Incentives and Selection in Promotion Contests:
Is It Possible to Kill Two Birds With One Stone?*

Rudi Stracke[a], Wolfgang Höchtl[b], Rudolf Kerschbamer[c], and Uwe Sunde[a]

[a]Seminar for Population Economics, University of Munich (LMU)

[b]Austrian National Bank, Vienna

[c]Department of Economics, University of Innsbruck

February 4, 2014

Abstract

This paper investigates whether a designer can improve both the incentive provision and the selection performance of a promotion contest by making the competition more (or less) dynamic. A comparison of static (one-stage) and dynamic (two-stage) contests reveals that this is not the case. A structural change that improves the performance in one dimension leads to a deterioration in the other dimension. This suggests that modifications of the contest structure are an alternative to strategic handicaps. A key advantage of structural handicaps over participant-specific ones is that the implementation of the former does not require prior identification of worker types.

JEL-Classification: M52, J33

Keywords: Promotion Contests, Heterogeneity, Incentives, Selection, Handicapping

---

*Corresponding author: Rudi Stracke, Schackstr. 4, 80539 Munich, Germany; Tel.:+49-2180-1285; Email: Rudi.Stracke@econ.lmu.de.

1   Introduction

Most employment relationships are characterized by competition among employees for promotion to a better paid, more attractive position. While such promotion contests are often just a by-product of a given hierarchical structure, they are sometimes actively used as an instrument in the practice of human resource management (HRM) in professional occupations. For example, 'up-or-out' promotion policies are the norm in law or consulting firms, where vacant manager or partner positions are almost exclusively filled with insiders. Contests are also used to fill top management positions. A prominent example is Jack Welch, the former CEO of General Electric (GE), who designed the competition for his succession about six years before he actually left. Several candidates from inside GE knew that they were competing against each other for this job, and that they would either become the next CEO, or would have to leave the firm.[1] 'Up-or-out' promotion policies are also common in the competition between scientists: In each year, only the (relatively) best performing PhDs become assistant professors, and only the best among the assistant professors receive a tenured position, while less successful staff members have to leave.

In all these applications, promotion contests are used as a means to achieve two goals: First, the contest provides employees with incentives, since the prospect of moving up the ladder to higher levels within the same institution is a strong motivator for employees to exert effort in their current job.[2] Second, the contest helps to select the most able and productive candidate(s) for promotion.[3] The fact that both incentive provision and selection performance matter for corporations raises the question how these two goals are related to each other. Can contests be used as a device to maximize the incentives for effort provision while at the same time minimizing the probability that the "wrong" contestant wins? Or, in other words, can promotion contests be designed in such a way that they kill two birds with one stone?

This paper investigates how modifications of the contest structure affect the performance in these two dimensions. In particular, we analyze whether a designer can improve both the incentive provision and the selection performance of a promotion contest by making the competition more (or less) dynamic. A comparison of incentive provision and selection performance in a *static* (one-shot) and a *dynamic* (two-stage) promotion contest suggests that the two goals

---

[1] See Welch (2001) for details. Konrad (2009) provides an extensive discussion of this example.

[2] This aspect is particularly important in employment relationships plagued by moral hazard problems (Lazear and Rosen 1981). Alternatively, the contest may serve as a commitment device for the principal (Malcomson 1984). See Prendergast (1999) for a survey.

[3] That promotion contests can help employers to select high ability employees has previously been acknowledged in the literature – by Rosen (1986) and Waldman (1990), for example. According to Rosen, "the inherent logic [of promotion contests] is to determine the best contestants and to promote survival of the fittest" (p.701). Somewhat surprisingly, however, Rosen's seminal paper is all about optimal incentive provision across different stages of the contest.

are incompatible. While the dynamic format performs better in terms of aggregate equilibrium efforts, the static contest dominates with respect to selection. An additional hierarchy level in promotion contests is therefore beneficial for incentive provision, but detrimental for the accuracy in selection.

More generally, our results show that the trade-off between incentive provision and accuracy of selection established in previous work cannot be solved by structural variations: Modifications that improve the performance in one dimension deteriorate the performance in the other dimension. Thus, the variations in the contest structure considered in this paper have similar effects as strategic handicaps.[4] Intuitively, contest structures which amplify the degree of heterogeneity between strong and weak workers perform well in terms of selection, as heterogeneity discourages weak workers relatively more than it induces strong workers to slack off. Pooling of workers in a simultaneous interaction tends to increase the effective degree of heterogeneity, implying that the one-stage contest delivers the best selection performance. In contrast, the more the structure of the competition moderates heterogeneity between workers, the better is its performance in the incentive dimension, since heterogeneity decreases the incentives for effort provision for both strong and weak workers in absolute terms. The separation of employees into pair-wise interactions, for example, reduces the effective degree of heterogeneity, which then enhances incentives for effort provision.

From a practitioner's perspective, our results have two important implications: First, structural variations cannot ensure that a promotion contest serves both goals equally well. This suggests that a promotion contest should either be used to maximize incentives for effort provision in a preselected sample of similarly talented employees, or for the sorting of types if the talent of employees is not observable and alternative incentive schemes are available. Second, the finding that structural variations work similar to strategic handicaps is useful when the ability of contestants is unobservable and the designer cares about one of the two objectives only. While participant-specific handicapping schemes require information about the ability of each contestant, or at least some (possibly imperfect) signal of participants' ability, such information is not necessary to employ structural handicaps.

The remainder of this paper is structured as follows. We first discuss the related literature in section 2. Section 3 introduces the formal model. Section 4 compares the equilibrium measures for incentive provision and selection performance of different contest formats. Section 5 discusses the intuition for and the limitations of our results, and section 6 concludes.

---

[4]Although the exposition focuses on personnel policies, and in particular on the promotion contest application, our findings are equally important for rent-seeking contests. Note, however, that the interpretation may be different in that case if effort inputs are assumed to be wasteful.

2 Related Literature

Our paper contributes to three strands of the literature on contests. It is most closely related to previous work on the relation between incentives and selection in promotion contests. Baker, Jensen, and Murphy (1988) were the first to investigate whether promotion contests ensure that employees end up in those jobs for which they are best suited, assuming that skill requirements differ qualitatively across hierarchy levels.[5] Our approach is different from theirs in that we consider situations where talent requirements are qualitatively similar across hierarchy levels, but where the ability to perform the tasks differs across workers. In other words, while the heterogeneity in Baker et al. is 'horizontal', the heterogeneity considered here is 'vertical'. Obvious examples for settings where vertical heterogeneity matters are law and consultancy firms. Similarly, top managers as well as CEOs perform very similar tasks, and both assistant and tenured professors teach and conduct research. Regarding the assumption that the ability to perform the same task differs across workers, our paper is closely related to work by Tsoulouhas, Knoeber, and Agrawal (2007), who study a one-stage promotion contest where insiders and outsiders compete for a CEO position. Assuming that both the quality of the promoted agent and the provision of incentives matter for the designer, they find that the two goals are conflicting if the ability of outsiders is higher than the ability of insiders. While this result has a similar flavor as the one established in our paper, the focus of their study is on optimal handicapping in a setting where selection involves both insiders and outsiders, but only effort provision by insiders is beneficial for the organization. In contrast, we consider how structural variations of within-firm competition affect incentive provision and selection performance. Our results show that the trade-off between these two goals established by Tsoulouhas, Knoeber, and Agrawal (2007) for the competition between insiders and outsiders is also present in purely internal promotion contests if a company employs workers of different types. In particular, we find that changes in the contest structure by making the within-firm competition more (or less) dynamic through additional (or less) hierarchy levels cannot improve the performance in both dimensions simultaneously.

The present paper are contributes to the literature on optimal handicapping.[6] While it is well known that strategic handicaps in a static interaction between two heterogeneous agents may either improve incentives or selection, depending on whether the designer handicaps the 'favorite' or the 'underdog', a designer with a given goal might find this insight of limited

---

[5]In their words, '...talents for the next level in the hierarchy are not perfectly correlated with talents to be the best performer in the current job' (p. 602). The best salesman, for example, can be a bad manager, which leads to the so-called Peter Principle. See also Prendergast (1993) and Bernhardt (1995), who also consider the matching performance of promotion contests.

[6]The seminal paper on handicapping in heterogeneous tournaments is Lazear and Rosen (1981), a more recent contribution that also cites some of the earlier work in this field is Gürtler and Kräkel (2010).

help – participant-specific handicapping requires prior identification of workers' productivities, which is often infeasible in reality. The structural variations studied in the present paper have similar effects as strategic handicaps but the important advantage that knowledge of individual productivities is not necessary. In that regard, the results of our paper are more closely related to recent work by Ridlon and Shin (2013). They analyze optimal handicapping strategies in repeated contests among heterogeneous workers under the assumption that the designer has imperfect information about the ability of her employees. Both the contest setting under investigation and the intuition for the results obtained are different from those in the present paper. Ridlon and Shin consider a repeated pairwise interaction and allow the principal to handicap conditional on the outcome of the first round contest. Thus, the strategic effect of a handicap is complicated by the fact that workers adjust their effort choice in the first round to the expected handicapping policy in the second round contest.

