4, we can get the hazard rates of the players in $\mathcal{M}$ and check the conditions that have to be satisfied for players in $\mathcal{S}$.

This generates two types of equilibria. Depending on $\mathcal{S}$ being empty or not.

Finally, each of these equilibria could eventually have some player with an atom in her strategy. However, from 3, even if a player has an atom at zero, there is an equilibrium in which nobody has an atom, but everyone in $\mathcal{M}$ is leaving with the same hazard rate. Also, conditional on the exit time being greater than zero, the two equilibria mentioned before are exactly the same.

In this section, we give definition of the equilibria mentioned before, discuss its existence and explore some of their properties.

**Definition 10.** A NWOA equilibrium is an equilibrium in which $T = 0$. 
This equilibrium always exist, and consist of one player leaving at zero, and all the other players playing any strategy that doesn’t have any positive probability of leaving at zero.

In this equilibrium, if player $j$ is leaving at zero, her expected utility is zero, and everyone else is $V_{ij}$

We call it NWOA because there is no observed war of attrition in equilibrium.

\textbf{Definition 11.} A SESE equilibrium is one in which $\mathcal{A} = \emptyset$, $\mathcal{M} = I$ and $\mathcal{S} = \emptyset$, and all players have a strategy with constant hazard rate that satisfies:

$$\sum_{j \in I, j \neq i} \lambda_j \Delta^j_i = 1, \forall i \in I$$

Defining $\Delta_I = (\Delta^j_i, i, j \in I)$ with $\Delta^i_i = 0$, the above conditions is equivalent to have a positive solution for $\lambda$ in

$$\Delta_I \lambda = 1$$

with $1$ a vector of ones.

The solution will exist depending on the shape of $\Delta_I$. Because we haven’t restricted $\Delta_I$, we can not say much about it.

However, if an equilibrium like this exist, then every players’ expected utility is zero.

We call this equilibrium SESE because when all players are symmetric, it is the only evolutionary stable equilibrium (Haigh and Cannings 1974).
If all conditions are satisfied, except \( A = \emptyset \), then we say the eq. is a SESE with an atom.

We showed that there can not be more than one player with an atom at zero. Call that player \( \tilde{i} \) and \( p \) her probability of leaving at zero. Then the expected ex-ante utility of that player in such equilibrium is going to be zero, but for the other players is going to be \( V_{\tilde{i}} \cdot p \), the utility they get if player \( \tilde{i} \) leaves, times the probability she does it.

**Definition 12.** A **STAYING** equilibrium is one in which \( A = \emptyset \), \(|M| > 1\) and \( S \neq \emptyset \), all players in \( M \) have a strategy with constant hazard rate that satisfies:

\[
\sum_{j \in M, j \neq i} \lambda_j \Delta^i_j = 1, \ \forall i \in M
\]

And players in \( S \) satisfy

\[
\sum_{j \in M} \lambda_j \Delta^i_j \geq 1, \ \forall s \in S
\]

The first conditions states that players that are mixing should be indifferent between leaving and staying. The second conditions is a consequence of mixing players having constant hazard rates. In order for a player to have incentives of staying, the instant costs of staying has to be greater than the cost.

As in the previous definition, the indifference condition will depend on a system of linear equations to have a positive solution, but that has to satisfy an inequality.
At first, it might look as harder conditions to satisfy than in a SESE. However, since there are much more combinations of players in $\mathcal{M}$ and $\mathcal{A}$ that could be checked, it is not clear that it is easier to find a STAYING equilibrium than a SESE one.\footnote{These combinations are all that have at least 2 players in $\mathcal{M}$, which is going to be $2^N - 1 - 1 - N$, the number of all non empty subsets of players, that does not include all players, minus the subsets with only one player.}

All players mixing will have ex-ante utility of zero, all player mixing will have positive ex-ante expected utility.

If all conditions are satisfied, except $\mathcal{A} = \emptyset$, then we say the eq. is a STAYING equilibrium with an atom. Only a player that is in $\mathcal{M}$ might have an atom. In such case, that player will have an ex ante expected utility of zero, and all other players will have a greater ex ante expected utility because of the probability of that player leaving at zero.

**Definition 13.** We say an equilibrium has GAPS if $T \neq 0$, and there is at least one player in $\mathcal{M}$ with a support that is not an interval.

For two players, we know that such equilibrium does not exist. However, for more players, it might exists. What we know from the previous lemmas allows us to see for each particular game if an equilibrium could exist. The important thing to check the existence is to remember Lemma 3 that on each interval in which the same player are mixing, the hazard rate must be constant. Therefore, in this eq. we will have constant hazard rates per segments. Also, using Lemma 4, if a player is not mixing on an interval greater than zero, then it must be that at first her expected value of leaving decreases, and at the end of the interval increases, because only in that situation she will be indifferent between leaving at the beginning and end of the interval.
Also, for every staying equilibrium, if the SESE equilibrium exists, then an equilibrium in which players in $S$ start mixing at any time $t^* > 0$ with the hazard rate of the SES equilibrium, and players from $M$ mix with the hazard rate of the staying equilibrium until $t^*$ and with the SESE eq. afterwards, will still be an equilibrium. (With gaps).

Moreover, using the previous paragraph and the previous insight, we know that for three players that is the only type of equilibrium with gaps that can arise.

2.3.3. Example

A modified version of Example 2 from section 2.2, defined by:

$$\Delta = \begin{bmatrix} 0 & \beta & \beta + \epsilon \\ \beta & 0 & \beta \\ \beta & \beta & 0 \end{bmatrix}$$

As we said before, an NWOA equilibrium always exist. In order to find a SESE equilibrium, we would have to invert the matrix.

The inverse is

$$\Delta^{-1} = \frac{1}{2\beta + \epsilon} \begin{bmatrix} -1 & \frac{\epsilon + \beta}{\beta} & 1 \\ 1 & -\frac{\epsilon + \beta}{\beta} & \frac{\epsilon + \beta}{\beta} \\ 1 & 1 & -1 \end{bmatrix}$$

And the resulting solution for $\lambda$ in a SESE equilibrium is:

$$\lambda = \begin{bmatrix} \frac{1}{2\beta + \epsilon} \\ \frac{1}{2\beta + \epsilon} \\ \frac{1}{2\beta + \epsilon} \end{bmatrix}$$

Then, for positive $\epsilon$ and $\beta$, a SESE exist.

Let’s find a WESE in which player 2 and 3 are mixing and player 1 stays forever. In that case:
\[ \Delta_I = \begin{bmatrix} 0 & \beta \\ \beta & 0 \end{bmatrix}, \quad \Delta_I^{-1} = \begin{bmatrix} 0 & \frac{1}{\beta} \\ \frac{1}{\beta} & 0 \end{bmatrix} \text{ and } \lambda = \begin{bmatrix} \frac{1}{\beta} \\ \frac{1}{\beta} \end{bmatrix} \]

It will be an equilibrium as long as:

\[ V_{12}\lambda_2 + V_{13}\lambda_3 > 1 \]

Or

\[ \frac{2\beta + \varepsilon}{\beta} > 1 \]

Then an equilibrium with player 1 staying forever, and both player 2 and 3 mixing exists.

Finally, in order to have an equilibrium with either player 2 or 3 staying forever, and the rest stayin, we would have to solve for

\[ \Delta_I = \begin{bmatrix} 0 & \beta + \varepsilon \\ \beta & 0 \end{bmatrix}, \quad \Delta_I^{-1} = \begin{bmatrix} 0 & \frac{1}{\beta+\varepsilon} \\ \frac{1}{\beta+\varepsilon} & 0 \end{bmatrix} \] and

\[ \lambda_I = \begin{bmatrix} \frac{1}{\beta} \\ \frac{1}{\beta+\varepsilon} \end{bmatrix} \]

It will be an equilibrium as long as:

\[ V_{31}\lambda_1 + V_{32}\lambda_3 > 1 \]

which is true since:
\[ 1 + \frac{\beta}{\beta + \varepsilon} > 1 \]

Since all possible staying and SESE equilibrium exists, the multiplicity of equilibrium is even bigger if we take into account equilibria in which until some time \( t^* \) two players mix, and afterwards all players mix.

### 2.4. Symmetric Valuations

Recall that a game is symmetric if
\[
\sum_{j \in I \setminus \{i\}} \Delta^j_i = C, \forall i \in I, \text{ for some } C > 0.
\]

Because in that is still a general setting, there is not much what we can say. However, since our results depend mainly on the solutions of a linear system, we can use results from linear programming to get some properties of these games.

**Lemma 14.** If a game is symmetric, then a SESE exists. If \( \Delta \) is invertible, it is unique.

**Proof:**

We are going to use a Lemma from linear programming:

**Lemma 15** (Farkas). Let \( A \in \mathbb{R}^{m \times n} \) and \( b \in \mathbb{R}^m \). Then, exactly one of the following statements is true:

1. There exists an \( x \in \mathbb{R}^n \) such that \( Ax = b \) and \( x \geq 0 \).
2. There exists an \( y \in \mathbb{R}^m \) such that \( A'y \geq 0 \) and \( b'y < 0 \)

In our case, \( b \) is the vector of ones, \( A \) is \( \Delta \), and \( n = m = |I| \).
Suppose we are in case 2 of Farkas’s lemma. That would mean that there is a \( y \) such that \( \Delta' y \geq 0 \) and \( \sum y_i < 0 \).

But \( (\Delta' y)_j = \sum_{i \in I} \Delta^i_j y_i \geq 0 \).

Adding those terms we get:

\[
\sum_{j \in I} (\Delta' y)_j = \sum_{j \in I} \sum_{i \in I} \Delta^i_j y_i
= \sum_{i \in I} y_i \sum_{j \in I} \Delta^i_j
= C' \sum_{i \in I} y_i
< 0
\]

Which is a contradiction. Then we are in the first case of Farkas lemma, which implies that there exists some \( \lambda \geq 0 \) such that \( \Delta \lambda = 1 \).

2.5. Independent Valuations

Even though the case with no externalities has already been studied, the previous results can be strengthened and give us properties of the equilibria that haven’t been discussed previously. Recall that a stage game with no externalities satisfies \( \Delta^i_j = \Delta^i_k = \Delta_i \). The difference with the general case is that the inverse of \( \Delta_I \) can be computed.

2.5.1. Asymmetric Valuations

We first restrict attention to cases where the players valuations are independent and fixed, but could differ across players. We present stronger results that do not apply in the general case.
**Property 16.** Consider a stage WoA with \( n \) players. Then for any mixed strategies equilibrium, on any interval reached with positive probability where the set of players mixing does not change and is equal to \( P \), the hazard rate of all players in \( P \) is constant and satisfies:

\[
\lambda_i = -\frac{1}{\Delta_i} + \frac{1}{|P| - 1} \sum_{j \in P} \frac{1}{\Delta_j}
\]

This property comes using Lemma 3, and inverting \( \Delta \) to solve the linear system.

**Property 17.** A SESE equilibrium will exist as long as:

\[
\lambda_i = -\frac{1}{\Delta_i} + \frac{1}{|I| - 1} \sum_{j \in I} \frac{1}{\Delta_j} > 0, \quad \forall i \in I
\]

The lowest possible \( \lambda_i \) comes from the player with the lowest \( \Delta_i \), since \( \lambda_i - \lambda_j = \frac{1}{\Delta_j} - \frac{1}{\Delta_i} \).

Then, if \( \lambda_i > 0 \) for that player, a SESE equilibrium exist. However, playing with the expression, that is going to happen as long as

\[
\sum_{j \in I, j \neq i} \frac{1}{\Delta_j} > (|I| - 2) \frac{1}{\Delta_i}
\]

Then, if the lowest valuation is very low compared to the other players valuation, a SESE equilibrium might not exist.

Also, if the number of players increases, and the valuations are bounded, then it becomes harder for SESE to exist.

Define,

\[
\lambda^0 = \min_i \frac{1}{|I| - 1} \sum_{j \neq i} \frac{1}{V_j} - \frac{|I| - 2}{|I| - 1} \frac{1}{V_i}
\]
If $\lambda^0 < 0$, a SESE equilibrium does not exist. But for $|I|$ big enough, that is what happens since $\frac{1}{|I|-1} \sum_{j \neq i} \frac{1}{V_j} - \frac{1}{|I|-2} \frac{1}{V_i} = \frac{1}{|I|-1} \sum_{j \neq i} \frac{1}{V_i} - \frac{1}{V_i} + \frac{1}{|I|-1} \frac{1}{V_i}$, as long as the average of the inverse of the other valuations does not converge to the inverse of the lowest valuation, $\lambda_0$ will be eventually less than zero. If that is the case, a SESE equilibrium will not exist.

**Property 18.** A WESE equilibrium will exist as long as:

1. All players $i$ in $M$ are mixing with positive constant hazard rate equal to:

   $$\lambda_i = -\frac{1}{\Delta_i} + \frac{1}{|M|-1} \sum_{j \in M} \frac{1}{\Delta_j}$$

2. For all players $j$ in $S$ the following inequality is satisfied:

   $$\frac{\Delta_i}{|M| - 1} \sum_{j \in M} \frac{1}{\Delta_j} \geq 1$$

There is always going to be an equilibrium like this. Take the two players with lowest valuation, wlog, players 1 and 2. Then, the hazard rates for those players is going to be $\frac{1}{\Delta_2}$ and $\frac{1}{\Delta_1}$. Then, for the rest of the players, we have to check that

$$\Delta_i \left( \frac{1}{\Delta_1} + \frac{1}{\Delta_2} \right) > 1$$

Which is true by construction.

### 2.5.2. Symmetric Valuations

Recall that a game is symmetric and independant if $\Delta_{ij} = V$, $\forall i, j \in I$. In that case, we can go further on the simplifications.
Property 19. In a symmetric and independent war of attrition:

(1) A SESE equilibrium always exists.

(2) All possible WESE equilibrium exist.