Finally, this paper is also related to recent contributions on optimal contest design. Ryvkin and Ortmann (2008) address the selection performance of different contest structures, but in contrast to our paper they discard the effect of this variation on incentives.[7] Groh, Moldovanu, Sela, and Sunde (2012) consider both objectives, but they focus on different seedings in a dynamic contest, while we investigate how dynamic contests relate to static ones when participants are heterogeneous. The only existing comparison of static and dynamic contests with heterogeneous participants by Stracke (2013) restricts attention to incentives and discards the selection performance.

3   The Model

3.1   A Promotion Contest with Heterogeneous Workers

Consider an entity that uses a promotion contest to fill a vacant high-level position which is of value P to workers from lower ranks.[8] Four risk neutral workers compete for the open position on the internal labor market. While working on their actual position, they are evaluated relative to their competitors, and the worker with the best performance is promoted at the end of the evaluation period. Workers know that they are being evaluated, and they are

---

[7]Brown and Minor (2011) provide an empirical test of the selection performance of two-stage pairwise elimination contests, which we analyze theoretically.

[8]The value of being promoted may include both monetary components (promotions imply higher wages) and non-monetary aspects (e.g., concerns for status or power). Higher wages may either be chosen by the contest-designing organization, or they may result from competition between organizations. Our modelling approach is consistent both with the concept of classic promotion tournaments á la Lazear and Rosen (1981) and market-based tournaments in the spirit of Waldman (1984). For a recent comparison of these two concepts, see Waldman (2013).

perfectly informed about their own productivity and the productivity of their colleagues.[9] To keep the theoretical analysis tractable, we assume that workers are of two different types: two workers are highly productive ("strong") and two workers are less productive ("weak"). Each worker provides effort to increase his chances for a promotion. The organizing entity of the contest, the "principal", cannot directly observe individual efforts, but receives only a noisy ordinal performance signal. As a consequence, the promotion probability of a given worker is increasing in the worker's own effort and decreasing in the effort(s) provided by the worker's immediate opponent(s). Since the signal is noisy, however, the worker with the highest effort does not win with certainty. Specifically, it is assumed that the promotion probability p_i for worker i in a contest interaction is given by the ratio of own effort x_i over the effort(s) provided by all workers who participate in the respective competition.[10] Denoting the effort(s) provided by the immediate competitor(s) of worker i by $X_{-i} = \sum_{m \neq i} x_m$, where $m \in N$ and N is the set of competitors, the promotion probability reads

$$p_i(x_i, X_{-i}) = \begin{cases} \frac{x_i}{x_i + X_{-i}} & \text{if} \quad x_i, X_{-i} > 0 \\ \frac{1}{\#N} & \text{if} \quad x_i, X_{-i} = 0 \end{cases}, \tag{1}$$

where $\#N$ is the number of workers participating in the contest interaction.

While we do not explicitly model the payoff function of the principal, we assume that it has two arguments. First, the payoff of the principal is increasing in effort provision by employees, i.e., higher work effort by employees translates into higher profits for the principal, and vice versa. Second, it is important for the profitability of the firm that productive rather than unproductive workers are promoted, as 'up-or-out' promotion policies imply that losers of the promotion competition are lost for the organization.[11] Thus, the payoff of the principal is increasing in the accuracy in selection of the promotion contest. In sum, we assume that the principal pursues the following two objectives:

1. Maximization of aggregate effort by all workers (Incentive Provision).

2. Maximization of the probability that a strong worker wins (Selection).

---

[9]In most professional occupations, the first promotion possibility for new hires is after one or two years. Therefore, workers who compete on the internal labor market for open positions usually know each other due to ongoing interactions in the workplace. The promotion contest for the succession of Jack Welch mentioned in the introduction, for example, lasted for six years.

[10]In technical terms we use a Tullock (1980) contest success function (CSF) with discriminatory power one. This format is equivalent to a perfectly discriminating CSF with multiplicative noise that follows the exponential distribution. See Konrad (2009) for details (p. 52).

[11]This holds even if losers are allowed to stay in the corporation. Workers who lose the competition for a promotion are certainly discouraged, such that they often apply at different companies, i.e., they will leave voluntarily.

Figure 1: Design Options Available to the Contest Designer
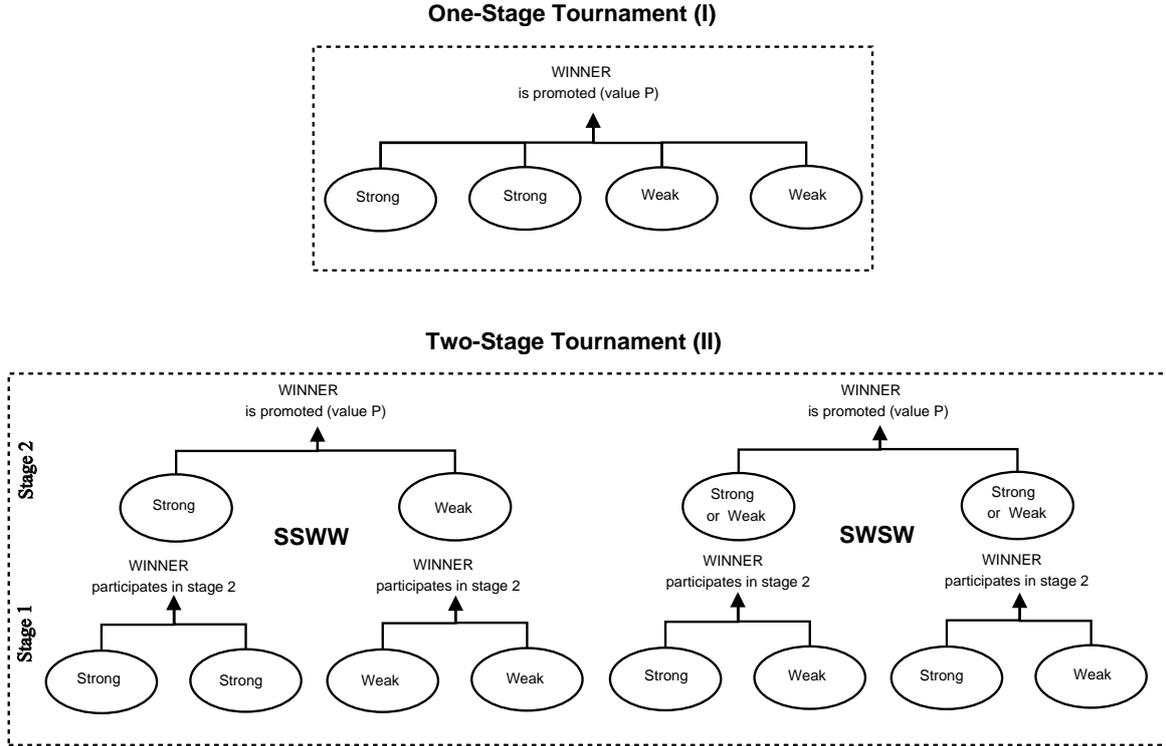
Arguably, the provision of incentives is an important goal for any corporation, and the prospect of being promoted to a better paid, more attractive position can be used to motivate and incentivize workers. At the same time, it is the inherent logic of promotion to promote productive and fire unproductive workers, in particular if 'up-or-out' promotion policies are used. One of the goals of our analysis is to find out whether a designer can improve both the incentive provision and the selection performance of a promotion contest by making the competition more (or less) dynamic. For this purpose, we compare the incentive and selection properties of a static (one-stage) contest with those of a dynamic (two-stage) pairwise elimination contest. The two different formats are depicted in Figure 1, which also shows that two different seedings are possible in the dynamic specification: Either a strong worker competes against another strong worker (and a weak worker against another weak worker) in the parallel stage-1 interactions (denoted setting SSWW); or both stage-1 interactions are mixed in terms of the productivity of the competing workers (setting SWSW). In the comparison of the static and the dynamic contest format, we assume that workers' productivities are not observable. Therefore, the seeding in stage 1 of the dynamic format is random: Setting SSWW occurs with probability 1/3, while the probability that setting SWSW realizes is 2/3.[12]

---

[12] After the first chosen worker has been eliminated from the pool of four workers, the probability that the next worker is of the same type is 1/3 (since only one of the remaining workers is of the same type), while the

3.2 Equilibrium Behavior by Workers in the One-Stage Contest

The one-stage contest, denoted I in the following, is a special case of the model developed by (and extensively discussed in) Stein (2002). Specifically, we consider a setting with only two worker types and assume that the effort costs of strong workers, $c_{\mathtt{S}}$, are lower than the effort costs of weak workers, $c_{\mathtt{W}}$ ($c_{\mathtt{S}} \leq c_{\mathtt{W}}$).[13] Since the one-stage contest is a simultaneous move game, the natural solution concept is Nash Equilibrium. In an equilibrium, worker i with marginal effort costs c_i chooses his/her effort $x_i \geq 0$ so as to maximize the expected payoff $\Pi_i(\mathtt{I})$, taking the total effort of all other workers, X_{-i}, as given. Formally, the optimization problem of worker i reads as follows:

$$\max_{x_i \geq 0} \Pi_i(x_i, X_{-i}) = \frac{x_i}{x_i + X_{-i}} P - c_i x_i.$$

The first-order conditions, together with symmetry, yield individual equilibrium efforts

$$x_{\mathtt{S}}^*(\mathtt{I}) = \begin{cases} \frac{3(2c_{\mathtt{W}} - c_{\mathtt{S}})}{4(c_{\mathtt{S}} + c_{\mathtt{W}})^2} P & \text{if } \frac{c_{\mathtt{W}}}{c_{\mathtt{S}}} < 2 \\ \frac{1}{4c_{\mathtt{S}}} P & \text{if } \frac{c_{\mathtt{W}}}{c_{\mathtt{S}}} \geq 2 \end{cases} \quad \text{and} \quad x_{\mathtt{W}}^*(\mathtt{I}) = \begin{cases} \frac{3(2c_{\mathtt{S}} - c_{\mathtt{W}})}{4(c_{\mathtt{S}} + c_{\mathtt{W}})^2} P & \text{if } \frac{c_{\mathtt{W}}}{c_{\mathtt{S}}} < 2 \\ 0 & \text{if } \frac{c_{\mathtt{W}}}{c_{\mathtt{S}}} \geq 2 \end{cases} \quad (2)$$

for strong and weak workers, respectively.[14] Equilibrium efforts determine both the incentive and the selection properties of the promotion contest. Our measure for the incentive provision performance of the one-stage contest, denoted $\mathcal{E}(\mathtt{I})$, is the sum of individual equilibrium efforts. Since two workers are strong and two are weak, we obtain

$$\mathcal{E}(\mathtt{I}) = 2x_{\mathtt{S}}^* + 2x_{\mathtt{W}}^*. \tag{3}$$

While the incentive provision measure depends on the absolute value of equilibrium efforts, winning probabilities depend on the ratio of equilibrium efforts. The selection performance, $\mathcal{S}(\mathtt{I})$, is defined as the probability that a strong worker wins the contest. The equilibrium winning probability of a strong worker must be multiplied by two, since two strong workers participate in the promotion contest, which gives

$$\mathcal{S}(\mathtt{I}) = \frac{x_{\mathtt{S}}^*}{x_{\mathtt{S}}^* + x_{\mathtt{W}}^*}. \tag{4}$$

---

probability that the next worker is of the other type is 2/3 (because two of the three remaining workers are of the other type).

[13] We model heterogeneity in terms of effort cost differences. Results for heterogeneity in valuations or in the mapping from effort to winning probabilities are analogous. Proofs are available from the authors upon request.

[14] A detailed derivation of these expressions is provided in Appendix A.

Table 1: Stage-2 Equilibrium Efforts of Different Agents

homogeneous interaction		heterogeneous interaction	
Strong (S)	Weak (W)	Strong (S)	Weak (W)
$x_{\texttt{S2}}^*(\texttt{SS}) = \frac{P}{4c_{\texttt{S}}}$	$x_{\texttt{W2}}^*(\texttt{WW}) = \frac{P}{4c_{\texttt{W}}}$	$x_{\texttt{S2}}^*(\texttt{SW}) = \frac{c_{\texttt{W}}}{(c_{\texttt{S}}+c_{\texttt{W}})^2}P$	$x_{\texttt{W2}}^*(\texttt{SW}) = \frac{c_{\texttt{S}}}{(c_{\texttt{S}}+c_{\texttt{W}})^2}P$

3.3 Equilibrium Behavior by Workers in the Two-Stage Contest

The relevant solution concept for the two-stage contest is Subgame Perfect Nash Equilibrium. The equilibrium is derived by backward induction. First, all possible stage-2 interactions must be solved. With four workers of two types, there are three potential stage-2 games, namely SS (both workers are strong), WW (both workers are weak), and SW (one strong and one weak worker). As before, we assume that the effort costs of strong workers, $c_{\texttt{S}}$, are lower than the effort costs of weak workers, $c_{\texttt{W}}$ ($c_{\texttt{S}} \leq c_{\texttt{W}}$). The formal optimization problem of worker i with effort cost c_i who competes against worker $j \neq i$ in stage 2 reads

$$\max_{x_{i2} \geq 0} \Pi_{i2}(x_{i2}, x_{j2}) = \frac{x_{i2}}{x_{i2} + x_{j2}}P - c_i x_{i2},$$

where x_{i2} and x_{j2} are individual efforts by workers i and j, respectively. A detailed solution of all stage-2 games is provided in Appendix B1. The analysis reveals that the equilibrium effort of each worker depends on the worker's own productivity and on the productivity of the worker's opponent.[15] The formal expressions of stage-2 equilibrium efforts are displayed in Table 1. In that table, $x_{\texttt{S2}}^*(\texttt{SS})$ denotes stage-2 equilibrium effort of a strong worker in the homogeneous interaction SS; $x_{\texttt{W2}}^*(\texttt{WW})$ denotes stage-2 equilibrium effort of a weak worker in the homogeneous interaction WW; $x_{\texttt{S2}}^*(\texttt{SW})$ and $x_{\texttt{W2}}^*(\texttt{SW})$ are the equilibrium efforts of strong and weak workers, respectively, in the heterogeneous stage-2 interaction SW. Since stage-2 equilibrium efforts solve the last stage of the game, we can now move forward to stage 1.

Setting SSWW. The stage-1 interactions in setting SSWW ensure that one strong and one weak worker reach stage 2 with certainty.[16] Consequently, SW is the only possible constellation on stage 2. Thus, each strong worker knows that, conditional on reaching stage 2, the opponent will be a weak worker, while each weak worker anticipates that the interaction on stage 2, if reached, will involve a strong competitor. The only reward for winning stage 1 is the par-

---

[15]The inverse of effort costs, $\frac{1}{c_i}$, can be interpreted as a measure of i's productivity.

[16]Stein and Rapoport (2004) study a framework that is similar to our SSWW scenario.

ticipation in stage 2, in which workers may then receive the promotion of value P. Thus, the expected equilibrium payoffs of a stage-2 interaction SW for strong and weak workers, $\Pi^*_{\text{S2}}(\text{SW}) \equiv \Pi_{i2}(x^*_{\text{S2}}(\text{SW}), x^*_{\text{W2}}(\text{SW}))$ and $\Pi^*_{\text{W2}}(\text{SW}) \equiv \Pi_{i2}(x^*_{\text{W2}}(\text{SW}), x^*_{\text{S2}}(\text{SW}))$, respectively, determine the continuation values for which workers compete in stage 1. This becomes clear when considering the optimization problem of some strong worker i, who competes with the second strong worker j:

$$\max_{x_{i1} \geq 0} \Pi_i(\text{SSWW}) = \frac{x_{i1}}{x_{i1} + x_{j1}} \Pi^*_{\text{S2}}(\text{SW}) - c_{\text{S}} x_{i1}.$$

The probability that worker i participates in stage 2, where participation is worth $\Pi^*_{\text{S2}}(\text{SW})$ in equilibrium, is increasing in his/her stage-1 effort x_{i1}. Similarly, the two weak workers compete for participation in stage 2, which is worth $\Pi^*_{\text{W2}}(\text{SW})$ for them. Let $x^*_{\text{S1}}(\text{SSWW})$ and $x^*_{\text{W1}}(\text{SSWW})$ denote the stage-1 equilibrium efforts in setting SSWW by strong and weak workers. They are given by

$$x^*_{\text{S1}}(\text{SSWW}) = \frac{c_{\text{W}}^2}{4c_{\text{S}}(c_{\text{S}} + c_{\text{W}})^2} P \quad \text{and} \quad x^*_{\text{W1}}(\text{SSWW}) = \frac{c_{\text{S}}^2}{4c_{\text{W}}(c_{\text{S}} + c_{\text{W}})^2} P,$$

respectively.[17] Then, the incentive measure for setting SSWW of the two-stage contest format, denoted $\mathcal{E}(\text{SSWW})$, is the sum of individual equilibrium efforts over all participants and both stages. Summing-up the equilibrium efforts of the two strong and the two weak workers in stage 1, and the equilibrium efforts of one strong and one weak worker in stage 2, we obtain:

$$\mathcal{E}(\text{SSWW}) = \underbrace{2[x^*_{\text{S1}}(\text{SSWW}) + x^*_{\text{W1}}(\text{SSWW})]}_{\text{stage 1 effort}} + \underbrace{x^*_{\text{S2}}(\text{SW}) + x^*_{\text{W2}}(\text{SW})}_{\text{stage 2 effort}}. \tag{5}$$

The selection measure, i.e., the probability that a strong worker wins the contest, is determined by relative effort provision of stage-2 participants. As mentioned previously, one strong and one weak worker compete in stage 2, independently of stage-1 outcomes. Therefore, the selection measure $\mathcal{S}(\text{SSWW})$ depends on the ratio of stage-2 equilibrium efforts $x^*_{\text{S2}}(\text{SW})$ and $x^*_{\text{W2}}(\text{SW})$:

$$\mathcal{S}(\text{SSWW}) = \frac{x^*_{\text{S2}}(\text{SW})}{x^*_{\text{S2}}(\text{SW}) + x^*_{\text{W2}}(\text{SW})}. \tag{6}$$

Setting SWSW. Since both stage-1 interactions are mixed in setting SWSW, the composition of the stage-2 competition is uncertain; any one of the three stage-2 games SS, WW, and SW is possible, as shown in Figure 1. Consequently, the solution of this setting is complicated by the fact that stage-1 continuation values are determined endogenously. To illustrate this complication, assume that a strong worker i and an arbitrary weak worker j compete in stage

---

[17]The derivation of equilibrium efforts is provided in Appendix B2.