Replacing on the previous properties, a SESE equilibrium will exist as long as

A SESE equilibrium will exist as long as:

\[ \lambda = \frac{|I|}{|I| - 1} \frac{1}{V} > 0, \forall i \in I \]

Which is always true. For a given \( M \), a WESE with only those players mixing and the rest staying forever will exist if:

\[ \lambda = \frac{|I|}{|M| - 1} \frac{1}{V} > 0, \forall i \in M \]

and

\[ \frac{V}{|M| - 1} \sum_{j \in M} \frac{1}{V} = \frac{|M|}{|M| - 1} > 1 \]

Which is always true.

2.6. Time-varying valuations

In this section, we relax those assumptions by allowing the value that each player assigns to each object to vary over time, i.e., \( V_{ij}(t) \) and \( c(t) \).

The strongest assumption of the classical WoA game is that the value that each player assigns to each object is independent of time. Implicitly there is an assumption that there is no discounting, and that the waiting costs are constant over time.
Also, the most important property of the war of attrition studied before, is that whenever there is an equilibrium with some players mixing, their hazard rate must be constant, at least by intervals.

This result comes from comparing the benefit of staying in the game for a little longer with the cost of doing so. Because in the classical game none of those quantities depends on time, the hazard rate of the other players must be constant in order for a player to mix.

However, even though when both the cost and the benefits change over time the previous analyses is no longer true, the intuition still holds. For a player to mix in an interval, she has to be indifferent between leaving at any point in that interval. For that to happen, in particular, she has to be indifferent between leaving at $t$ and $t + dt$. If she waits until $t + dt$, her benefit increases depending on the expected gains. That gains can be approximated as

$$
\sum_{j \text{ mixing in } t} V_{ij}(t)\lambda_j(t)dt
$$

And the increase in the cost of waiting a little longer to leave, can be approximated by $c'(t)dt$.

Then, we end up with a harder to solve system of equation at $t$, but one that has the same flavor as before.

In the classical WoA the assumption is that and $c(t) = t$. This assumption implies that the ratio of the value of staying in the game divided by the marginal cost of staying in the game is constant, i.e., $\eta(t) \equiv V_{ij}(t)/c'(t) = V_{ij} = \eta$. 
Intuitively, this means that the trade-off between staying in the game or exiting is constant over time, thus the hazard function of exit is constant over time.

If $\eta(t)$ is not constant over time, the hazard function would not be constant, and the distribution of exit times would not be exponential. However, if $\eta(t)$ is 'well-behaved' all we need to do is to find a distribution whose hazard function is equal to $\eta(t)$, i.e., $h(t) = \eta(t)$.

2.7. Discussion

It is natural to think on allocation externalities in this type of games. Complexity of economic agents make it necessary to take this into account for elimination games.

The current models of the War of Attrition don’t allow for that, and in this paper we explore how would things change when we consider these externalities in the simplest possible way. Our results gives us more insight about the game, and enrich the possibilities for estimations.

However, there are many pros and cons of using this model in an applied setting. Among the good things is that it is easy to estimate, since there is a direct relationship between the hazard rates and the primitives of the model. Then, if the valuations are modeled in a simple way, and the exit times are observable, the model can be calibrated to fit the data. It is important to notice that in general, we will not observe all exit times, but only the first exit time; and that in the general case, even if we could estimate hazard rates, there are $n(n - 1)$ parameter to be obtained. However, this are problems that a smart applied economist could work around.
On the cons of this model for applied settings is the fact that there is a huge multiplicity of equilibrium. This hasn’t stopped researcher before, but in this setting we should be careful to test how sensible results are to the type of equilibrium selection, and justify its selection.

There are equilibria that can be easily dismissed under weak assumptions for this purpose. Like all that end at zero (NWOA), and the ones with atoms at zero (conditional on the game not finishing at zero, these are either SESE or Staying).
CHAPTER 3

Dynamic Tullock Contest
3.1. Introduction

Contests are predominantly used to model environments in which agents exert costly effort to increase their chances of winning a prize. Following Tullock (1975), a variety of papers have applied the “Tullock-contest” framework to the analysis of political competition, lobbying, sport tournaments, patent races, litigation, advertising, among other applications.¹

A large segment of the contest literature has focused on static environments. Despite the insights drawn from one-stage contests, many situations are dynamic and the fact that agents face each other on multiple occasions matters. Often, agents can utilize their effort from one contest in future contests. In advertising, for example, the resources spent on one campaign affect the starting point of the next one. In R&D races, the knowledge acquired during the development of one product is used in the development of future ones.

In this paper, we extend the static contest model by studying sequential contests with effort accumulation. In contrast to many papers on sequential contests, we allow for contest-specific prizes and cost of effort, as well as a contest success function, the function that maps effort to the probability of winning, not necessarily homogeneous. The homogeneity assumption, although widely used, could be violated in environments where firms compete using the outcome of a production process with multiple inputs (Rai and Sarin (2009)). In our analysis of equilibrium, we consider the whole set of possible deviations, and not only interior local deviations as in Schmitt et al. (2004) or Baik and Lee (2000).² When the accumulated effort cannot be disposed, global deviations

¹ See also Buchanan et al. (1980), Tullock (2001), and Jia et al. (2013).
²Baik and Lee (2000) only consider interior equilibrium, although their setting is slightly different to ours, as we do not study contests for participation. Schmitt et al. (2004) is closer to our model. In their analysis, they restrict to cases such that the equilibrium is always interior. We show that focusing on
are important because agents cannot always reach the interior equilibrium by exerting positive effort.

We provide novel predictions arising from the trade-off between dynamic competition and cost savings. In our model, agents take into consideration that exerting effort in the first contest they are choosing their head start for the second contest. The trade-off faced by the agents is similar to the one explored by Yıldırım (2005), where agents can revise their previous effort choices. The intuition behind the equilibrium outcome is simple: if exerting effort in the first contest is relatively inexpensive, compared to the cost of exerting effort in the second contest, agents have strong incentives to exert effort in the first contest, as their effort will be accumulated for future use. In fact, agents exert the same equilibrium level of effort that they would in a static contest where the prize equals the sum of the prizes of the two contests and the cost of effort is the cost of the first period contest. On the other hand, when agents anticipate lower cost of effort in future contests, they have fewer incentives to exert effort in the first contest. In this case, however, they exert more effort in the first contest than they would in a static contest with the same prize and cost of effort. Pal (1991) shows that the same kind of results arise if firms compete in a Cournot duopoly rather than in a contest.

We apply our equilibrium results to study two issues that have typically been examined in the contest literature: amount of rents dissipated and the aggregate effort in equilibrium. Nitzan (1994) summarizes the impact of different modeling assumptions (number

local interior deviations is restrictive, because in some cases the continuation strategy is a corner solution. Even more, this restriction hides interesting equilibrium dynamics.
of agents, risk attitude, asymmetries) on the amount of rents dissipated. Our model generalizes some of these results and offer novel predictions on equilibrium behavior, rent dissipation, and aggregate effort.

We derive conditions over the primitives of the model to characterize the amount of rent dissipation and aggregate effort in sequential contests with and without effort accumulation. We find that restricting the contest success function to the widely used “power” form (or the “logit” form) has significant qualitative implications. In fact, under the homogeneity assumption, the equilibrium level of effort is linear with respect to the parameters of the model. In this case, both aggregate effort and rent dissipation are unaffected by the possibility of effort accumulation. Once we drop the homogeneity assumption, and depending on the convexity of the equilibrium level of effort with respect to the parameters of the contest, we find that rent dissipation with effort accumulation can be larger, smaller, or the same as in sequential contests without effort accumulation. We also study to what extent effort accumulation changes the aggregate effort across contests. For instance, when the cost of effort increases over time, the aggregate effort across contests is lower when effort accumulation is allowed. For a different set of primitives, we find that allowing for effort accumulation increases the aggregate effort.

Finally, an important part of the contest literature is devoted to contest design. Our results show how to design a contests where effort accumulates if the final goal is to maximize the aggregate effort or to minimize rent dissipation. We also provide some comparative statics to understand situations where the policy tool is whether to allow for effort to accumulate among contests. The growing experimental literature on contests could
benefit from the new insights provided by our results, as it would allow experimenters to better explain their findings (Schmitt et al. (2004)).

The chapter is structured as follows: In the next section we briefly review the related literature. In Section 3.3 we present the model, assumptions, and the equilibrium characterization. In Section 3.4 we show novel results on rent dissipation and aggregate effort. Section 3.5 presents an extension of the sequential contests with variable carryover rate. Section 3.6 discusses our findings and conclusions.

3.2. Literature review

Our paper is related to the extensive literature of dynamic contests, where current effort affects the outcome of future contests through some channel. One of these channels is by obtaining an advantage in future contests. The advantage could come directly from the level of effort exerted in the current contest, as in Clark and Nilssen (2013), or implicitly by winning the contest, as in Möller (2012), or by generating asymmetries as in Clark et al. (2012). Another channel, is by allowing agents to store part of their current effort to use it in future contests. The closest paper to this setting is Schmitt et al. (2004), which presents a model and experimentally tests it. The main drawback of their model is that prizes and the cost of effort is the same for all the contests, the restriction to the linear Tullock-contests, and the requirement of sufficiently small carryover rates. Our model explore some interesting dynamics that are missing under their assumptions. Baik and Lee (2000) studies the effect of carryovers on elimination contests. In this setting, only the winners of a prior selection contests are able to compete for the prize. They find that rent dissipation increases with the carryover rate. In our setting there is no elimination,
all the agents are active at every contest, and our assumptions are less restrictive. We also study deviations that are not considered in Baik and Lee (2000), since they assume only interior local deviations. In a different setting but somehow related, Grossmann et al. (2011) analyzes an infinitely repeated Tullock-contest with effort accumulation, exploring the transitional dynamics in a log-linearization around the steady state.

Yildirim (2005) presents similar ideas, although in a different setting. In his model, agents can revise their effort before the end of the contest and there is only one prize. However, the ability of agents to revise their effort is equivalent to starting a contest with some effort or head start. Apart from having multiple contests with different prizes and cost of effort, we also provide results on rent dissipation and aggregated effort and we weaken the assumption of homogeneity in the contest success function. Regarding rent dissipation, Nitzan (1994) provides a summary of the literature.

An important feature of a contest is the mapping between effort and scores. The literature has made different assumptions regarding this mapping. Jia et al. (2013) discusses the different classes of contest success functions used in the literature. Skaperdas (1996) presents axioms to justify different contest success functions. In particular, the axiom of homogeneity is what delivers the “power” form. Clark and Riis (1998) extends these axioms to allow for heterogeneous abilities and Rai and Sarin (2009) allows for multiple investments.

The literature on contest design has studied different aspects of the design. In Moldovanu and Sela (2001), the contest designer chooses the size of the prizes, while in Moldovanu and Sela (2006a) the designer can choose the type of contest (elimination, simultaneous,
etc). Gershkov and Perry (2009) studies the optimal number of reviews before, for instance, a promotion. Although our focus is not design, we explore applications in which the carryover rate is the design tool.

More broadly, our paper contributes to the literature of endogenous head starts in contests. Kirkegaard (2012) studies the effect of head starts and handicaps on total effort. Siegel (2014b) finds the equilibrium for all-pay contests with head starts.

Finally, there is a growing literature on experimental contest. Parco et al. (2005) empirically tests an elimination contest with budget constraints, finding more rent dissipation than the theory predicted. Sheremeta (2010) compares the performance of one-stage versus two-stages elimination contests with carryovers. They find that two-stage contests generate higher revenue and higher dissipation rates than a one-stage contest. Schmitt et al. (2004) test a closely related setting, finding that their model does not fully explains their empirical findings.

3.3. Model

There are two agents competing sequentially in two contests. In each contest, agents simultaneously choose how much effort they want to exert. Effort is converted into scores by a strictly increasing and (weakly) concave function $g(\cdot)$, with $g(0) = 0$, and $g'(0) > 0$.\(^3\) We do not necessarily restrict $g(\cdot)$ to be homogeneous, which is a typical assumption in the contest literature.\(^4\) As we will show in section 3.4, rent dissipation and aggregate effort crucially depend on the shape of the score function.

\(^3\)The assumption $g(0) = 0$ can be weakened in some cases. For examples, our results are still valid for the “logit” form, although in that case $g(0) > 0$.

\(^4\)This assumption leads to a contest success function that is homogeneous (of degree zero). For example, $g(x) = x + \sqrt{x}$ is outside the standard framework and we will show the qualitative implications of this assumption.
In the first contest, agents are symmetric in their abilities and in their cost of exerting effort. There is a period-specific cost of effort $c_t$. If agent $i$ exerts $W_i$ units of effort in period 1, her score is $g_i = g(W_i)$ and the probability of winning the period-specific prize $V_i$ is given by

$$p(W_i, W_j) = \begin{cases} 
\frac{1}{2} & g(W_i) + g(W_j) = 0 \\
g(W_i) & g(W_i) + g(W_j) = 0 \\
g(W_i) + g(W_j) & \text{otherwise}
\end{cases}.$$  

Only the winner receives the prize, while the loser gets zero. The main feature of the baseline model is that the effort exerted by agents in the first contest is freely, and fully, carried over to the second contest.\(^5\) In other words, an agent that exerts $W$ units of effort in the first contest and $x$ units in the second one, competes with scores $g(W)$ in the first contest and with score $g(W + x)$ in the second one. In the first period, the agent pays a cost of $c_1 \cdot W$ to exert effort $W$, while in the second period the agent only pays $c_2 \cdot x$, the cost of the additional effort.

Given a history of effort in the first contest, $(W_i)_{i=1,2}$, and the additional effort exert in the second period, $(x_i)_i$, agent $i$ wins the second contest with probability $p(W_i + x_i, W_j + x_j)$. Thus, the effort chosen in the first contest acts as a head start in the second period. Competition in the first contest will endogenously determine the head start in the second contest.

Apart from effort accumulation and generally non-homogeneous scoring function $g(\cdot)$, our model features period-specific prizes and cost of effort. The trade-off introduced by

\(^5\)In section 3.5, we discuss the case of variable carryover rate.
period-specific parameters comes from the agents wanting to exert effort when is relatively inexpensive to do so, while the competition effect and the accumulation of effort pushes agents to exert effort in earlier periods.