1 for the right to participate in stage 2. Simultaneously, strong worker k and weak worker l compete for the remaining stage-2 slot in the other stage-1 interaction. Then, the optimization problems of workers i and j read

$$\max_{x_{i1}\geq 0}\Pi_i(\texttt{SWSW}) = \frac{x_{i1}}{x_{i1}+x_{j1}}\underbrace{\left[\frac{x_{k1}}{x_{k1}+x_{l1}}\pi^*_{\texttt{S2}}(\texttt{SS})+\frac{x_{l1}}{x_{k1}+x_{l1}}\pi^*_{\texttt{S2}}(\texttt{SW})\right]}_{\equiv P_i(x_{k1},x_{l1})}-c_{\texttt{S}}x_{i1}$$

$$\max_{x_{j1}\geq 0}\Pi_j(\texttt{SWSW}) = \frac{x_{j1}}{x_{i1}+x_{j1}}\underbrace{\left[\frac{x_{k1}}{x_{k1}+x_{l1}}\pi^*_{\texttt{W2}}(\texttt{SW})+\frac{x_{l1}}{x_{k1}+x_{l1}}\pi^*_{\texttt{W2}}(\texttt{WW})\right]}_{\equiv P_j(x_{k1},x_{l1})}-c_{\texttt{W}}x_{j1}\ .$$

Thus, the continuation values $P_i(x_{k1}, x_{l1})$ and $P_j(x_{k1}, x_{l1})$ of workers i and j, respectively, depend on the behavior of workers k and l in the parallel stage-1 interaction. The reason for this interdependence is that the expected equilibrium payoffs of the workers differ across the three potential stage-2 interactions $\texttt{SS}$, $\texttt{WW}$, and $\texttt{SW}$.[18] Obviously, the same holds for the continuation values $P_k(x_{i1}, x_{j1})$ and $P_l(x_{i1}, x_{j1})$ of workers k and l in the second stage-1 interaction. This implies that the two heterogeneous stage-1 interactions are linked through continuation values that are determined endogenously. This interesting technical complication is discussed in Appendix B.2, where we provide a detailed closed-form solution to this setting.[19] Using stage-1 equilibrium efforts $x^*_{\texttt{S1}}(\texttt{SWSW})$ and $x^*_{\texttt{W1}}(\texttt{SWSW})$ by strong and weak workers, respectively, we can compute the incentive measure for setting $\texttt{SWSW}$, denoted by $\mathcal{E}(\texttt{SWSW})$. The resulting expression reads:

$$\mathcal{E}(\texttt{SWSW}) = \underbrace{2*[x^*_{\texttt{S1}}(\texttt{SWSW})+x^*_{\texttt{W1}}(\texttt{SWSW})]}_{\text{stage 1 effort}} +$$
$$\underbrace{2*\{\pi^2 x^*_{\texttt{S2}}(\texttt{SS})+(1-\pi)^2 x^*_{\texttt{W2}}(\texttt{WW})+\pi(1-\pi)[x^*_{\texttt{S2}}(\texttt{SW})+x^*_{\texttt{W2}}(\texttt{SW})]\}}_{\text{stage 2 effort}}\ . \qquad (7)$$

In this expression, $\pi = \frac{x^*_{\texttt{S1}}(\texttt{SWSW})}{x^*_{\texttt{S1}}(\texttt{SWSW})+x^*_{\texttt{W1}}(\texttt{SWSW})}$ stands for the probability that a strong worker wins against the weak opponent in stage 1. This probability determines the likelihood for a particular stage-2 configuration: The stage-2 participants are both strong with probability π^2, both weak with probability $(1-\pi)^2$, and of different types with probability $2\pi(1-\pi)$. The probability π that a strong worker wins in stage 1 is also relevant for the selection performance of setting

---

[18]Conditional on reaching stage 2, workers of both types have a higher expected payoff from meeting a weak rather than a strong opponent; that is, $\Pi^*_{\texttt{W2}}(\texttt{WW}) > \Pi^*_{\texttt{W2}}(\texttt{SW})$ and $\Pi^*_{\texttt{S2}}(\texttt{SW}) > \Pi^*_{\texttt{S2}}(\texttt{SS})$. Details are provided in Appendix B.2.

[19]See equations (B.8) and (B.8) in Appendix B.2.

SWSW, $\mathcal{S}(\texttt{SWSW})$. It is defined as

$$\mathcal{S}(\texttt{SWSW}) \quad = \quad \pi^2 + 2\pi(1-\pi)\frac{x^*_{\texttt{S2}}(\texttt{SW})}{x^*_{\texttt{W2}}(\texttt{SW}) + x^*_{\texttt{S2}}(\texttt{SW})}. \tag{8}$$

Intuitively, a strong worker is promoted if either both strong workers win their stage-1 interactions, which happens with probability π^2, or if only one strong worker wins in stage 1, and subsequently also in stage 2.

Random Seeding. If the principal chooses the dynamic format and has no information about the productivity of the single workers, the seeding in stage 1 is random. In this case, setting SSWW occurs with probability $1/3$, and setting SWSW realizes with probability $2/3$. Consequently, the expected incentive provision measure for the two-stage promotion contest, denoted $\mathcal{E}(\texttt{II})$, is a weighted average of total effort provision in the two settings. Formally,

$$\mathcal{E}(\texttt{II}) = \frac{\mathcal{E}(\texttt{SSWW}) + 2 \cdot \mathcal{E}(\texttt{SWSW})}{3}. \tag{9}$$

Obviously, the same holds for the selection measure $\mathcal{S}(\texttt{II})$, which is a weighted average of $\mathcal{S}(\texttt{SSWW})$ and $\mathcal{S}(\texttt{SWSW})$. Formally,

$$\mathcal{S}(\texttt{II}) = \frac{\mathcal{S}(\texttt{SSWW}) + 2 \cdot \mathcal{S}(\texttt{SWSW})}{3} . \tag{10}$$

4 Comparing Incentive and Selection Properties

Using the above results on equilibrium behavior of workers, we can compare one- and two-stage contests (I versus II) to investigate how structural modifications of the contest affect incentive provision and selection performance.

Incentive provision and selection performance are identical in the one-stage and the two-stage contest if all workers are homogeneous, i.e., of the same productivity. The equality in terms of selection performance follows trivially from the homogeneity assumption: Either all workers are weak, and the probability that a strong worker wins is zero in both formats, or all workers are strong, and the probability that a strong worker wins is one in both formats. That the contest structure does not affect aggregate effort provision in the homogeneous case is less obvious. However, that this holds for the specification considered here has already been established by Gradstein and Konrad (1999).[20] Consequently, the comparison of these two structures in our world with heterogeneous workers allows for a rigorous investigation of the

---

[20]An intuition for this result is provided by Amegashie (2000).

effects of heterogeneity on workers' behavior in one-stage and two-stage contests. A formal comparison of the incentive measures $\mathcal{E}(\text{I})$ and $\mathcal{E}(\text{II})$, and the selection measures $\mathcal{S}(\text{I})$ and $\mathcal{S}(\text{II})$, respectively, delivers the following proposition:

Proposition 1. *If the cost of effort is strictly higher for weak than for strong workers,*

(a) *aggregate effort is strictly higher in the two- than in one-stage contest, i.e.,*

$$c_{\text{W}} > c_{\text{S}} \implies \mathcal{E}(\text{I}) < \mathcal{E}(\text{II});$$

(b) *the probability that a strong agent receives the promotion is strictly higher in the one- than in the two-stage contest, i.e.,*

$$c_{\text{W}} > c_{\text{S}} \implies \mathcal{S}(\text{I}) > \mathcal{S}(\text{II}).$$

Proof. See Appendix C. □

According to Proposition 1, a designer cannot improve both the incentive provision and the selection performance of a promotion contest at the same time by making the competition more (or less) dynamic. Thus, the trade-off between these two performance dimensions established by previous work cannot be solved by structural variations of the promotion contest. Modifications that improve the performance in one dimension lead to a deterioration of performance in the other dimension. In particular, we find that the two-stage contest (with random seeding) dominates the one-stage format in terms of incentive provision, whereas the opposite holds for selection performance.

Figure **??** provides a graphical illustration of the main result. Panel (a) plots the incentive measures of both contest formats, $\mathcal{E}(\text{I})$ and $\mathcal{E}(\text{II})$, as a function of the effort costs of weak workers, c_{W}, holding the effort costs of strong workers and the value of the promotion fixed at one ($c_{\text{S}} = 1$ and $P = 1$). The figure shows that the aggregate effort provision in the two-stage contest (indicated by the dotted line) is always above aggregate effort in the one-stage contest (the solid line), reflecting the statement in part (a) of Proposition 1. The difference is highest at the kink of the one-stage incentive measure (where weak workers drop-out voluntarily, see Appendix A for details), and decreases subsequently. For extremely high values of c_{W}, aggregate effort provision approaches 0.5 in both contest formats. The selection performance of both contest formats is illustrated in Panel (b) of Figure **??**, which plots the probability that a strong worker wins the respective contest, $\mathcal{S}(\text{I})$ and $\mathcal{S}(\text{II})$, as a function of the effort costs of

weak workers.[21] The figure shows that the one-stage format dominates the two-stage one in terms of its selection performance, confirming part (b) of the proposition. Again, the difference between the two contest formats is highest at the kink of the one-stage selection measure where weak workers drop out, and decreases for higher degrees of heterogeneity. Even if the costs of effort are five times as high for weak than for strong workers, however, the probability that a strong worker wins is still almost ten percentage points higher in the one-stage than in the two-stage contest. Only for extreme values of heterogeneity (when $c_{\mathtt{W}} \to \infty$), $\mathcal{S}(\mathtt{II})$ approaches $\mathcal{S}(\mathtt{I})$.[22]

To understand the trade-off between the two goals we consider, one has to distinguish between *absolute* and *relative* incentives for effort provision. *Relative* incentives (in terms of the ratio of the workers' efforts) determine the selection performance of a contest design: The lower the equilibrium efforts of weak workers are *relative* to the equilibrium efforts of strong workers, the better is the selection performance of a contest. This implies that the accuracy in selection is increasing in the degree of heterogeneity between workers. *Absolute* incentives (that is, the sum of the workers' efforts) determine total effort. It is well known that *absolute* incentives are decreasing in the degree of heterogeneity.[23] This is confirmed by our findings: In the static and dynamic contest format, total effort provision decreases when the gap between the effort costs of weak workers and those of strong workers increases. Taken together, these considerations imply that the curves for incentive provision displayed in panel (a) of Figure **??** are downwards sloping, while the curves for selection shown in panel (b) are upwards sloping. These considerations are not informative about the position of one curve relative to the other, however. For this comparison, the following analogy is helpful: Contest structures that amplify the degree of heterogeneity between strong and weak workers perform better in terms of selection, as heterogeneity discourages weak workers relatively more than it induces strong workers to slack off. At the same time, the more a contest design moderates the heterogeneity between types, the better is its performance in the incentive dimension, since heterogeneity decreases the incentives for effort provision for both strong and weak workers in absolute terms. Consequently, the structural variation considered in this paper works analogously to a strategic handicap – the strategic advantage of strong over weak workers is higher in the static than in the dynamic promotion contest, which is a different way of saying that the static contest format handicaps weak workers by its structure.

---

[21]The effort costs of strong workers are again normalized to one.

[22]This is not shown in Panel (b) of Figure **??**, but a formal derivation is straightforward. Details available upon request.

[23]See Lazear and Rosen (1981) and Gürtler and Kräkel (2010) for theoretical work addressing this issue, and Sunde (2009) or Brown (2011) for empirical evidence.

Figure 2: Performance in One-Stage and Two-Stage Contests
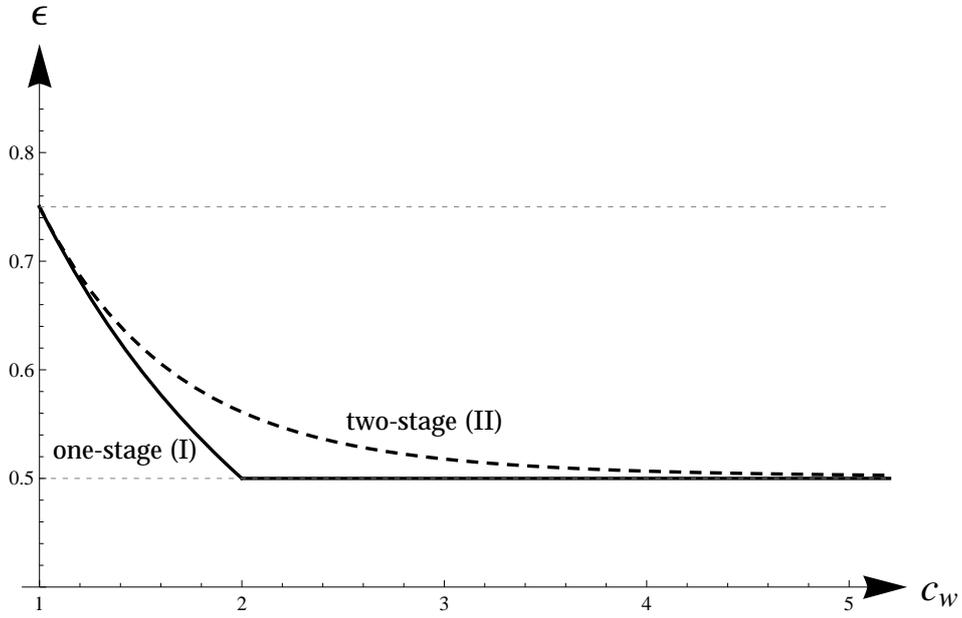
(a) Incentive Provision
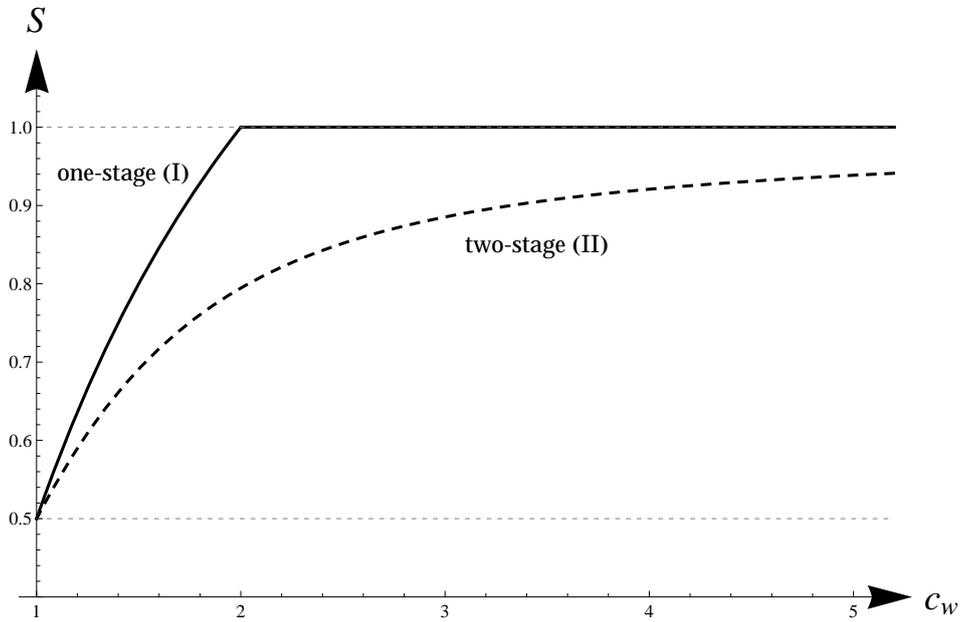
(b) Selection

Notes: Panel (a) plots expressions (3) and (9) with $c_\mathsf{s} = 1$ and $P = 1$; panel (b) plots (6) and (8) under the same assumption.

14

5   Discussion and Additional Results

The main goal of handicapping strategies discussed in the existing literature is to reduce the adverse effect of heterogeneity on incentives. Such handicapping strategies require identification of worker types, however, which is often impossible. Structural handicapping has the key advantage of being feasible even if identification of abilities is impossible. Thus, it can also be used in applications where selection is of primary interest to the designer. In this case, it should be the weak rather than the strong worker who is handicapped, which is just the opposite of what the literature on handicaps for the maximization of incentive provision suggests.[24]

Recent work by Groh et al. (2012) suggests that the trade-off between incentive provision and selection performance disappears in some settings. They consider a two-stage pairwise elimination contest with heterogeneous contestants and investigate how the seeding of types in stage 1 affects aggregate incentives and selection performance. When comparing equilibrium behavior in settings SSWW and SWSW, they find that SWSW dominates in terms of incentive provision and in terms of selection performance. This result is surprising in light of the previous discussion, since the seeding of types in stage 1 can also be seen as a handicapping strategy: Setting SWSW handicaps weak workers, as they must beat a strong competitor to stay in the contest, which simplifies the advancement of strong workers to the second round. In contrast, setting SSWW handicaps strong workers; they have to win against a strong competitor, while the seeding ensures that one weak worker makes it to the second stage for sure. This reasoning is confirmed by the following result, which formally compares the incentive measures $\mathcal{E}(\mathtt{SSWW})$ and $\mathcal{E}(\mathtt{SWSW})$, and the selection measures $\mathcal{S}(\mathtt{SSWW})$ and $\mathcal{S}(\mathtt{SWSW})$ for the case where workers are heterogeneous:

Proposition 2. *If the cost of effort is strictly higher for weak than for strong workers,*

(a) aggregate effort is strictly higher in setting SSWW than in setting SWSW, i.e.,

$$c_W > c_S \;\Rightarrow\; \mathcal{E}(\mathit{SSWW}) > \mathcal{E}(\mathit{SWSW});$$

(b) the probability that a strong agent receives the promotion is strictly higher in setting SWSW