To analyze the model, we will rely on known results for a single contest which we will embed in the dynamic contest with effort accumulation. In the next section, we summarize general properties of a standard simultaneous Tullock Contest. \(^6\)

### 3.3.1. Static Tullock Contest

We will refer as the static “Tullock” contest when the contest success function is given by (3.1).\(^7\) Conditional on \(s\), the effort exerted by the rival agent, the payoff for an agent that exerts effort \(x\) in the standard static Tullock contest is given by

\[
\pi(x, s|\theta) = \theta p(x, s) - x,
\]

where \(\theta = \frac{V}{c}\) is the value of the prize relative to the cost of effort. It is easy to verify that for fixed parameters \(s\) and \(\theta\) the function \(\pi(x|s, \theta)\) is strictly concave, which implies that the best response function

\[
R(s|\theta) = \arg \max_{x \geq 0} \pi(x|s, \theta)
\]

is well defined.

**Lemma 20.** The best response function satisfy the following properties:

(1) \(R(s|\theta) \leq \theta\) and \(R(s|\theta)\) is differentiable.

---

\(^6\)Most of the proofs are omitted in the main text and left to the Appendix.

\(^7\)We are somehow abusing terminology since Tullock contests are typically characterized the homogeneous score function \(g(x) = x^\alpha\).
(2) \( R(s|\theta) \) is strictly increasing in \( \theta \) for a fixed \( s \).

(3) \( R(s|\theta) \) is strictly increasing in \( s \) if and only if \( R(s|\theta) > s \).

(4) Exists \( \delta > 0 \) such that \( R(s|\theta) > s \) for all \( s \in (0, \delta) \).

(5) Exists a unique \( x^*(\theta) \) such that \( R(x^*(\theta)|\theta) = x^*(\theta) \).

The properties in Lemma 20 characterize the shape of the best response function \( R(s|\theta) \), as shown in Figure 3.1 below.

![Figure 3.1. Best response function, for a static Tullock contest with parameter \( \theta \).](image)

\( R(s|\theta) \) reaches its maximum at \( x^*(\theta) \), increases at \( s \), for \( s < x^*(\theta) \), and decreases at \( s \), for \( s > x^*(\theta) \). The unique symmetric equilibrium in pure strategies in the standard static Tullock contest with two agents is \( x^*(\theta) \), for \( \theta > 0 \), and \( x^*(0) = 0 \). This solution is valid as long as \( F(x) = \frac{g(x)}{g'(x)} > \frac{x}{2} \), for all \( x > 0 \).\(^8\) A useful property for our analysis is the following:

**Lemma 21.** \( x^*(\theta) \) is increasing in \( \theta \).

**Proof.** Consider \( \theta' > \theta \). By property 2 in Lemma 20, \( x^*(\theta) = R(x^*(\theta)|\theta) < R(x^*(\theta)|\theta') \).

Since \( x^*(\theta) < R(x^*(\theta)|\theta') \), by property 3 in Lemma 20 \( R(\cdot|\theta') \) is increasing at \( x^*(\theta) \). This implies that \( R(x^*(\theta)|\theta') < R(x^*(\theta')|\theta') = x^*(\theta') \). Thus, \( x^*(\theta) < x^*(\theta') \). \(\square\)

\(^8\)This condition is analogous to the standard condition \( \alpha < 2 \) for the case \( g(x) = x^\alpha \).
3.3.2. Static Tullock Contest with Head Starts

We start by finding the equilibrium in the second stage of the game. Let \( W_i \) and \( W_j \) be the first stage’s effort levels. In the second stage, agent \( i \) chooses her effort \( x_i \) to solve

\[
\max_{x_i \geq 0} V_2 \cdot p(x_i + W_i, x_j + W_j) - c_2 \cdot x_i
\]

Notice that we are assuming \( x_i \geq 0 \), so agents cannot dispose effort that comes from past periods.\(^9\) The second period’s effort choice can be written as the following constrained maximization problem

\[
\max_{u_i \geq W_i} \pi(u_i, x_j + W_j|\theta_2),
\]

where \( \theta_2 = \frac{V_2}{c_2} \). By Lemma 20, it is easy to verify that the best response function will be:

\[
x_i(x_j|W_i, W_j) = \max\{0, x^*(\theta_2) - W_i\}.
\]

Replacing the best response in the corresponding regions we have,

**Proposition 22.** In the second period, with head starts \((W_i, W_j)\) the equilibrium efforts are given by

\[
(x_i^*, x_j^*) = \begin{cases} 
(x^*(\theta_2) - W_i, x^*(\theta_2) - W_i) & \text{if } \max\{W_i, W_j\} \leq x^*(\theta_2) \\
(0, R(W_i|\theta_2) - W_j) & \text{if } W_i > x^*(\theta_2), \ W_j \leq R(W_i|\theta_2) \\
(R(W_j|\theta_2) - W_i, 0) & \text{if } W_i \leq R(W_j|\theta_2), \ W_j > x^*(\theta_2) \\
(0, 0) & \text{otherwise}
\end{cases}
\]

\(^9\)Allowing agents to dispose past effort give agents more incentives to exert effort early on relative to our results.
If both agents have large head starts, they will exert no effort in the second stage. Intuitively, the marginal benefit is always less than $c_2$, since the large head starts shift the benefits to the region where the marginal benefits have decreased considerably (by concavity of $p$). If both agents have similar and small head starts, they will exert effort until the total effort equals $x^*(\theta)$, the solution of a one shot contest. However, if the difference in head starts is large enough and one of the agents exerts more than $x^*(\theta)$ in the first contest, the other agent does not exert enough effort to reach $x^*(\theta)$. In fact, the agent that is behind exerts strictly less effort compared to the agent that is ahead. Therefore, in the second stage we observe asymmetric equilibria when the head starts are asymmetric.

We have four relevant regions, described in Figure 3.2. For each one of these regions, we derive the corresponding equilibrium payoffs in the second stage.

Figure 3.2. Regions for different cases of second period payoffs.

(1) Region I = \{(W_1, W_2) : \max\{W_1, W_2\} < x^*(\theta_2)\}. In this region, both agents have exerted less effort than the equilibrium effort of the static contest in the
second period. Thus, agents will exert effort until they reach that static equilibrium level \( x^*(\theta_2) \). In the symmetric equilibrium, each agent wins the prize \( V_2 \) with probability \( \frac{1}{2} \). Hence, the period 2 payoff for each agent is given by:

\[
\pi_1(W_1, W_2) = \frac{V_2}{2} - c_2x^*(\theta_2) + c_2W_1
\]

\[
\pi_2(W_1, W_2) = \frac{V_2}{2} - c_2x^*(\theta_2) + c_2W_2
\]

(2) Region II.1 = \{ \( (W_1, W_2) : W_1 > x^*(\theta_2), W_2 < R(W_1|\theta_2) \) \}. In this region, only agent 2 has exerted less effort than the equilibrium effort of the static contest in the second period. Agent 1 has exerted ‘too much’ effort in the previous contests. Thus, agent one does not exert effort and agents 2 best responds to \( W_1 \). Period 2’s payoff for each agent is given by:

\[
\pi_1(W_1, W_2) = V_2p(W_1, R(W_1|\theta_2))
\]

\[
\pi_2(W_1, W_2) = V_2p(R(W_1|\theta_2), W_1) - c_2 \cdot (R(W_1|\theta_2) - W_2)
\]

(3) Region II.2 = \{ \( (W_1, W_2) : W_1 < R(W_2|\theta_2), W_2 > x^*(\theta_2) \) \}. Analogous to the previous case, only agent 1 will exert effort. Period 2’s payoff for each agent is given by:

\[
\pi_1(W_1, W_2) = V_2p(R(W_2|\theta_2), W_2) - c_2 \cdot (R(W_2|\theta_2) - W_1)
\]

\[
\pi_2(W_1, W_2) = V_2p(W_2, R(W_2|\theta_2))
\]

(4) Region III = \{ \( (W_1, W_2) : W_1 > x^*(\theta_2), W_2 > x^*(\theta_2) \) \}. Both agents have exerted more effort than the equilibrium in the static game, therefore they exert zero
effort in this period. Period 2’s payoff for each agent is given by:

$$\pi_1(W_1, W_2) = V_2 p(W_1, W_2)$$
$$\pi_2(W_1, W_2) = V_2 p(W_2, W_1)$$

It is important to notice that, conditional on history, the set of possible deviations in the second stage are not limited to deviations within one region only. Some papers, such as Baik and Lee (2000), compute the equilibrium assuming that agents only choose deviations within region I. This is what we call ‘local deviations’. We do not impose this restriction and we look for equilibrium considering all possible deviations. For example, if $W_1 = 0$ and agent 2’s effort is low, agent 1 can obtain the payoff of region I exerting some effort, or region II.1 exerting more effort, or region III exerting even more effort. By the properties of the game, the payoff function within each region is continuous. Lemma 23 shows the payoff is continuous everywhere.

**Lemma 23.** The continuation payoff $\pi_i(W_i, W_j)$ is continuous.

### 3.3.3. Endogenous Head Starts

In the first stage, agents choose their level of effort considering three effects: Firstly, effort will be used to compete in for a prize $V_1$ in the contest of period 1. Secondly, effort exerted in the first period will be used as a head start in the second contest. Thirdly, the cost of effort today is $c_1$ but tomorrow is $c_2$. Thus, in the first stage, agent $i$ solves:

$$\max_{W_i \geq 0} V_1 \cdot p(W_i, W_j) - c_1 \cdot W_i + \pi_i(W_i, W_j)$$
The continuation payoffs depend on the region where \((W_i, W_j)\) belong to. Thus, we have to consider which region the agents will end up at the second period, giving the effort exerted at the first period.

- The first period payoff in Region I is given by:

\[
V_1 p(W_i, W_j) - (c_1 - c_2)W_i + U_2^*,
\]

where \(U_2^* = \frac{V_2}{2} - c_2x^*(\theta_2)\) is the equilibrium payoff in a standard Tullock contest with parameter \(\theta_2\). Notice that, independent of the allocation inside region I, the agents will always reach the static Nash equilibrium in the second stage by exerting the additional effort in stage two. Because each unit of effort in period 1 is freely used in period 2, the total cost of the aggregate effort is \(W_i\) at cost \(c_1\) and \(x^* - W_i\) at cost \(c_2\). Thus, the net cost of each unit in the first period is \(c_1 - c_2\).

- The first period payoff in Region II.i is given by:

\[
V_1 p(W_i, W_j) - c_1W_i + V_2 p(W_i, R(W_i|\theta_2)).
\]

An allocation inside region II.i corresponds to the payoff of the contest of the first stage plus the payoff of the leader of a Stackelberg game. Thus, in this region, agent \(i\)'s effort is used to compete for the prize \(V_1\) in the period 1, and freely used in the second period as the investment of a Stackelberg leader.

- The first period payoff in Region II.j is given by:

\[
V_1 p(W_i, W_j) - (c_1 - c_2)W_i + V_2 p(R(W_j|\theta_2), W_j) - c_2 \cdot R(W_j|\theta_2).
\]
An allocation inside region II,j corresponds to the payoff of the contest of the first stage plus the payoff of the follower of a Stackelberg game. Agent i’s effort is chosen to compete in the contest of period 1, but also by taking into account that in next period he/she would be the follower of a Stackelberg game, and that he/she will have to pay a cost of $c_2$ for each additional unit of effort.

- The first period payoff in Region III is given by:

$$ (V_1 + V_2) \cdot p(W_i, W_j) - c_1 \cdot W_i $$

An allocation inside region III corresponds to playing the two sequential contest with the same effort level. Because effort in the first period must be freely used in the second period, agents choose effort in the first period as if they were playing a single contest with prize $V_1 + V_2$ and marginal cost of effort equal to $c_1$.

Given these payoffs, we can now compute the equilibrium level of effort at the first contest. Proposition 24 analyzes the case $c_2(V_1 + V_2) < c_1 V_2$ and Proposition 25 the case $c_2(V_1 + V_2) \geq c_1 V_2$. Putting these results together, Proposition 26 characterizes the subgame perfect equilibrium of the game. All of the proofs are in the Appendix.

**Proposition 24.** When the condition $c_2(V_1 + V_2) < c_1 V_2$ holds, agent i’s best response in the first contest is

$$ BR_i(W_j) = R(W_j | \theta_I), \text{ where } \theta_I = \frac{V_1}{c_1 - c_2}. $$

By properties of the best response function, there is a unique equilibrium in pure strategies given by $x^*(\theta_I)$. 
When condition $c_2(V_1 + V_2) < c_1V_2$ holds, the unique equilibrium in the first stage and corresponds to $x^*(\theta_I)$. That is, the first stage of the sequential contest with effort accumulation is (strategically) equivalent to a static Tullock contest with parameter $\theta_I$. Also, this condition implies that $\theta_I < \theta_2$ and therefore, by Proposition 22, the second stage equilibrium effort is $x^*(\theta_2) - x^*(\theta_I)$. The aggregate effort in both periods is $x^*(\theta_2)$.

**Proposition 25.** When $W_2 \geq x^*(\theta_2)$ and the condition $c_2(V_1 + V_2) \geq c_1V_2$ holds, agent 1’s best response in the first contest is

$$BR_1(W_2) = R(W_2|\theta_{III}), \text{ where } \theta_{III} = \frac{V_1 + V_2}{c_1}.$$ 

When agent $j$’s effort is above $x^*(\theta_2)$, agent $i$’s best response is to play as if the prize was $V_1 + V_2$. In this case, because agent $j$ has exerted so much effort already, in the next contest agent $i$ will not exert extra effort. With these results we can characterize the unique subgame perfect equilibrium of the game.