---

[24]To our knowledge, the only exception is the paper by Ridlon and Shin (2013). They show that it can sometimes be optimal to handicap the first-round loser, even though this tends to make the second-round competition more heterogeneous and therefore less intense. The reason is that handicapping the first-round loser implies that both workers provide more effort in the first-period competition as winning this interaction is more attractive. As Ridlon and Shin show, the total effect on incentives across both periods depends on the degree of heterogeneity.

than in setting `SSWW`, *i.e.,*

$$c_W > c_S \;\Rightarrow\; \mathcal{S}(\texttt{SSWW}) < \mathcal{S}(\texttt{SWSW}).$$

Proof. See Appendix C. □

Hence, there is a trade-off between incentive provision and selection performance across the two settings `SSWW` and `SWSW` for the theoretical specification we consider, in contrast to what Groh et al. (2012) observe in their specification. This difference can be explained by properties of the contest technology. Groh et al. (2012) consider a perfectly discriminating contest where a marginal lead in effort translates into a sure win. Thus, weak workers must outperform their strong competitors in the effort dimension to win the competition in Groh et al., while the ordinal performance measure is determined by both effort and a random component in the lottery contest considered here. The strategic disadvantage for weak workers is therefore more pronounced in the perfectly discriminating contest. This difference has no effect on the qualitative selection properties of a contest, but it does affect incentive provision: In line with what our model suggests, Groh et al. find that setting `SWSW` dominates `SSWW` in terms of selection. However, their results also show that setting `SWSW` dominates in the incentive dimension, in contrast to what we find in our model. Consequently, a structural handicap for strong workers has a detrimental rather than a positive impact on incentive provision in their model. While this seems surprising on first sight, it makes perfect sense intuitively. The strategic disadvantage of weak workers in the perfectly discriminating format makes it optimal to exclude all weak types from the competition – total effort is (weakly) higher in a pair-wise interaction of the two strong workers than in any seeding variant of the multi-stage contest.

The previous comparison of seeding variants in two different contest models suggests that structural modifications may also be able to improve both the incentive properties and the selection performance of a contest scheme if the baseline setting induces suboptimal participation of weak types. In this special case, structural handicaps for strong workers may affect incentive provision negatively rather than positively in contests with heterogeneous workers, such that the trade-off disappears.[25] This is rather the exception than the rule in promotion contests, since participation of weak types is only suboptimal if their strategic disadvantage is extreme.

---

[25]The expected equilibrium effort by weak workers is positive in setting `SSWW`, and zero in `SWSW`. See Groh et al. (2012) for details.

6   Concluding Remarks

This paper has investigated how structural variations of a contest affect incentive provision and selection performance in promotion contests. A comparison of *static* (one-shot) and *dynamic* (two-stage) contests suggests that the two goals are incompatible. Thus, a designer cannot improve both the incentive provision and the selection performance of a promotion contest by making the competition more (or less) dynamic. This implies that multiple instruments should be used whenever both goals are equally important.

Another important implication is that structural variations work like strategic handicaps. This insight is highly relevant for applications where contest schemes are either used only for incentive provision or only for selection purposes. In contrast to handicapping strategies that are discussed in the existing literature, a designer who relies on structural handicaps is not required to identify the abilities of competing workers – or at least to observe some signal of their productivities – to implement the handicap. Therefore, structural handicaps might help to improve the performance of contest schemes whenever the types of competing workers are unobservable by the designer. While the analysis in this paper compares two of the most prominent structures, it is only the first step toward a more general analysis of structural handicaps, which constitutes an interesting avenue for future research.

References

AMEGASHIE, J. A. (2000): "Some Results on Rent-Seeking Contests with Shortlisting," *Public Choice*, 105, 245–253.

BAKER, G. P., M. C. JENSEN, AND K. J. MURPHY (1988): "Compensation and Incentives: Practice vs. Theory," *Journal of Finance*, 43, 593–616.

BERNHARDT, D. (1995): "Strategic Promotion and Compensation," *Review of Economic Studies*, 62, 315–339.

BROWN, J. (2011): "Quitters Never Win: The (Adverse) Incentive Effects of Competing with Superstars," *Journal of Political Economy*, 119, 982–1013.

BROWN, J., AND D. MINOR (2011): "Selecting the Best: Spillover and Shadows in Elimination Tournaments," *NBER Working Paper No. 17639*.

CORNES, R., AND R. HARTLEY (2005): "Asymmetric Contests with General Technologies," *Economic Theory*, 26, 923–946.

GRADSTEIN, M., AND K. A. KONRAD (1999): "Orchestrating Rent Seeking Contests," *Economic Journal*, 109, 536–545.

GROH, C., B. MOLDOVANU, A. SELA, AND U. SUNDE (2012): "Optimal Seedings in Elimination Tournaments," *Economic Theory*, 49, 59–80.

GÜRTLER, O., AND M. KRÄKEL (2010): "Optimal Tournament Contracts for Heterogeneous Workers," *Journal of Economic Behavior & Organization*, 75, 180 – 191.

KONRAD, K. (2009): *Strategy and Dynamics in Contests*. Oxford University Press.

LAZEAR, E. P., AND S. ROSEN (1981): "Rank-Order Tournaments as Optimal Labor Contracts," *Journal of Political Economy*, 89, 841–864.

MALCOMSON, J. M. (1984): "Work Incentives, Hierarchy, and Internal Labor Markets," *Journal of Political Economy*, 92, 486–507.

NTI, K. O. (1999): "Rent-Seeking with Asymmetric Valuations," *Public Choice*, 98, 415–430.

PRENDERGAST, C. (1993): "The Role of Promotion in Inducing Specific Human Capital Aquisition," *Quarterly Journal of Economics*, 108, 523–534.

——— (1999): "The Provision of Incentives in Firms," *Journal of Economic Literature*, 37, 7–63.

RIDLON, R., AND J. SHIN (2013): "Favoring the Winner or Looser in Repeated Contests," *Marketing Science*, 32, 768–785.

ROSEN, S. (1986): "Prizes and Incentives in Elimination Tournaments," *American Economic Review*, 76, 701–715.

RYVKIN, D., AND A. ORTMANN (2008): "The Predictive Power of Three Prominent Tournament Formats," *Management Science*, 54, 492–504.

Stein, W. E. (2002): "Asymmetric Rent-Seeking with More than Two Contestants," *Public Choice*, 113, 325–336.

Stein, W. E., and A. Rapoport (2004): "Asymmetric Two-Stage Group Rent-Seeking: Comparison of Two Contest Structures," *Public Choice*, 124, 309–328.

Stracke, R. (2013): "Contest Design and Heterogeneity," *Economics Letters*, 121, 4–7.

Sunde, U. (2009): "Heterogeneity and Performance in Tournaments: A Test for Incentive Effects using Professional Tennis Data," *Applied Economics*, 41, 3199–3208.

Tsoulouhas, T., C. R. Knoeber, and A. Agrawal (2007): "Contests to Become CEO: Incentives, Selection and Handicaps," *Economic Theory*, 30, 195–221.

Tullock, G. (1980): "Efficient Rent-Seeking," in *Toward a Theory of the Rent-Seeking Society*. J.M. Buchanan and R.D. Tollison and G. Tullock (Eds.). Texas A&M Press, College Station, p. 97-112.

Waldman, M. (1984): "Job Assignments, Signalling, and Efficiency," *Rand Journal of Economics*, 15, 255–267.

——— (1990): "Up-or-Out Contracts: A Signaling Perspective," *Journal of Labor Economics*, 8, 230–250.

——— (2013): "Classic Promotion Tournaments versus Market-Based Tournaments," *Internation Journal of Industrial Organization*, 31, 198–210.

Welch, J. (2001): *Jack: Straight From the Gut*. Warner Books, New York.

Appendix

A Solution of the One-Stage Contest

Due to symmetry, it suffices to solve the optimization problem of one strong and one weak worker. Without loss of generality, we consider the strong worker i and the weak worker k and obtain

$$
\begin{aligned}
\max_{x_i \geq 0} \Pi_i(\mathtt{I}) &= \frac{x_i}{x_i + X_{-i}} P - c_{\mathtt{S}} x_i, \\
\max_{x_k \geq 0} \Pi_k(\mathtt{I}) &= \frac{x_k}{x_k + X_{-k}} P - c_{\mathtt{W}} x_k,
\end{aligned}
$$

where $X_{-i} = x_j + x_k + x_l$, and $X_{-k} = x_i + x_j + x_l$. This leads to the two first-order optimality conditions $X_{-i}P = c_{\mathtt{S}}(x_i + X_{-i})^2$ and $X_{-k}P = c_{\mathtt{W}}(x_k + X_{-k})^2$. Combining these conditions with symmetry reveals that the relation

$$
x_{\mathtt{W}}^* = \frac{2c_{\mathtt{S}} - c_{\mathtt{W}}}{2c_{\mathtt{W}} - c_{\mathtt{S}}} x_{\mathtt{S}}^* \tag{A.1}
$$

holds in an interior NE. Since the equilibrium efforts cannot be negative, a corner solution (with $x_{\mathtt{W}}^* = 0$) applies for $c_{\mathtt{W}} \geq 2c_{\mathtt{S}}$. In other words, weak workers drop out from the competition voluntarily for large differences in productivity (for $c_{\mathtt{W}} \geq 2c_{\mathtt{S}}$), leaving the two strong workers as the only contenders for the prize. Taking these considerations into account, the equilibrium efforts of strong and weak workers are given by

$$
x_{\mathtt{S}}^*(\mathtt{I}) = \begin{cases} \frac{3(2c_{\mathtt{W}} - c_{\mathtt{S}})}{4(c_{\mathtt{S}} + c_{\mathtt{W}})^2} P & \text{if } \frac{c_{\mathtt{W}}}{c_{\mathtt{S}}} < 2 \\ \frac{1}{4c_{\mathtt{S}}} P & \text{if } \frac{c_{\mathtt{W}}}{c_{\mathtt{S}}} \geq 2 \end{cases} \quad \text{and} \quad x_{\mathtt{W}}^*(\mathtt{I}) = \begin{cases} \frac{3(2c_{\mathtt{S}} - c_{\mathtt{W}})}{4(c_{\mathtt{S}} + c_{\mathtt{W}})^2} P & \text{if } \frac{c_{\mathtt{W}}}{c_{\mathtt{S}}} < 2 \\ 0 & \text{if } \frac{c_{\mathtt{W}}}{c_{\mathtt{S}}} \geq 2 \end{cases}. \tag{A.2}
$$

B Solution of the Two-Stage Contest

B.1 Solution for Stage 2

Assume that two workers i and j with cost of effort c_i and c_j, respectively, compete against each other in stage 2. The optimization problems of these two workers, who maximizes their stage-2 payoff Π_{i2} and Π_{j2}, respectively by choosing an optimal level of effort x_{i2} (x_{i2}), read as

follows:[26]

$$\max_{x_{\text{i2}} \geq 0} \Pi_{\text{i2}} = \frac{x_{\text{i2}}}{x_{\text{i2}} + x_{\text{j2}}} P - c_{\text{i}} x_{\text{i2}},$$

$$\max_{x_{\text{j2}} \geq 0} \Pi_{\text{j2}} = \frac{x_{\text{j2}}}{x_{\text{i2}} + x_{\text{j2}}} P - c_{\text{j}} x_{\text{j2}}.$$

First order conditions are necessary as well as sufficient in any pair-wise interaction for the lottery CSF (see Nti, 1999, or Cornes and Hartley, 2005). The combination of first-order conditions implies equilibrium efforts

$$x_{\text{i2}}^* = \frac{c_{\text{j}}}{(c_{\text{i}} + c_{\text{j}})^2} P \quad \text{and} \quad x_{\text{j2}}^* = \frac{c_{\text{i}}}{(c_{\text{i}} + c_{\text{j}})^2} P, \tag{B.1}$$

respectively. Inserting optimal actions in the two objective functions gives the expected equilibrium payoffs

$$\Pi_{\text{i2}}^* = \frac{c_{\text{j}}^2}{(c_{\text{i}} + c_{\text{j}})^2} P \quad \text{and} \quad \Pi_{\text{j2}}^* = \frac{c_{\text{i}}^2}{(c_{\text{i}} + c_{\text{j}})^2} P \tag{B.2}$$

Equations (B.1) and (B.2) characterize equilibrium efforts and payoffs, respectively, for any possible combination of types, i.e., for SS, SW, and WW. In the main text of this paper, the particular contest environment considered (that is, SS, SW, or WW) is in parentheses – as in $x_{i2}(\text{SS})$ or $\Pi_{i2}(\text{SS})$, for example – or is omitted when there is no risk of confusion.

B.2 Solution for Stage 1

Setting SSWW. Due to symmetry of the optimization problems, it suffices to solve the optimization problem of one strong worker (i or k), and one weak worker (j or l). Without loss of generality, we consider the maximization problems of workers i and j,

$$\max_{x_{i1} \geq 0} \Pi_i(\text{SSWW}) = \frac{x_{i1}}{x_{i1} + x_{k1}} \Pi_{\text{S2}}^*(\text{SW}) - c_{\text{S}} x_{i1},$$

$$\max_{x_{j1} \geq 0} \Pi_j(\text{SSWW}) = \frac{x_{j1}}{x_{j1} + x_{l1}} \Pi_{\text{W2}}^*(\text{SW}) - c_{\text{W}} x_{j1}.$$

The optimization problem for strong workers is similar to the one considered in stage-2 interaction SS; the only difference is the prize, which now amounts to $\Pi_{\text{S2}}^*(\text{SW})$ rather than P. Analogously, weak workers face the same situation as in stage-2 interaction WW with a different

---

[26]Throughout the paper the first subscript of the variables Π and x indicates the player, while the second subscript indicates the stage.

prize ($\Pi_{W2}^*(SW)$ instead of P). Consequently, first-order and symmetry conditions deliver stage-1 equilibrium efforts

$$x_{S1}^*(\text{SSWW}) \equiv x_{i1}^*(\text{SSWW}) = x_{k1}^*(\text{SSWW}) = \frac{c_W^2}{4c_S(c_S + c_W)^2} P \tag{B.3}$$

$$x_{W1}^*(\text{SSWW}) \equiv x_{j1}^*(\text{SSWW}) = x_{l1}^*(\text{SSWW}) = \frac{c_S^2}{4c_W(c_S + c_W)^2} P. \tag{B.4}$$

Setting SWSW. We assume (without loss of generality) that workers i and k are strong, whereas workers j and l are weak, and that the two pairwise stage-1 interactions are between workers i and j, and between workers k and l, respectively. We start by considering the decision problem of strong worker i and weak worker j. Both workers choose their optimal stage-1 effort, given equilibrium behavior in any potential stage-2 interaction. The optimization problems are

$$\max_{x_{i1} \geq 0} \Pi_i(\text{SWSW}) = \frac{x_{i1}}{x_{i1} + x_{j1}} \underbrace{\left[\frac{x_{k1}}{x_{k1} + x_{l1}} \Pi_{S2}^*(\text{SS}) + \frac{x_{l1}}{x_{k1} + x_{l1}} \Pi_{S2}^*(\text{SW}) \right]}_{\equiv P_i(x_{k1}, x_{l1})} - c_S x_{i1}$$

$$\max_{x_{j1} \geq 0} \Pi_j(\text{SWSW}) = \frac{x_{j1}}{x_{i1} + x_{j1}} \underbrace{\left[\frac{x_{k1}}{x_{k1} + x_{l1}} \Pi_{W2}^*(\text{SW}) + \frac{x_{l1}}{x_{k1} + x_{l1}} \Pi_{W2}^*(\text{WW}) \right]}_{\equiv P_j(x_{k1}, x_{l1})} - c_W x_{j1} .$$

The continuation values $P_i(x_{k1}, x_{l1})$ and $P_j(x_{k1}, x_{l1})$ of workers i and j, respectively, depend on the behavior of workers k and l in the other stage-1 interaction. Similarly, the continuation values $P_k(x_{i1}, x_{j1})$ and $P_l(x_{i1}, x_{j1})$ of workers k and l depend on the behavior of workers i and j. Therefore, the two stage-1 interactions are linked through endogenously determined continuation values. The reason is that expected equilibrium payoffs for workers differ across the three potential stage-2 interactions SS, WW, and SW. Conditional on reaching stage 2, workers of both types have a higher expected payoff from meeting a weak rather than a strong opponent, since $\Pi_{W2}^*(\text{WW}) > \Pi_{W2}^*(\text{SW})$ and $\Pi_{S2}^*(\text{SW}) > \Pi_{S2}^*(\text{SS})$. However, each worker takes the probability that the opponent is of a certain type as given, since it is determined in the parallel stage-1 interaction. The first-order conditions for the interaction between i and j read

$$x_{j1} P_i(x_{k1}, x_{l1}) - c_S(x_{i1} + x_{j1})^2 = 0 \quad \text{and} \quad x_{i1} P_j(x_{k1}, x_{l1}) - c_W(x_{i1} + x_{j1})^2 = 0.$$

The respective conditions for the other stage-1 interaction between workers k and l are

$$x_{l1} P_k(x_{i1}, x_{j1}) - c_S(x_{k1} + x_{l1})^2 = 0 \quad \text{and} \quad x_{k1} P_l(x_{i1}, x_{j1}) - c_W(x_{k1} + x_{l1})^2 = 0 .$$

Combining the four conditions, we obtain two expressions that define a relation between equilibrium effort choices of workers *within* each interaction, namely

$$\frac{x_{i1}}{x_{j1}} = \frac{c_{\text{W}}}{c_{\text{S}}} \frac{P_i(x_{k1}, x_{l1})}{P_j(x_{k1}, x_{l1})} \quad \text{and} \quad \frac{x_{k1}}{x_{l1}} = \frac{c_{\text{W}}}{c_{\text{S}}} \frac{P_k(x_{i1}, x_{j1})}{P_l(x_{i1}, x_{j1})}, \tag{B.5}$$

respectively. These expressions show that each stage-1 interaction is a contest between workers with different costs and endogenously determined valuations of winning. While the costs of effort differ by construction, differences of the value for winning are a result of the contest structure: Reaching stage 2 is more valuable for strong than for weak workers.