**Proposition 26.** The unique subgame perfect equilibrium is symmetric and given by

**Effort in the first contest:**

$$W^*_i = \begin{cases} 
  x^*(\theta_I) & \text{if } c_2(V_1 + V_2) < c_1V_2 \\
  x^*(\theta_{III}) & \text{if } c_2(V_1 + V_2) \geq c_1V_2.
\end{cases}$$

**Effort in the second contest:**

$$x^*_i = \begin{cases} 
  x^*(\theta_2) - x^*(\theta_I) & \text{if } c_2(V_1 + V_2) < c_1V_2 \\
  0 & \text{if } c_2(V_1 + V_2) \geq c_1V_2.
\end{cases}$$
Intuitively, firms would want to exert as much effort as possible in the first contest, because that investment is used in both periods. However, firms need to trade-off the competition and cost-saving incentives. If the cost of exerting effort in the second contest is low, firms are not willing to pay for all the effort at the first contest. The marginal cost of effort in the first contest is reduced by the investment in cost saving at the second contest. The net cost of effort is then $c_1 - c_2$. The balance between competition and cost-savings is given by the condition

$$c_2(V_1 + V_2) < c_1V_2 \iff \frac{c_2}{V_2} < \frac{c_1}{V_1 + V_2}.$$ 

In particular, it is not enough that $c_2 < c_1$ to guarantee positive effort in the second contest. The marginal cost over the size of the potential prize is what drives the trade-off. In the next section, we show that restricting attention to the widely used functional form $g(x) = x^\alpha$ is not without loss of generality. In particular, we focus on two important characteristic of a contest: rent dissipation and aggregate effort. By allowing increasing and concave scoring functions (not necessarily homogeneous), we show that the shape of this function has drastic implications on the rent dissipation and the aggregate effort in sequential contests where effort accumulates.

### 3.4. Rent Dissipation and Aggregate Effort across Contests

In this section, we use the equilibrium found in the previous section to study rent dissipation and aggregate effort across contests with and without carryover. Notice that if the two contest have the same prize and cost of effort, then the rent dissipation and the aggregate effort in both contests are the same (up to a constant). However, when prizes
and the cost of effort is different across contests, these two concepts do not necessarily coincide. We show that period-specific costs and the functional form of the score function are important objects to determine the amount of rent dissipation and aggregate effort.

We will compare the sequential contest with effort accumulation with a sequential contests without effort accumulation. The unique pure strategy equilibrium of a sequential contests without effort accumulation, such that contest $t$ has prize $V_t$ and cost $c_t$, is given by the standard static Tullock symmetric equilibrium $x^*(\theta_t)$, where $\theta_t = \frac{V_t}{c_t}$.

### 3.4.1. Rent Dissipation

The rent dissipation is defined as agents’ expenditure over the sum of the prizes. In sequential contests (with and without effort accumulation), it is given by

$$\text{Rent Dissipation} = 2 \cdot \begin{cases} \frac{c_1 x^*(\theta_1) + c_2 x^*(\theta_2)}{V_1 + V_2} & \text{without effort accumulation}, \\ \frac{c_1 W_t^* + c_2 x^*_t}{V_1 + V_2} & \text{with effort accumulation}, \end{cases}$$

where $W_t^*$ and $x_t^*$ are given in Proposition 26. The amount of rents dissipated in contests with effort accumulation depend of the shape of $x^*(\theta)$. Lemma 27 establishes the relation between the score function $g(\cdot)$, a the primitive of the model, and the convexity of $x^*(\theta)$. It also derives conditions for the monotonicity of the equilibrium expenditure.

**Lemma 27.** Let $g(\cdot)$ be increasing and concave, $F(x) = \frac{g(x)}{g'(x)}$ and $h(c) = c \cdot x^* \left( \frac{V}{c} \right)$. Then,

1. $\text{sign} \left( \frac{d^2 x^*(\theta)}{d\theta^2} \right) = -\text{sign} \left( F''(x^*(\theta)) \right)$.

2. $h'(c) > 0$ if and only if $F''(x) > 0$.
The first part of the lemma established conditions on the primitives that guarantee convexity of the equilibrium level of effort \( x^*(\theta) \). The second part of the proposition establishes that the expenditure of agents in a static Tullock contest with parameter \( \theta = \frac{V}{c} \) is increasing when the cost of effort increases. The next proposition compares the rent dissipation in contests with and without effort accumulation.

**Proposition 28.** Let \( F(x) = \frac{g(x)}{g'(x)} \). In a contest with effort accumulation:

1. When \( c_2(V_1 + V_2) < c_1V_2 \), there is more (less) rent dissipation than in a contest without effort accumulation iff \( F'' > 0 \) (<0).

2. When \( c_2(V_1 + V_2) > c_1V_2 \), there is more rent dissipation than in a contest without effort accumulation if \( c_1 > c_2 \) and \( F'' < 0 \), and less rent dissipation if \( c_1 > c_2 \) and \( F'' < 0 \).

**Proof.** Assume that the condition \( c_2(V_1 + V_2) < c_1V_2 \) holds. The total expenditure in a sequential contest with effort accumulation for each agent is \( (c_1 - c_2)x^*(\theta_1) + c_2x^*(\theta_2) \), while in a contest without effort accumulation is \( c_1x^*(\theta_1) + c_2x^*(\theta_2) \). Agents spend more (dissipate more rents) in with effort accumulation if and only if

\[
(c_1 - c_2)x^* \left( \frac{V_1}{c_1 - c_2} \right) > c_1x^* \left( \frac{V_1}{c_1} \right).
\]

By Lemma 27, \( c \cdot x^* \left( \frac{V}{c} \right) \) is increasing, and \( x^*(\cdot) \) is concave iff \( F'' > 0 \). By the monotonicity of \( h(c) \) and the fact that \( c_1 - c_2 < c_1 \), we obtain the result.

For the second part, note that if \( x^*(\cdot) \) is convex, and \( x^*(0) = 0 \), for any \( \lambda \in (0, 1) \) we have \( x^*(\lambda \theta) \leq \lambda x^*(\theta) \). Then, taking \( \lambda = \frac{a}{a+b} \) we have \( x^*(a) + x^*(b) \leq x^*(a+b) \), since \( a = \lambda(a+b) \) and \( b = (1-\lambda)(a+b) \). Since \( F'' > 0 \), we also know that \( h(c) = c x^* \left( \frac{V}{c} \right) \) is
increasing. When \( c_2 < c_1 \) we have,

\[
c_1 x^* \left( \frac{V_1}{c_1} \right) + c_2 x^* \left( \frac{V_2}{c_2} \right) < c_1 \left[ x^* \left( \frac{V_1}{c_1} \right) + x^* \left( \frac{V_2}{c_1} \right) \right] \\
\leq c_1 x^* \left( \frac{V_1 + V_2}{c_1} \right)
\]

Similarly, when \( x^*(\cdot) \) is concave and \( c_1 < c_2 \) we obtain the other result.

\[
c_1 x^* \left( \frac{V_1}{c_1} \right) + c_2 x^* \left( \frac{V_2}{c_2} \right) > c_1 \left[ x^* \left( \frac{V_1}{c_1} \right) + x^* \left( \frac{V_2}{c_1} \right) \right] \\
\geq c_1 x^* \left( \frac{V_1 + V_2}{c_1} \right)
\]

□

This result establishes the relation between the functional form of \( g(\cdot) \) and the rent dissipation. Traditionally, the literature has focused on \( g(x) = x^\alpha \) and identical costs across periods. Corollary 29 shows that, when \( g(x) = x^\alpha \), the total rent dissipation equals \( \frac{\alpha^2}{2} \) regardless of the effort accumulation.

**Corollary 29.** When \( g(x) = x^\alpha \), the rent dissipation in sequential contests with and without effort is the same and equals

\[
Rent Dissipation = \frac{\alpha}{2}.
\]

**Proof.** When \( g(x) = x^\alpha \), \( F(x) = \frac{x}{\alpha} \) and \( F''(x) = 0 \). By Proposition 28, the rent dissipation is the same in the contests with or without effort accumulation. □

When \( g(x) = x^\alpha \), the equilibrium level of effort \( x^*(\theta) \) is linear with respect to the contest parameter \( \theta \). In this case, dynamic considerations either from cost savings or
effort accumulation do not affect the amount of rent dissipation. However, as Proposition 29 shows, outside from the standard repeated Tullock contest setting, in general, both dynamic considerations affect the amount of rents dissipated.

3.4.2. Aggregate Effort across Contests

Similarly to the rent dissipation results, we can study the aggregate effort in contests with and without effort accumulation. In many applications, the variable of interest is precisely the aggregate effort, so understanding the role of effort accumulation is relevant. In sequential contests without effort accumulation, the aggregate equilibrium effort for each agent equals

\[ x^\ast (\theta_1) + x^\ast (\theta_2) . \]

The following proposition compares the aggregate level of effort in sequential contests with and without effort accumulation.

**Proposition 30.** Let \( F(x) = \frac{g(x)}{g'(x)} \). The aggregate effort across time in contests with effort accumulation:

1. When condition \( c_2(V_1 + V_2) < c_1 V_2 \) holds, is lower than the aggregate effort without accumulation, and independent of the shape of the score function \( g(\cdot) \).
2. When condition \( c_2(V_1 + V_2) \geq c_1 V_2 \) holds, is higher (lower) if \( F'' < 0 \ (>0) \) and \( c_1 < c_2 \ (c_1 > c_2) \) than the aggregate effort without accumulation.

**Proof.** When condition \( c_2(V_1 + V_2) < c_1 V_2 \) holds, and there is effort accumulation, the aggregate equilibrium effort for each agent equals \( x^\ast \left( \frac{V_1}{c_2} \right) \), which is obviously lower than the aggregate effort without accumulation.
When condition $c_2(V_1 + V_2) \geq c_1 V_2$ holds, and there is effort accumulation, the aggregate equilibrium effort for each agent equals $x^\ast\left(\frac{V_1 + V_2}{c_1}\right)$. The comparison between these equilibrium levels will depend on the convexity of $x^\ast(\cdot)$ and the cost of effort. If $x^\ast(\cdot)$ is convex, since $x^\ast(0) = 0$, for any $\lambda \in (0, 1)$ we have $x^\ast(\lambda \theta) \leq \lambda x^\ast(\theta)$. Then $x^\ast(a) + x^\ast(b) \leq x^\ast(a + b)$, since $a = \frac{a}{a+b}(a+b)$ and $b = \frac{b}{a+b}(a+b)$. Thus, if $c_1 < c_2$ we have

$$x^\ast\left(\frac{V_1}{c_1}\right) + x^\ast\left(\frac{V_2}{c_2}\right) \leq x^\ast\left(\frac{V_1}{c_1} + \frac{V_2}{c_2}\right) < x^\ast\left(\frac{V_1 + V_2}{c_1}\right).$$

Similarly, when $x^\ast(\cdot)$ is concave and $c_1 > c_2$ we have:

$$x^\ast\left(\frac{V_1}{c_1}\right) + x^\ast\left(\frac{V_2}{c_2}\right) \geq x^\ast\left(\frac{V_1}{c_1} + \frac{V_2}{c_2}\right) > x^\ast\left(\frac{V_1 + V_2}{c_1}\right).$$

Proposition 30 shows that the aggregate effort across time in contests with effort accumulation can be higher or lower than the aggregate effort in contests with no accumulation. In the first case, when agents exert effort in both periods, the aggregate effort is always lower when effort accumulates and this is independent of shape of the score function $g$. The reason is that the cost saving incentive dominates the competition effect, so agents in total exert the same effort as a single static contest with parameter $\theta_2$. However, when cost of effort in the first contest is lower than the cost on the second contest, agents have incentives to exert all the effort in the first period. Depending on the convexity $x^\ast(\cdot)$, characterized in Lemma 27, the aggregate effort with effort accumulation can be higher or lower than in sequential contests without accumulation.
3.5. Extension: Head Starts with Variable Carryover rate

In this section, we analyze situations in which the effort exerted in the first contest does not perfectly carries over to the second on. We study sequential contests where only a fraction $\delta$ of the effort in the first contest is carried over to the second one. We allow for $\delta \in [0, \infty)$, allowing for depreciation of effort ($0 < \delta < 1$) and also growth of effort ($\delta > 1$). For simplicity, we restrict to the case $g(x) = x^\alpha$ with $\alpha < 2$.

**Proposition 31.** Consider the score function $g(x) = x^\alpha$, with $\alpha < 2$. If a proportion $\delta \in [0, \infty)$ of the effort exerted in the first contest accumulates to the second one, there is a unique equilibrium, which is symmetric, characterized by first contest’s effort:

$$
(\delta W_i)^* = \begin{cases} 
x^*(\theta^\delta_{I}) & \text{if } \delta c_2(V_1 + V_2) < c_1 V_2 \\
x^*(\theta^\delta_{III}) & \text{if } \delta c_2(V_1 + V_2) \geq c_1 V_2
\end{cases}
$$

and second contest’s effort:

$$
x^*_i = \begin{cases} 
x^*(\theta_2) - x^*(\theta^\delta_{I}) & \text{if } \delta c_2(V_1 + V_2) < c_1 V_2 \\
0 & \text{if } \delta c_2(V_1 + V_2) \geq c_1 V_2
\end{cases}
$$

where $\theta^\delta_{III} = \delta \frac{V_1 + V_2}{c_1}$, $\theta_2 = \frac{V_2}{c_2}$ and $\theta^\delta_{I} = \delta \frac{V_1}{c_1 - \delta c_2}$.

**Proof.** The proof is a direct consequence of the previous propositions. Proposition 22 gives us the best response in the second contest for arbitrary head starts. Evaluating at $(\delta W_i, \delta W_j)$, we obtain the second stage equilibrium. For the first stage, we work with the variables $(\delta W_i, \delta W_j)$ instead of $(W_i, W_j)$. When $g(\cdot)$ is homogeneous (of any degree), notice that $p(\delta W_i, \delta W_j) = p(W_i, W_j)$. The first stage payoffs with carryover rate $\delta$ are
given by
\[ \max_{W_i \geq 0} V_1 \cdot p(\delta W_i, \delta W_j) - \frac{c_1}{\delta} \cdot \delta W_i + \pi_i(\delta W_i, \delta W_j) \]

With a change of variables, this is the same problem we solved for in Proposition 26, but with a different cost \( \frac{c_1}{\delta} \). Directly from Proposition 26 we find the result. \( \square \)

The result in this proposition is intuitive. For a low (high) value of the parameter \( \delta \), agents exert less (more) effort in the first period. Although we are restricting to the widely used case \( g(x) = x^\alpha \), losing the effects of accumulation of effort on rent seeking and aggregate effort, we can illustrate how different carry over rates affect rent dissipation and aggregate effort.

**Corollary 32.** When \( g(x) = x^\alpha \), the rent dissipation in sequential contests with and without effort is the same and equals \( \frac{\alpha}{2} \) for any carryover rate \( \delta \in [0, 1] \).

**Proof.** We can compute the level of rent dissipation by explicitly solving for the equilibrium levels in Proposition 31. \( \square \)

Corollary 32 shows that for the case \( g(x) = x^\alpha \) the carryover rate has no effect on the amount of rents dissipated.

To study aggregate effort, we define \( \delta^* = \frac{c_1 V_2}{c_2(V_1 + V_2)} \) as the threshold carryover value at which effort switches from being exerted in the first contest only, to effort being exerted in both contests. In this case, Proposition 31 implies the following closed form solution:

\[
W_i^* = \begin{cases} 
\frac{\alpha V_1}{4(c_1 - \delta c_2)} & \text{if } \delta < \delta^* \\
\frac{\alpha (V_1 + V_2)}{4c_3} & \text{if } \delta \geq \delta^*
\end{cases}, \quad 
 x_i^* = \begin{cases} 
\frac{\alpha (c_1 V_2 - \delta c_2 (V_1 + V_2))}{4c_2(c_1 - \delta c_2)} & \text{if } \delta < \delta^* \\
0 & \text{if } \delta \geq \delta^*
\end{cases}
\]
From Proposition 30, we know that allowing for effort accumulation affects the level of aggregate effort across contests. The next corollary, derives the aggregate effort as a function of $\delta$ for the case $g(x) = x^\alpha$.

**Corollary 33.** When $c_1 < c_2$ (or $c_1 > c_2$) the aggregate effort across contests increases (decreases) for $\delta \in [0, \delta^*)$ and is constant $\delta > \delta^*$.

Figure 3.3 shows the aggregate effort for different cases. When $\delta > \delta^*$ agents exert all the effort at the first contest, because the carryover is strong enough. When $\delta$ is small, agents have incentives to exert effort in both contests.

![Figure 3.3. Aggregate effort for the two contests as a function of $\delta$.](image)

3.6. Conclusion

This paper studies sequential contests with effort accumulation. Our model allows for heterogeneous contest characteristics (prizes and costs), and a contest success function that is not necessarily homogeneous, the typical assumption in the literature. As a first contribution, we characterized the subgame perfect equilibrium of the game, taking into account all possible deviations, i.e. not restricting the parameters to obtain an interior solution. Depending on the parameters of the model, only two types of equilibria can
arise. In one case, all the effort is exerted in the first contest. This occurs when agents expect higher costs of effort in the second contest, so the incentive to accumulate effort dominates. Instead, if agents expect lower cost of effort in the second contest, effort is spread among the two contests. Agents need some effort to compete in the first period, but they do not have incentives accumulate effort, since effort is cheaper in the second contest. We show that this basic trade-off of competition and cost savings is governed by a condition that involves the cost of effort normalized by the value of the potential prizes. Thus, departing from a setting in which all contests have the same prizes and costs allow us to discover new dynamic equilibrium behavior. A similar result was obtained by Pal (1991) in a Cournot duopoly setting instead of a contest.

Our second contribution is to study to what extent effort accumulation affects the rent dissipation and the aggregate effort. We find that the functional form used in the contest, as well as the cost of effort, govern the amount of rents dissipated and the aggregate effort. We show that under the traditional contest success functions $p(x, y) = \frac{x^{\alpha_{x}} y^{\alpha_{y}}}{x^{\alpha_{x}} + y^{\alpha_{y}}}$, the effect of accumulating effort on rent dissipation and aggregate effort vanishes. Non-homogeneity arises in some cases when the score players use to compete is the outcome of a production function with multiple inputs (Rai and Sarin (2009)). In particular, we show effort accumulation affects rent dissipation only when $x^{\ast}(\theta)$—the equilibrium level of effort in a static contest of parameter $\theta$—is convex (or concave). This can only happen if we consider scoring functions outside the class of homogeneous functions. Under the traditional homogeneity assumption, $x^{\ast}(\theta)$, is linear in $\theta$ and the effect of accumulating effort on rent dissipation and aggregate effort disappears. We provide conditions on
the primitives of the model that characterize the effort accumulation affects the rent
dissipation and the aggregate effort.

Finally, we provide some extensions when effort of one contest is not perfectly carried
over to the next one. Further extensions of our model could incorporate asymmetries
among players, different number of periods and players, or private information. Adding
more periods or players is challenging, since the number of continuation games grows
rapidly with the number of players or periods. However, with more periods, we the main
intuition will be preserve: a condition that determines when to exert equilibrium will
involve the cost of effort over prizes in the period where that effort is used up. Adding
private information seems a research direction that could bring novel insights. We leave
these questions open for future work.
References


APPENDIX A

Ex-ante Expected Utility in an All-Pay Auction

A.1. Appendix I

A.1.1. One strong uninformed player

In this case we assume there is one strong (uninformed) player with $V_1 = \{v\}$, and a weak (informed) one with $V_2 = \{v_1, ..., v_N\}$, where $v > v_{j+1} > v_j, \forall j \in \{1, 2, ..., N - 1\}$. We define $p_j$ as the probability of the informed player having valuation $v_j$.

We are interested on finding the distribution of types that maximizes the expected payoff of the informed player.

A.1.1.1. Equilibrium. The equilibrium in an all pay auction is unique and can be calculated with the same algorithm we have used so far.

Define $i^*$ as the lowest type of player 2 that bids above zero with positive probability. This type will satisfy:

\begin{itemize}
  \item $v_{i^*} \left(1 - \sum_{i > i^*} p_i \frac{v}{v_i}\right) < p_{i^*} v$
  \item $p_j v < v_j \left(1 - \sum_{i > j} p_i \frac{v}{v_i}\right), \forall j > i^*$
\end{itemize}

Define $d_i$ as the upper bound of the BR set of the informed player when its type is $v_i$. Then for every $i < i^*$, $d_i = 0$. 

\[
d_i = \begin{cases} 
0 & i < i^* \\
v_i^* \left(1 - \sum_{i > i^*} p_i \frac{v_i}{v_i^*}\right) & i = i^* \\
d_{i-1} + p_i v & i > i^*
\end{cases}
\]

And finally, the eq. strategies are mixed, with density functions:

\[
g_1(x) = \begin{cases} 
\frac{1}{v_i} & x \in [d_{i-1}, d_i] \\
0 & \text{o.w.}
\end{cases}
\]

And for the informed one, when \( i \geq i^* \)

\[
g_2(x, i) = \begin{cases} 
\frac{1}{p_i v} & x \in [d_{i-1}, d_i] \\
0 & \text{o.w.}
\end{cases}
\]

For this player, if the signal is lower than \( i^* \), she will not participate (bid 0 with probability one). Also, in case she is \( i^* \), she will randomize in zero with probability

\[
1 - \frac{v_i^*}{p_{i^*} v} \left(1 - \sum_{i > i^*} p_i \frac{v_i}{v_i^*}\right)
\]

The payoffs of the informed player when she has type \( i \) is:

\[
U_i = \begin{cases} 
0 & i \leq i^* \\
G_1(d_i)v_i - d_i = G_1(d_{i-1})v_i - d_{i-1} & i > i^*
\end{cases}
\]
With

\[
G_1(d_i) = \begin{cases} 
0 & i < i^* \\
1 - \sum_{j > i} \frac{p_j v_j}{v_i} & \text{o.w.}
\end{cases}
\]

And the uninformed player has an expected payoff of

\[
U = v - d_N
\]

A.1.1.2. Optimal Probabilities. If we want to know the distribution that will maximize the expected payoff of the informed player, \((E(U) = \sum p_j U_j)\), we don’t want to deal with changes in \(i^*\). The following lemma helps us with that:

Lemma 34. In this environment, a game with \(i^* > 1\) is payoff equivalent to one in which \(p_i^* = 0, \forall i < i^*, \ p_i^* = p_i, \forall i > i^*\) and \(p_i^* = \sum \ p_i\)

Proof. This is almost direct. We are changing probabilities of \(i \leq i^*\), but that does not change any \(d_j, j \geq i^*\), nor \(G_1(d_i)\) for any \(i > i^*\). Therefore all payoffs for types \(j > i^*\) are the same. Type \(i^*\) still get’s zero, and the uninformed player keeps getting the same expected payoff. \(\square\)

This allows us to focus our attention to cases when \(i^* = 1\), when all types make player 1 to participate.

Given a set of \(N\) types, we can select subsets of that, and maximize subject to all types participating. Then, get the subset that reaches the maximum expected payoff and get the optimal distribution. This procedure might seem difficult but for the next section makes things much easier. Each maximization problem is of a continuous function in a compact set, therefore it is guaranteed to have a solution. And then, by taking the max
of a finite sequence, the maximum will always exist. Also, it is easy to check that in this maximization problem the objective functions are infinitely differentiable.

We will prove the previous statements next:

A.1.1.3. Existence.

**Lemma 35.** Solving the original problem is equivalent to solving \( \max_{i=1,\ldots,N} G_i \), where \( G_i \) is defines as the maximum expected utility when the only available types are the ones greater or equal than \( i \), and type \( i \) is "almost" always participating. The condition for the first thing to hold is \( \sum_{j \geq i} p_j = 1 \) and for the second to hold is the previous condition plus \( 1 > \sum_{j > i} p_j \frac{v_j}{w_j} \).

**Proof.** By the previous lemma, the original problem is equivalent to solving \( \max_{i=1,\ldots,N} \hat{G}_i \), where \( \hat{G}_i \) is the maximum expected utility when the only available types are the ones greater or equal than \( i \), and type \( i \) is always participating. The condition for the first thing to hold is \( \sum_{j \geq i} p_j = 1 \) and for the second to hold is the previous condition plus \( 1 > \sum_{j > i} p_j \frac{v_j}{w_j} \).

What we want to prove is that if we change the strict inequality we still have the same thing. In case we have the non strict inequality, then it is like saying that the \( i \) does not participate, but the next type is participating. Therefore if the maximum of \( G \) is attained when we have an equality on the second condition, by the previous lemma, such case is included when we maximize conditional on the next type being the first to participate. Therefore we can use the non strict inequality instead of the strict one. \( \square \)

Now, in every sub problem, as we said before, we are maximizing a continuous function in a compact set, therefore a solution exist, and finally a solution to the original problem also exists.
A.1.1.4. Optimal Dist. We will show that the solution involves weights only on both the highest and lowest type. The way to prove this will be in the following way. First, we will show that in the sub optimization problems we presented before (max expected utility conditional on the first type participating) for \( N \) different types, if \( N > 3 \) then the solution involves \( N - 1 \) types with positive probabilities. After that we will show that for \( N = 3 \) the solution has weight in only 2 types, and will calculate the expected payoff and how does that depends on the possible types. This shows that for the original problems, the solution must involve only two probabilities greater than zero. Then we will get the pair that maximizes the payoff.

Lemma 36. Conditional on the lowest type "almost" always participating, if \( N > 3 \), there is no interior solution. In other words, either \( p_j = 0 \) for some \( j \), or the participation condition holds with equality, which means that the first type does not participate.

Proof. See Appendix II. \( \square \)

The solution not being interior can mean three things. First, one of the probabilities can be 1. This is not possible because it leaves the informed player with zero payoff, and therefore there are other option that are better. Second, the solution involves at least one probability that is 0. This is, the solution puts not negative weight in at most \( N - 1 \) types. And third, the first type is not participating. In such case, using lemma 1, we have that the solution is equivalent to having no positive probability on the first type.

With this, using an induction argument, we have showed that the solution the original problem can not have weight in more than 3 different types. Now we will see what is the optimal thing to do in such case.
Lemma 37. Suppose the informed player can have payoff \( v_1, v_2 \) or \( v_3 \). Then the optimal distribution is \( p_3 = \frac{v_3}{2v} \), \( p_2 = 0 \), which yields expected payoff \( \frac{v_3}{4v}(v_3 - v_1) \)

Proof. See Appendix II \( \square \)

This lemma gives us the final step. The implication is that the optimal distribution will have positive weight only in 2 different types. Also, with the exact formula for the final distribution, we can decide which of the candidates is the one that gives the greatest expected payoff. \( \frac{v_3}{4v}(v_3 - v_1) \) is increasing in \( v_3 \), but decreasing in \( v_1 \), therefore the optimal distribution will have positive weight in the lowest and highest types, with \( p_l = 1 - p_h \), \( p_h = \frac{v_3}{2v} \).

A.1.2. One strong informed player

In this case we are going to assume that the strong player is the informed one. The uninformed player has \( V_1 = \{v\} \), and the informed one \( V_2 = \{v_1, ..., v_N\} \), where \( v < v_j < v_{j+1} \), \( \forall j \in \{1, 2, ..., N - 1\} \). Again, we define \( p_j \) as the probability of the informed player having valuation \( v_j \).

We are interested on finding the distribution of types that maximizes the expected payoff of the informed player.

A.1.2.1. Equilibrium. It is easy to check that the unique equilibrium will induce all types of player 2 to be active. This comes from

\[
\frac{1}{v} - \sum p_i \frac{p_i}{v_i} > 0
\]
Because $v < v_i, \forall i$. This implies that the upper bound of the best response set of the informed player when her type is $v_i$ is 

$$d_i = \sum_{j \leq i} p_j v,$$

for $i \in \{1, ..., N\}$, $d_0 = 0$. And the density functions of the eq. strategies are:

$$g_1(x) = \begin{cases} 
\frac{1}{v_i} & x \in [d_{i-1}, d_i] \\
0 & \text{o.w.}
\end{cases}$$

for the uninformed player, and

$$g_2(x, i) = \begin{cases} 
\frac{1}{p_i v} & x \in [d_{i-1}, d_i] \\
0 & \text{o.w.}
\end{cases}$$

for the informed one when her type is $v_i$

Also, player 1 will have an atom at zero, of size $1 - \sum_{i \geq 1} \frac{p_i v}{v_i}$

**A.1.2.2. Optimal probabilities.** We want to find out which is the distribution will leave player 2 with the highest ex-ante expected utility. To do that we will first show that if we take a little mass from the probability of a low type and give it to a higher type, then the utility of each type of player 2 increases. This implies that if we have two distributions, then if one stochastically dominates the other, the first leaves player 2 with a greater ex-ante expected utility. Finally, because the distribution that has all the mass on the highest type stochastically dominates all the others, that one maximizes the ex-ante expected utility of player 2.

Let's start with the first statement:
Lemma 38. Given the distribution of types \( \{ p_i : i = 1, \ldots, n \} \), then a new distribution \( \{ p'_i : i = 1, \ldots, n \} \) with \( p'_k = p_k - e, p'_j = p_j + e, p_i = p'_i \) for \( i \neq j, k \), and \( j > k \), will induce higher or equal utility for each possible type of the informed player.

Proof. The utility of player 2 when her type is \( v_i \) under the original distribution is

\[
\left( 1 - \sum_{l>i} \frac{p_l v}{v_l} \right) v_i - \sum_{l \leq i} p_l v
\]

Now we have several cases to check. If \( i < k \), then

\[
\left( 1 - \sum_{l>i} \frac{p'_l v}{v_l} \right) v_i - \sum_{l \leq i} p'_l v = \left( 1 - \sum_{l>i} \frac{p_l v}{v_l} + e v \left( \frac{1}{v_k} - \frac{1}{v_j} \right) \right) v_i - \sum_{l \leq i} p_l v
\]>

\[
\left( 1 - \sum_{l>i} \frac{p_l v}{v_l} \right) v_i - \sum_{l \leq i} p_l v
\]

If \( k \leq i < j \):

\[
\left( 1 - \sum_{l>i} \frac{p'_l v}{v_l} \right) v_i - \sum_{l \leq i} p'_l v = \left( 1 - \sum_{l>i} \frac{p_l v}{v_l} \right) v_i - \sum_{l \leq i} p_l v - e v \frac{v_i}{v_j} + ev
\]>

\[
\left( 1 - \sum_{l>i} \frac{p_l v}{v_l} \right) v_i - \sum_{l \leq i} p_l v
\]

And finally, if \( i \geq j \), then

\[
\left( 1 - \sum_{l>i} \frac{p'_l v}{v_l} \right) v_i - \sum_{l \leq i} p'_l v = \left( 1 - \sum_{l>i} \frac{p_l v}{v_l} \right) v_i - \sum_{l \leq i} p_l v
\]

\( \square \)
This lemma is saying that if we can it is always better to have more probability in the biggest type. The implication that

**A.1.2.3. Alternative.** The utility that the informed player gets when her type is $v_i$ is constant in her best response set, therefore, we can calculate it at $d_i$ or $d_{-i}$, and it is going to be

$$U_i = G_i v_i - d_i = G_{i-1} v_i - d_{i-1}$$

Where $G_i$ is the probability that player 1 makes effort lower or equal than $d_i$. It is straightforward to see that $U_i$ is going to be strictly increasing in $i$. However, in this specific case, $U_n = v_n - v$ independent of the distribution of types, in case $p_n > 0$ Therefore, it is clear that the distribution that will maximize the ex-ante expected utility of player 2 is the one that has all the weight in $v_n$, or $p_n = 1$.

**A.1.3. One informed player, not strong or weak**

The setting is almost the same as before, the only thing that changes is that, $v$ is between $v_1$ and $v_n$. Therefore, the uninformed player is not weak or strong. We will see that the distribution that maximize the ex-ante expected payoff of the informed player is going to be some sort of mix between the two previous cases we have seen so far.

**A.1.3.1. Equilibrium.** In this setting we could have two different situations, depending on which player has an atom at zero.

**A.1.3.2. Informed with the atom.** When the informed player has an atom, the equilibrium is going to be exactly as in section 1. However, to be in this case we need a couple of conditions to hold.
First condition assure us that there is a type for the informed player that has an atom at zero, and the second condition that such type is active. When the second equation is binding, it means that $i^*$ is not active, which means that our $i^*$ is actually $i^* + 1$ or something greater.

A.1.3.3. Uninformed with the atom. In this case, the uninformed player gets a payoff of zero, and the equilibrium is just like the one in section 2. However, in order for this to happen, we need that

\begin{equation}
1 - \sum_{i=i^*+2}^{n} \frac{p_i v_i}{v_i} > \frac{p_{i^*+1} v_{i^*+1}}{v_{i^*+1}}
\end{equation}

If it hold with equality, we are still in the same case (the uninformed player getting zero expected payoff)

A.1.3.4. Optimal probability. To find the optimal distribution we will first study the two previous cases separately, and then solve the original problem.

A.1.3.5. Uninformed player with an atom at zero. As we know, the expected payoff of the informed player when her type is $v_i$ is increasing on $i$. Because the length of the union of best response sets is always $v$ on this case, any distribution forces an equilibrium with an atom for the uninformed player will leave the informed one with an
ex ante payoff lower than $v_n - v$. But that is what she would get with a distribution that has all the mass on $v_n$. Also, because $v_n > v$, that distribution satisfies the condition to be in this set

**A.1.3.6. Informed player with an atom at zero.** This case is considerable more difficult. What we will do is basically follow what we did on section 1. First, as before, we only need to focus on distributions that have the first type active (with an atom but doing an effort greater than zero with positive probability for the lowest type). This is exactly as Lemma 34. It is only slightly different because in this case we have to be sure that the conditions for being in this case still hold.

Then, showing that there can not be more than 3 different types with positive probabilities is done in the same way. The idea is that the problem that we had in section 1 is the same that we have here, but with one less restriction. Therefore, if we look at our problem now, the solution is going to be either the same as the one without the restriction, or such that the restriction holds with equality. Because having the condition with equality implies that the first type is not active, the conclusion is the same.

Using a similar argument, when we solve for the case with only three types, we are doing the same as in section one but with more conditions. The idea is that the solution is going to be either the same as in section 1, or a solution with the new conditions holding with equality. In any case that means that there are only 2 types with positive mass.

What follows is solving the case for two types, and the use all we know about the problem to get the solution.
Two types. Suppose we have only two types \((v_1 < v < v_2)\) and we are conditioning on the lowest type having utility zero. Then, given the unique equilibrium, we have that the ex ante expected utility is:

\[
p_2 \left( 1 - \frac{p_2 v}{v_2} \right) (v_2 - v_1)
\]

We want to maximize that and compare it with the case in which the uninformed player has zero payoff. The condition for that to hold is

\[
1 - \frac{p_2 v}{v_2} \leq \frac{p_1 v}{v_1}
\]

Note that when this hold with equality, then we are in the case in which the length of the union of best responses is equal to \(v\). If that is the case, then having all the mass on \(v_2\) is better. Also, if the condition holds, then \(p_2\) must be less than one. Thus we need to maximize replacing \(p_1 = 1 - p_2\) and subject to \(p_2 \geq 0\). The solution is

\[
p_2^* = \min \left\{ \frac{v_2}{2v}, \frac{v - v_1}{v_2 - v_1}, \frac{v_2}{v} \right\}
\]

The second term is always less than 1. Now, if we are in the first case, then we have to compare the final utility with what we would get if all the mass is in \(v_2\). If we are in the second case, it means that the solution of this sub problem is to have a total length of \(v\), therefore we know that having all the mass in \(v_2\) dominates. Finally, we are in the first case whenever \(\frac{v_1 + v_2}{2} < v\), and on the second case otherwise. With this, the solution of the two types case is going to be
$$p_2^* = \begin{cases} \frac{v_2}{2v} & \text{if } v_n - v < \frac{1}{4} \frac{v_2}{v} (v_2 - v_1) \& \frac{v_1 + v_2}{2} < v \\ 1 & \text{o.w.} \end{cases}$$

And the expected utility is:

$$U^*(v_2, v_1) = \begin{cases} \max \left\{ v_2 - v, \frac{1}{4} \frac{v_2}{v} (v_2 - v_1) \right\} & \text{if } \frac{v_1 + v_2}{2} < v \\ v_2 - v & \text{o.w.} \end{cases}$$

**N types.** If we have $N$ types, call $i^*$ the highest type that is less than $v$. We know that the solution will involve at most two types with positive mass, but which one? First, if the pair is above $v$, then it does not matter, it is like using all the mass in $v_n$. If the pair is less than $v$, then the best is to use $i^*$ and 1, getting $\frac{1}{4} \frac{v_2}{v} (v_{i^*} - v_1)$, and if it involves one greater and one less, then the optimal utility will be given by $U^*(.,.)$. So, which is the best?

We have three cases.

The first is when $\frac{v_n + v_1}{2} \leq v$ In such case, the optimal pair is $n$ and 1. This because $U^*(v_n, v_1) \geq v_n - v, \frac{1}{4} \frac{v_2}{v} (v_n - v_1)$, therefore better than anything another pair could get.

The second case is when for all $i > i^*, \frac{v_i + v_1}{2} > v$. In such case, the best we can do by using pairs that have something bigger and lower than $v$ is to put all the weight in $v_n$. But we do not know how that compares to what we would get if we are using a pair lower than $v$, therefore the utility will be $U^* = \max \left\{ v_n - v, \frac{1}{4} \frac{v_2}{v} (v_{i^*} - v_1) \right\}$. But, $v_n - v > v - v_1$, so $v_n - v > v_{i^*} - v_1$, and because $\frac{1}{4} \frac{v_2}{v} < 1$, in this case the solution is $U^* = v_n - v$, all the mass on $v_n$. 


Finally, the third case is when there is some \( l > i^* \) that has \( \frac{v_l + v_1}{2} < v \) but that does not hold for \( n \). In such case, we know that because \( \frac{v_l + v_1}{2} < v \leq \frac{v_n + v_1}{2} \), \( v_n - v \geq 0.5(v_n - v_1) \), and \( v_n - v_1 > v_l - v_1 \). Also, \( v_l < 2v \), if it wasn’t the case, then \( \frac{v_l + e}{2} \geq v \) for all \( e \), which can not be. This implies that \( \frac{1}{4}v(v_l - v_1) < v_n - v \) and that \( U^*(v_l, v_1) < v_n - v \). Also, it is direct to see that \( U^*(v_l, v_1) \) is greater than what you would get using any pair that is less than \( v \).

To summarize, we will end up with two cases:

\[
U^* = \begin{cases} 
\max \left\{ v_n - v, \frac{v_n}{4} (v_n - v_1) \right\} & \text{if } \frac{v_n + v_1}{2} < v \\
v_n - v & \text{o.w.}
\end{cases}
\]

This is, it will either put all the mass on \( v_n \) or distribute it between \( v_n \) and \( v_1 \). Finally, reducing the cases, what we have is that:

\[
U^* = \begin{cases} 
\frac{v_n}{4} (v_n - v_1) & \text{if } \frac{v_n + v_1}{2} < v \text{ and } (2v - v_n)^2 > v_1 v_n \\
v_n - v & \text{o.w.}
\end{cases}
\]

\[
p_3^* = \begin{cases} 
\frac{v_n}{2v} & \text{if } \frac{v_n + v_1}{2} < v \text{ and } (2v - v_n)^2 > v_1 v_n \\
1 & \text{o.w.}
\end{cases}
\]

And \( p_1^* = 1 - p_n^* \).
A.2. Appendix II

Lemma 36. Conditional on the lowest type “almost” always participating, there is no interior solution. In other words, either \( p_j = 0 \) for some \( j \), or the participation condition holds with equality, which means that the first type does not participate.

Proof. The proof of this lemma is as follows, first, assume that there is an optimal interior distribution that puts positive probability on all types, and for which the lowest type is participating. Then,

\[
\mathbb{E}(U) = \sum_{i>1} U_ip_i
\]

Because we are in the optimal case, and it is interior (in the sense that all probabilities are greater than zero, and type 1 is participating), there can not be any small change in probabilities \( p_j, j > 1 \) that can lead to an increase in the expected payoff. In other words, the derivative of the expected payoff must be zero for all \( p_j, j > 1 \).

\[
\frac{\partial \sum_{i>1} U_ip_i}{\partial p_j} = U_j + \sum_{i>1} \frac{\partial U_i}{\partial p_j}p_i
\]

\[
= 0, \quad \forall j > 0
\]

With
\[
\frac{\partial U_i}{\partial p_j} = \frac{\partial}{\partial p_j} \left( 1 - \sum_{k>i} \frac{p_k v}{v_k} \right) v_i - \frac{\partial d_i}{\partial p_j}
\]

\[
= \begin{cases} 
-\frac{v_i}{v_j} v_j + v_1 \frac{v}{v_j} & j > i \\
 v_1 \frac{v}{v_j} - v & j \leq i
\end{cases}
\]

and

\[
U_j = v_j \left( 1 - \sum_{i>j} \frac{p_i v}{v_i} \right) - v_1 \left( 1 - \sum_{i>1} \frac{p_i v}{v_i} \right) - \sum_{j \geq i>1} p_i v
\]

therefore:

\[
\frac{\partial}{\partial p_j} \sum_{i>1} U_i p_i = v_j \left( 1 - \sum_{i>j} \frac{p_i v}{v_i} \right) - v_1 \left( 1 - \sum_{i>1} \frac{p_i v}{v_i} \right) - \sum_{j \geq i>1} p_i v + (1 - p_1) v \frac{v}{v_j} - \sum_{1<i<j} p_i v i \frac{v}{v_j}
\]

\[
= v_j \left( 1 - \sum_{i>j} \frac{p_i v}{v_i} \right) - v_1 \left( 1 - \sum_{i>1} \frac{p_i v}{v_i} \right) - \sum_{i>1} p_i v + (1 - p_1) v \frac{v}{v_j} - \sum_{1<i<j} p_i v i \frac{v}{v_j}
\]

\[
= v_j \left( 1 - \sum_{i>j} \frac{p_i v}{v_i} \right) + (1 - p_1) v \frac{v}{v_j} - \sum_{1<i<j} p_i v i \frac{v}{v_j} - v_1 \left( 1 - \sum_{i>1} \frac{p_i v}{v_i} \right) - \sum_{i>1} p_i v
\]

As we said, if all the probabilities are greater than zero, then it must be the case that all the derivatives are equal to zero.
Call

\[ C(p) = v_1 \left(1 - \sum_{i > 1} \frac{p_i v}{v_i}\right) + \sum_{i > 1} p_i v \]

then

\[ v_{j+1} \frac{\partial \mathbb{E}[U]}{\partial p_{j+1}} - v_j \frac{\partial \mathbb{E}[U]}{\partial p_j} = (v_{j+1}^2 - v_j^2) \left(1 - \sum_{i > j} \frac{p_i v}{v_i}\right) - (v_{j+1} - v_j) C(p) \]

\[ = (v_{j+1} - v_j) \left[(v_{j+1} + v_j) \left(1 - \sum_{i > j} \frac{p_i v}{v_i}\right) - C(p)\right] \]

In a solution with positive probabilities, this must be equal to zero, which implies that

\[ (v_{j+1} + v_j) \left(1 - \sum_{i > j} \frac{p_i v}{v_i}\right) = C(p) \]

However, the LHS is strictly increasing on \( j \), but the RHS is constant on \( j \), which is a contradiction. We can apply this for any case in which we have more than 4 possible valuations. Therefore, any solution will have at most 3 probabilities different than 0. \( \square \)

**Lemma 37.** Suppose the informed player can have payoff \( v_1, v_2 \) or \( v_3 \). Then the optimal distribution is \( p_3 = \frac{v_3}{2v}, p_2 = 0 \), which yields expected payoff \( \frac{v_3}{4v}(v_3 - v_1) \)

**Proof.** Suppose the informed player can have payoff \( v_1, v_2 \) or \( v_3 \). Conditional on type \( v_1 \) participating, the expected payoff is:

\[ E(U) = p_2(v_2 - v_1) \left(1 - \frac{p_2 v}{v_2} - \frac{p_3 v}{v_3}\right) + p_3 \left(v_3 - p_3 v - p_2 v - v_1 \left(1 - \frac{p_2 v}{v_2} - \frac{p_3 v}{v_3}\right)\right) \]
Let’s study this function. If we forget about other considerations (when this expression is valid), we can compute the optimal $p_2$ given $p_3$.

\[
\frac{\partial \mathbb{E}(U)}{\partial p_2} = (v_2 - v_1) \left( 1 - 2 \frac{p_2 v}{v_2} - \frac{p_3 v}{v_3} \right) + p_3 v_1 \frac{v}{v_2} - p_3 v
\]

\[= (v_2 - v_1) \left( 1 - 2 \frac{p_2 v}{v_2} - \frac{p_3 v}{v_3} - p_3 \frac{v}{v_2} \right)\]

\[
\frac{\partial^2 \mathbb{E}(U)}{\partial p_2^2} = -2(v_2 - v_1) \frac{v}{v_2} < 0
\]

Therefore, given $p_3$, the optimal $p_2$ is given by:

\[p_2^*(p_3) = \max \left\{ 0, \frac{v_2}{2v} - p_3 \left( 1 + \frac{v_2}{v_3} \right) \right\}\]

When $p_3 \left( 1 + \frac{v_2}{v_3} \right) \geq \frac{v_2}{2v}$ the max is attained for $p_3 = \frac{v_2}{2v}$, and involves $p_2 = 0$. The expected payoff for that case is $\frac{v_2}{4v} (v_3 - v_1)$. \(^1\)

In case $p_3 \left( 1 + \frac{v_2}{v_3} \right) < \frac{v_2}{2v}$

\[\frac{d \mathbb{E}(U)}{dp_3} = \frac{\partial \mathbb{E}(U)}{\partial p_2} \frac{dp_2}{dp_3} + \frac{\partial \mathbb{E}(U)}{\partial p_3}\]

\[= \frac{\partial \mathbb{E}(U)}{\partial p_3}\]

\(^{1}\)Replacing, the objective function is strictly concave.
\[
\frac{d^2 \mathbb{E}(U)}{dp_3^2} = \frac{\partial^2 \mathbb{E}(U)}{\partial p_2^2} \frac{dp_2}{dp_3} + \frac{\partial \mathbb{E}(U)}{\partial p_3} \\
= 2(v_2 - v_1) \frac{v}{v_2} \left(1 + \frac{v_2}{v_3}\right) - 2v \left(1 - \frac{v_1}{v_3}\right) \\
= 2v \left(\frac{v_2}{v_3} - \frac{v_1}{v_2}\right)
\]

We have two cases. If \(\frac{v_2}{v_3} \geq \frac{v_2}{v_2}\) then the solution is in the bounds where \(p_3\) can live. This is, \(p_3 = 0\) or \(p_3 \left(1 + \frac{v_2}{v_3}\right) = \frac{v_2}{2v}\). The second case yields less payoff than \(p_3 = \frac{v_2}{2v}\) (showed before), and the first has payoff \(\frac{v_2}{4v} (v_2 - v_1)\) which is less than the optimal thing in the previous cases.

If the function is concave, then we have a quadratic form. We will see that the max of that function is in the bounds. The derivative is strictly decreasing. Then in the right bound, when \(p_3 \left(1 + \frac{v_2}{v_3}\right) = \frac{v_2}{2v}, p_2 = 0\), the derivative is

\[
\frac{\partial \mathbb{E}(U)}{\partial p_3} = -p_2 (v_2 - v_1) \frac{v}{v_3} + \left(v_3 - p_2 v - 2p_3 v - v_1 \left(1 - \frac{p_2 v}{v_2} - \frac{2p_3 v}{v_3}\right)\right)
\]

\[
= \left(v_3 - 2p_3 v - v_1 \left(1 - \frac{2p_3 v}{v_3}\right)\right)
\]

\[
= (v_3 - v_1) \left(1 - \frac{2p_3 v}{v_3}\right)
\]

\[
= (v_3 - v_1) \left(1 - \frac{v_2}{v_3 + v_2}\right)
\]

\[
> 0
\]

\(^2\)The derivative of \(p_2\) with respect to \(p_3\) is constant.
Therefore the max is when \( p_3 \left( 1 + \frac{v_3}{v_3} \right) = \frac{v_3}{2v} \), but the payoff when that happens is dominated by \( p_3 = \frac{v_3}{2v} \). \( \square \)
APPENDIX B

War of Attrition with Information Externalities

Lemma 3. In any equilibrium, in any sub-interval $SI$ of $[0, T)$, if a set of players $P \subset \mathcal{M}$ is mixing with positive density in $SI$, and the rest is not mixing in $SI$, then the hazard rate of those players is constant in that sub-interval $(\lambda_i)$ and satisfies:

$$\sum_{j \in P \setminus \{i\}} \lambda_j \Delta^j_i = 1, \forall i \in P$$

Suppose we are in an equilibrium, and in $(a, b) \subset SI$ only players in $P$ are mixing, but the rest is not. Assuming players are mixing with $F_p(.)$ (all players), the utility of quitting at time $t$ for player $i$ in $P$ is:

If no player from $P^c$ has left, it is equal to:

$$\sum_{j \in P \setminus \{i\}} \int_0^x \left( \Delta^j_i - x_j \right) \prod_{q \in P \setminus \{i, j\}} (1 - F_q(x_j)) dF_j - x \prod_{j \in P \setminus \{i\}} (1 - F_j(x))$$

If a player from $P^c$ has quit before $x$, because there is no chance of her quitting after $a$ but before $t$ it is:

$$\mathbb{E}(\min_{j \in P \setminus \{i\}} x_j | \min_{j \in P^c} x_j < \min \{a, x_p\}, \forall p \in P)$$

The second term does not depend on $x$, and the probability of players in $c$ quitting before $x$ only depend on $a$. Therefore, the utility of leaving at $t$ is of the form:
\[ M(a) \left( \sum_{j \in P \setminus \{i\}} \int_0^x (\Delta_j^i - x_j) \prod_{q \in P \setminus \{i, j\}} (1 - F_q(x_j)) \, dF_j - x \prod_{j \in P \setminus \{i\}} (1 - F_j(x)) \right) + N(a) \]

Because that must be constant for \( x \in SI \), it turns that

\[ \sum_{j \in P \setminus \{i\}} \int_0^x (\Delta_j^i - x_j) \prod_{q \in P \setminus \{i, j\}} (1 - F_q(x_j)) \, dF_j - x \prod_{j \in P \setminus \{i\}} (1 - F_j(x)) \]

is constant around \( x \). Also, because there are no atoms in \( SI \), we know that all probability distributions are continuous. Then, we are allowed to take derivatives, getting:

\[ \sum_{j \in P \setminus \{i\}} (\Delta_j^i - x) \prod_{q \in P \setminus \{i, j\}} (1 - F_q(x)) \prod_{j \in P \setminus \{i\}} (1 - F_j(x)) + \sum_{j \in P \setminus \{i\}} x \prod_{q \in P \setminus \{i, j\}} (1 - F_q(x)) f_j(x) = 0 \]

Cancelling terms, and dividing by \( \prod_{j \in P \setminus \{i\}} (1 - F_j(x)) \), we get:

\[ \sum_{j \in P \setminus \{i\}} \Delta_j^i \frac{f_j(x)}{1 - F_j(x)} - 1 = 0 \]

Equivalent to

\[ \sum_{j \in P \setminus \{i\}} \Delta_j^i \lambda_j(x) - 1 = 0 \]

Since we have \(|P|\) unknowns and \(|P|\) equations, if solvable, the solution will not depend on \( x \), therefore it must be that the equilibrium strategies have a constant hazard rate, that satisfy:
\[
\sum_{j \in P \setminus \{i\}} \Delta^j_i \lambda_j - 1 = 0, \quad \forall i \in P
\]

Note that a distribution with constant hazard rate has a cumulative of \(F(x) = 1 - C \exp[-\lambda x]\), with \(C\) a constant, and \(\lambda\) the hazard rate.

**Lemma 4.** In any sub-interval \(SI\) of \([0,T)\), if a set of players \(P\) is mixing with constant hazard rate in \(SI\), and the rest is not mixing in \(SI\), then for a player \(i \notin P\) the value of leaving increases (decreases) in \(SI\) if and only if

\[
\sum_{j \in P} \lambda_j \Delta^j_i < (>1, \quad \forall i \in P
\]

For a player \(i\) in \(P^c\), her utility in \(SI\) is of the form:

\[
M(a) \left( \sum_{j \in P} \int_0^x \left( \Delta^j_i - x_j \right) \prod_{q \in P \setminus \{j\}} \left( 1 - F_q(x_j) \right) dF_j - x \prod_{j \in P} \left( 1 - F_j(x) \right) \right) + N(a)
\]

With \(M(a)\) positive. Then, in order for that to be monotone at \(x\), we have to focus on

\[
\sum_{j \in P} \int_0^x \left( \Delta^j_i - x_j \right) \prod_{q \in P \setminus \{j\}} \left( 1 - F_q(x_j) \right) dF_j - x \prod_{j \in P} \left( 1 - F_j(x) \right)
\]

Since in equilibrium it must be that strategies are continuous, it implies the previous expression is almost everywhere differentiable. Taking derivative we get:
\[
\frac{\partial U}{\partial x} = \sum_{j \in P} \Delta_j^i \prod_{q \in P \setminus \{j\}} (1 - F_q(x))f_j(x) - \prod_{j \in P} (1 - F_j(x))
\]

And since we are before \(T\),

\[
\frac{\partial U}{\partial x} = \prod_{j \in P} (1 - F_j(x)) \left( \sum_{j \in P} \Delta_j^i \frac{f_j(x)}{1 - F_j(x)} - 1 \right)
\]

Then, depending on the sign of \(\sum_{j \in P \setminus \{i\}} \Delta_j^i \lambda_j(x) - 1\), we get that the utility of leaving increases or decreases in \(x\).
APPENDIX C

Dynamic Tullock Contest

Proof of Lemma 20

Proof. (1) If $x > \theta$, then $\pi(x|y, \theta) < 0$. Notice that $x = 0$ implies that $\pi(x|y, \theta) \geq 0$, and therefore $R(s|\theta) \leq \theta$. The differentiability of the best response comes from the differentiability of $g(\cdot)$ and the implicit function theorem.

(2) It is a consequence of the monotone comparative static theorem.

(3) Deriving implicitly and using that:

$$\frac{\partial^2 \pi}{\partial x \partial y} = \frac{\theta g'(x)g'(y)}{(g(x) + g(y))^2} (g(x) - g(y)) \quad \text{and} \quad \frac{\partial^2 \pi}{\partial x^2} = \frac{\theta g(y)}{(g(x) + g(y))^2} (g''(x) - 2[g'(x)]^2)$$

we obtain

$$\frac{dR(s|\theta)}{ds} = \frac{g'(R(s|\theta))g'(s)(g(R(s|\theta)) - g(s))}{g(s) ([g'(R(s|\theta))]^2 - g''(R(s|\theta)))}.$$ 

(4) Otherwise, if $R(s|\theta) < s$ for all $s \in (0, \delta)$, then $R(s|\theta)$ would be decreasing in $(0, \delta)$, which would imply by continuity that $R(s|\theta) = 0$ for all $s$. This is a contradiction, since the marginal benefit of invest around zero is larger than the marginal cost, for $s$ small enough, as long as $g'(0) > \frac{g(s)}{\theta}$. So as long as $\theta > 0$, and $s$ is small, $R(s|\theta) > 0$. 
(5) Notice that for \( s > \theta \) we have \( R(s|\theta) < s \). The continuity of \( R(s|\theta) \) and its monotonicity implies that there exists a unique \( x^*(\theta) \) such that \( R(x^*(\theta)|\theta) = x^*(\theta) \).

\[ \square \]

**Proof of Lemma 23**

**Proof.** Clearly \( \pi_i(W_i, W_j) \) is continuous in the interior of the regions I, II.1, II.2, and III.

Consider \( W_2 < x^*(\theta_2) \). In the boundary of regions I and II.1 we have:

\[
\lim_{W_1 \to x^*(\theta_2)^-} \pi_1(W_1, W_2) = \frac{V_2}{2} = \lim_{W_1 \to x^*(\theta_2)^+} \pi_1(W_1, W_2)
\]

\[
\lim_{W_1 \to x^*(\theta_2)^-} \pi_2(W_1, W_2) = \frac{V_2}{2} - c_2x^*(\theta) = \lim_{W_1 \to x^*(\theta_2)^+} \pi_2(W_1, W_2)
\]

where we used that \( R(x^*(\theta_2)|\theta_2) = x^*(\theta_2) \). Consider a point \((\tilde{W}_1, R(\tilde{W}_1|\theta_2))\) in the frontier of regions II.1 and III. We have:

\[
\lim_{W_1 \to \tilde{W}_1^-} \pi_1(W_1, W_2) = \frac{V_2}{2} = \lim_{W_1 \to \tilde{W}_1^+} \pi_1(W_1, W_2)
\]

\[
\lim_{W_1 \to \tilde{W}_1^-} \pi_2(W_1, W_2) = V_2p(R(\tilde{W}_1|\theta_2), \tilde{W}_1) = \lim_{W_1 \to \tilde{W}_1^+} \pi_1(W_1, W_2)
\]

Analogously, using the symmetry of the problem, continuity is shown in the remaining regions. \( \square \)
Proof of Proposition 24

**Proof.** When $W_2 \leq x^*(\theta_2)$ agent 1 can only go to regions I, II.1, and III by changing its effort $W_1$. When $W_2 > x^*(\theta_2)$, agent 1 can only go to regions II.2, and III. We divide the proof in two lemmas:

**Lemma 39.** When $W_2 \leq x^*(\theta_2)$ and the condition $c_2(V_1 + V_2) < c_1V_2$ holds, agent 1’s the best response in the first contest is

$$BR_1(W_2) = R(W_2|\theta_I), \text{ where } \theta_I = \frac{V_1}{c_1 - c_2}.$$ 

**Proof.** First, define $\theta_I = \frac{V_1}{c_1 - c_2}$ and notice that $c_2(V_1 + V_2) < c_1V_2$ guarantees that $\theta_I < \theta_2$.

1) Notice that the payoff in region I corresponds to the payoff of a standard Tullock contest with parameter $\theta_I = \frac{V_1}{c_1 - c_2}$ plus a constant. This payoff function is increasing and concave, and its maximum is at $R(W_2|\theta_I)$. As a function of $W_1$, the payoff is increasing until $W_1 = R(W_2|\theta_I)$ and decreasing after that point.

2) Adding and subtracting $c_2W_1$ we can write the payoff in the region II.1 as:

$$V_1p(W_1, W_2) - (c_1 - c_2)W_1 + V_2p(W_1, R(W_1|\theta_2)) - c_2W_1.$$ 

This is the payoff of a standard Tullock contest with parameter $\theta_I$ plus the Stackelberg payoff

$$S(W_1) = V_2p(W_1, R(W_1|\theta_2)) - c_2W_1.$$
Since in region II.1 $W_1 > x^*(\theta_2) > x^*(\theta_I)$ we know that $V_{1p}(W_1, W_2) - (c_1 - c_2)W_1$ is decreasing in $W_1$. We now show that also $S(W_1)$ is decreasing in region II.1.

By the definition of $R(s|\theta_2)$ we have

$$\frac{V_2g(s)g'(R(s|\theta_2))}{(g(R(s|\theta_2)) + g(s))^2} = c_2 \Rightarrow \frac{V_2g(s)}{g(R(s|\theta_2)) + g(s)} = c_2 \frac{g(R(s|\theta_2)) + g(s)}{g'(R(s|\theta_2))}$$

This implies that

$$S(W_1) = c_2 \left[ \frac{g(R(W_1)) + g(W_1)}{g'(R(W_1|\theta_2))} - W_1 \right].$$

Taking derivative we get

$$\frac{S'(W_1)}{c_2} = R'(W_1|\theta_2) + \frac{g'(W_1)}{g'(R(W_1|\theta_2))} = \frac{[g(R(W_1|\theta_2)) + g(W_1)]g''(R(W_1|\theta_2))R'(W_1|\theta)}{[g'(R(W_1|\theta_2))]^2} - 1.$$ 

Notice that $R'(W_1|\theta_2) < 0$ because in region II.1 $x^*(\theta) < W_1$. This also implies (by property of $R$) that $W_1 > R(W_1|\theta_2)$. Since $g$ is concave, $g'$ is decreasing, which implies that $g'(W_1) \leq g'(R(W_1|\theta_2))$. Combining all these facts we get that $S'(W_1) < 0$.

Hence, the payoff in region II.1 is decreasing in $W_1$.

3) The payoff of region III corresponds to the payoff of a standard Tullock contest with parameter $\theta_{III} = \frac{V_1 + V_2}{c_1}$. Again, $c_2(V_1 + V_2) < c_1V_2$ implies that $\theta_{III} < \theta_2$ and therefore for $W_1 > x^*(\theta)$, which is the case in region III, the payoff is decreasing.

Thus, by continuity, the payoff is always decreasing for $W_1 > x^*(\theta)$. This implies that:
arg max \( V_i \cdot p(W_i, W_j) - c_1 \cdot W_i + \pi_i(W_i, W_j) \) = \arg \max_{0 \leq W_i \leq x^*(\theta)} V_i \cdot p(W_i, W_j) - c_1 \cdot W_i + \pi_i(W_i, W_j) = R(W_j|\theta_I)

Thus, the best response of in the first stage for \( W_2 \leq x^*(\theta) \) is given by \( R(W_2|\theta_I) \) \( \square \)

**Lemma 40.** When \( W_2 > x^*(\theta_2) \) and the condition \( c_2(V_1 + V_2) < c_1 V_2 \) holds, agent 1’s the best response in the first contest is

\[
BR_1(W_2) = R(W_2|\theta_I), \text{ where } \theta_I = \frac{V_1}{c_1 - c_2}.
\]

**Proof.** Consider \( W_2 > x^*(\theta_2) \) and notice that for any \( W_1 \leq R(W_2|\theta_2) \), the payoff corresponds to a Tullock with parameter \( \theta_I \) plus a constant that depends on \( W_2 \). In this region, the maximum payoff achieved at either \( W_1 = R(W_2|\theta_I) \) if \( R(W_2|\theta_I) < R(W_2|\theta_2) \) or at \( W_1 = R(W_2|\theta_2) \) otherwise. However, \( \theta_I < \theta_2 \) implies that \( R(W_2|\theta_I) < R(W_2|\theta_2) \) for all \( W_2 \). Then, the maximum is achieved at \( W_1 = R(W_2|\theta_I) \)

When \( W_1 > R(W_2|\theta_2) \), in region III, and the payoff is achieved at \( W_1 = R(W_2|\theta_2) \), because \( R(W_2|\theta_{III}) < R(W_2|\theta_2) \) and the payoff in region III is decreasing for \( W_1 > R(W_2|\theta_{III}) \).

By continuity, \( BR(W_2) = R(W_2|\theta_I) \). \( \square \)
Proof of Proposition 25

**Proof.** Notice that $c_2(V_1 + V_2) \geq c_1 V_2$ implies $\theta_2 \leq \theta_{III}$, and $\theta_I \geq \theta_2$. For a fixed $W_2 \geq x^*(\theta_2)$, agent 1’s payoff can only be in regions II.2 or III. In II.2, agent 1 payoff’s equals:

$$V_1 p(W_1, W_2) - (c_1 - c_2)W_1 + V_2 p(R(W_2|\theta_2), W_2) - c_2 R(W_2|\theta_2).$$

Notice that the payoff is increasing in $W_1$ for any fixed $W_2$. If $c_1 \leq c_2$, then it is clearly increasing. If $c_1 > c_2$, the payoff in an unrestricted domain is maximized at $W_1^* = R(W_2|\theta_I) > R(W_2|\theta_2)$. By properties of the standard Tullock contests, this payoff is increasing for $W_1 < R(W_2|\theta_2)$. Also, the payoff of region $III$ in an unrestricted domain is maximized at $W_1^* = R(W_2|\theta_{III}) > R(W_2|\theta_2)$. By continuity of the payoff, the best response given $W_2 \geq x^*(\theta)$ is reached $BR(W_2) = R(W_2|\theta_{III})$. ∎

Proof of Proposition 26

**Proof.** We have two cases:

a) $c_2(V_1 + V_2) < c_1 V_2$: From Proposition 24, the best responses in the first stage are equal to $R(\cdot|\theta_I)$. From Lemma 20 the equilibrium must be unique and equal to $x^*(\theta_I)$.

b) $c_2(V_1 + V_2) < c_1 V_2$: In this case, $\theta_{III} \geq \theta_2$, therefore $x^*(\theta_{III}) \geq x^*(\theta_2)$, and the best response to $x^*(\theta_{III})$ is $R(x^*(\theta_{III})|\theta_{III}) = x^*(\theta_{III})$ from Proposition 25. Thus, we know that $x^*(\theta_{III})$ is an equilibrium, by properties of the best response in Lemma 20.
To show that this is the unique equilibrium, suppose by contradiction that there is another equilibrium. We know that for $W_i > x^*(\theta_2)$ and $W_j > x^*(\theta_2)$ there is a unique equilibrium. Then, in any other equilibrium we must have $W_i \leq x^*(\theta_2)$ or $W_j \leq x^*(\theta_2)$. But $W_i \leq x^*(\theta_2)$ or $W_j > x^*(\theta_2)$ cannot be an equilibrium, since $BR(W_j) = R(W_j|\theta_{III}) > x^*(\theta_2) \geq W_i$. Similarly, $W_j \leq x^*(\theta_2)$ or $W_i > x^*(\theta_2)$ cannot be an equilibrium. Therefore, it must be the case that $W_i \leq x^*(\theta_2)$ and $W_j \leq x^*(\theta_2)$. The only way for this to be an equilibrium is that $W_i = R(W_j|\theta_I)$ and $W_j = R(W_i|\theta_I)$, otherwise agents would have a local deviation. The only solution for that system is $W_i = W_j = x^*(\theta_I)$. But $R(W_j|\theta_I) > x^*(\theta_2)$ and $R(W_i|\theta_I) > x^*(\theta_2)$, hence there is no other equilibrium.

Finally, the efforts in the second contest come from replacing the efforts of the first contest as head starts into proposition 22.

\[\Box\]

Proof of Lemma 27

**Proof.** From the first order condition we get

\[
\frac{\theta g(y)g'(x)}{(g(x) + g(y))^2} = 1.
\]

When $g(\cdot)$ is increasing and concave, there is a symmetric equilibrium $x^*(\theta)$ characterized by

\[
\frac{g(x^*(\theta))}{g'(x^*(\theta))} = \frac{\theta}{4}.
\]
We define $F(x) = \frac{g(x)}{g'(x)}$. Taking implicit derivative to the condition above, we get

$$F'(x^*(\theta)) \frac{dx^*(\theta)}{d\theta} = \frac{1}{4}.$$ 

Taking derivative of this condition we get

$$F''(x^*(\theta)) \left( \frac{dx^*(\theta)}{d\theta} \right)^2 + F'(x^*(\theta)) \frac{d^2x^*(\theta)}{d\theta^2} = 0.$$ 

It is easy to show that $F'(x) > 0$ when $g$ is increasing and concave. Therefore,

$$\frac{d^2x^*(\theta)}{d\theta^2} = -F''(x^*(\theta)) \left( \frac{dx^*(\theta)}{d\theta} \right)^2 \frac{F'(x^*(\theta))}{F'(x^*(\theta))} > 0.$$ 

Notice that $\text{sign} \frac{d^2x^*(\theta)}{d\theta^2} = -\text{sign}F''(x^*(\theta))$. Thus, when $F'' > 0$ if and only if $x^*(\theta)$ is concave.

Consider $h(c) = c \cdot x^* \left( \frac{V}{c} \right)$, and let $\theta = \frac{V}{c}$. Then,

$$h'(c) = x^* \left( \frac{V}{c} \right) + c \frac{dx^*(\theta)}{d\theta} \frac{d\theta}{dc} = x^*(\theta) - \theta \frac{dx^*(\theta)}{d\theta}.$$ 

By concavity, and using that $x^*(0) = 0$ we have that $h'(c) > 0$ iff $F'' > 0$. □

**Proof of Corollary 33**

**Proof.** Consider first the case $\delta < \delta^*$. From Proposition 31,

$$x^*_i + W^*_i = x^*(\theta_2) + (1 - \delta)W^*_i = \frac{\alpha V_2}{4c_2} + (1 - \delta)\frac{\alpha V_1}{4(c_1 - \delta c_2)}.$$
It is easy to show that \((1 - \delta)\frac{\alpha V_1}{4(c_1 - \delta c_2)}\) is strictly increasing (decreasing) if \(c_2 > c_1\) (\(c_2 < c_1\)).

Notice that for \(\delta = 0\), \(x_i^* + W_i^* = \frac{\alpha}{4}\left(\frac{V_2}{c_2} + \frac{V_3}{c_1}\right)\), and for \(\delta \geq \delta^*\), the aggregate effort is constant and equal to \(\frac{\alpha}{4}\left(\frac{V_1 + V_2}{c_1}\right)\). \(\square\)