We proceed now to the solution of the problem, which comprises two heterogeneous participants with regard to their effort costs and their valuation. As mentioned previously, any contest with two heterogeneous participants has a unique, interior equilibrium for the chosen contest success function (Cornes and Hartley 2005, Nti 1999. Consequently, each of the two pairwise stage-1 interactions has a unique equilibrium for each pair of continuation values. What remains to be shown is that the two expressions in (B.5) can be satisfied *jointly* such that both stage-1 interactions are satisfied simultaneously in equilibrium. Inserting the continuation values in (B.5) and simplifying gives

$$\frac{x_{i1}}{x_{j1}} = \frac{c_{\text{W}}}{c_{\text{S}}} \frac{(c_{\text{S}} + c_{\text{W}})^2 \frac{x_{k1}}{x_{l1}} + 4c_{\text{W}}^2}{4c_{\text{S}}^2 \frac{x_{k1}}{x_{l1}} + (c_{\text{S}} + c_{\text{W}})^2} \quad \text{and} \quad \frac{x_{k1}}{x_{l1}} = \frac{c_{\text{W}}}{c_{\text{S}}} \frac{(c_{\text{S}} + c_{\text{W}})^2 \frac{x_{i1}}{x_{j1}} + 4c_{\text{W}}^2}{4c_{\text{S}}^2 \frac{x_{i1}}{x_{j1}} + (c_{\text{S}} + c_{\text{W}})^2}. \tag{B.6}$$

System (B.6) consists of two equations in the two unknowns $\frac{x_{i1}^*}{x_{j1}^*}$ and $\frac{x_{k1}^*}{x_{l1}^*}$, respectively. Note that both equations are symmetric, since the two workers in each of the two stage-1 interactions face identical optimization problems. This implies that $x_{\text{S1}}^* \equiv x_{i1}^* = x_{k1}^*$ and $x_{\text{W1}}^* \equiv x_{j1}^* = x_{l1}^*$ do hold in the symmetric equilibrium.[27] Combining these conditions and (B.6) gives:

$$
\begin{aligned}
\frac{x_{\text{S1}}^*}{x_{\text{W1}}^*} &= \frac{c_{\text{W}}}{c_{\text{S}}} \frac{(c_{\text{S}} + c_{\text{W}})^2 \frac{x_{\text{S1}}^*}{x_{\text{W1}}^*} + 4c_{\text{W}}^2}{4c_{\text{S}}^2 \frac{x_{\text{S1}}^*}{x_{\text{W1}}^*} + (c_{\text{S}} + c_{\text{W}})^2} \\
\Leftrightarrow \quad 0 &= 4c_{\text{S}}^2 \left[\frac{x_{\text{S1}}^*}{x_{\text{W1}}^*}\right]^2 + \left(1 - \frac{c_{\text{W}}}{c_{\text{S}}}\right)(c_{\text{S}} + c_{\text{W}})^2 \left[\frac{x_{\text{S1}}^*}{x_{\text{W1}}^*}\right] - 4\frac{c_{\text{W}}^3}{c_{\text{S}}} \\
\Leftrightarrow \quad \frac{x_{\text{S1}}^*}{x_{\text{W1}}^*} &= F^*(c_{\text{S}}, c_{\text{W}}),
\end{aligned}
$$

---

[27] The symmetric equilibrium exists for any degree of heterogeneity and is unique. Intuitively, one must show that the graphs of the two relations in (B.6) have a unique intersection in the domain defined by $\frac{x_{j1}^*}{x_{i1}^*} \in [0, 1]$ and $\frac{x_{l1}^*}{x_{k1}^*} \in [0, 1]$. It suffices to consider this domain, since the assumption of lower costs of effort and the resulting higher value of winning of strong workers imply that $x_{i1}^* \geq x_{j1}^*$ and $x_{k1}^* \geq x_{l1}^*$, respectively. This follows from (B.5). A complete formal proof is available from the corresponding author upon request.

where

$$F^*(c_{\mathrm{S}}, c_{\mathrm{W}}) = \frac{(c_{\mathrm{W}} - c_{\mathrm{S}})(c_{\mathrm{S}} + c_{\mathrm{W}})^2 + \sqrt{64c_{\mathrm{W}}^3 c_{\mathrm{S}}^3 + (c_{\mathrm{S}} - c_{\mathrm{W}})^2(c_{\mathrm{S}} + c_{\mathrm{W}})^4}}{8c_{\mathrm{S}}^3}. \qquad (\mathrm{B.7})$$

$F^*(c_{\mathrm{S}}, c_{\mathrm{W}})$ is the ratio of stage-1 efforts of the two worker types, which is directly proportional to heterogeneity in costs and continuation values, as equation (B.5) shows. Therefore, $F^*(c_{\mathrm{S}}, c_{\mathrm{W}})$ can be interpreted as a measure for both the exogenous heterogeneity in effort costs between strong and weak workers and the endogenous heterogeneity between types that is due to different continuation values in stage 1. The expression $F^*(c_{\mathrm{S}}, c_{\mathrm{W}})$ allows us to disentangle and solve analytically the two interdependent stage-1 interactions. We start by considering the continuation values which satisfy

$$P_i(x_{\mathrm{S}1}^*, x_{\mathrm{W}1}^*) = P_k(x_{\mathrm{S}1}^*, x_{\mathrm{W}1}^*) = \frac{(c_{\mathrm{S}} + c_{\mathrm{W}})^2 F^*(c_{\mathrm{W}}, c_{\mathrm{S}}) + 4c_{\mathrm{W}}^2}{4(c_{\mathrm{S}} + c_{\mathrm{W}})^2[1 + F^*(c_{\mathrm{S}}, c_{\mathrm{W}})]} P \,,$$

$$P_j(x_{\mathrm{S}1}^*, x_{\mathrm{W}1}^*) = P_l(x_{\mathrm{S}1}^*, x_{\mathrm{W}1}^*) = \frac{(c_{\mathrm{S}} + c_{\mathrm{W}})^2 + 4c_{\mathrm{S}}^2 F^*(c_{\mathrm{W}}, c_{\mathrm{S}})}{4(c_{\mathrm{S}} + c_{\mathrm{W}})^2[1 + F^*(c_{\mathrm{S}}, c_{\mathrm{W}})]} P.$$

Note that $P_i(x_{\mathrm{S}1}^*, x_{\mathrm{W}1}^*) = P_k(x_{\mathrm{S}1}^*, x_{\mathrm{W}1}^*)$ and $P_j(x_{\mathrm{S}1}^*, x_{\mathrm{W}1}^*) = P_l(x_{\mathrm{S}1}^*, x_{\mathrm{W}1}^*)$ due to symmetry. Given these continuation values, stage-1 equilibrium efforts can be determined as

$$x_{\mathrm{S}1}^*(\mathrm{SWSW}) \equiv x_{i1}^*(\mathrm{SWSW}) = x_{k1}^*(\mathrm{SWSW}) = \frac{(c_{\mathrm{S}} + c_{\mathrm{W}})^2 F^*(c_{\mathrm{W}}, c_{\mathrm{S}})^2 + 4c_{\mathrm{W}}^2 F^*(c_{\mathrm{W}}, c_{\mathrm{S}})}{4c_{\mathrm{S}}(c_{\mathrm{S}} + c_{\mathrm{W}})^2[1 + F^*(c_{\mathrm{S}}, c_{\mathrm{W}})]^3} P \qquad (\mathrm{B.8})$$

$$x_{\mathrm{W}1}^*(\mathrm{SWSW}) \equiv x_{j1}^*(\mathrm{SWSW}) = x_{l1}^*(\mathrm{SWSW}) = \frac{(c_{\mathrm{S}} + c_{\mathrm{W}})^2 F^*(c_{\mathrm{W}}, c_{\mathrm{S}}) + 4c_{\mathrm{S}}^2 F^*(c_{\mathrm{W}}, c_{\mathrm{S}})^2}{4c_{\mathrm{W}}(c_{\mathrm{S}} + c_{\mathrm{W}})^2[1 + F^*(c_{\mathrm{S}}, c_{\mathrm{W}})]^3} P \,. \qquad (\mathrm{B.9})$$

C   Proofs of Propositions 1 and 2

Lemma 1. *Assume without loss of generality that $c_{\mathrm{W}} \geq c_{\mathrm{S}} = 1$ and define $f(c_{\mathrm{W}}) = \frac{5c_{\mathrm{W}}^3 + 2c_{\mathrm{W}}^2 + c_{\mathrm{W}}}{c_{\mathrm{W}}^2 + 2c_{\mathrm{W}} + 5}$. Then, the relation $F^*(1, c_{\mathrm{W}}) > f(c_{\mathrm{W}})$ does hold for all $c_{\mathrm{W}} > 1$, where $F^*(1, c_{\mathrm{W}})$ is defined as in (B.7). Furthermore, for $c_{\mathrm{W}} = 1$ it holds that $F^*(1, c_{\mathrm{W}}) = f(c_{\mathrm{W}})$.*

Proof. From equation (B.5), we know that $\frac{x_{i1}}{x_{j1}} = \frac{c_{\mathrm{W}}}{c_{\mathrm{S}}} \frac{P_i(x_{k1}, x_{l1})}{P_j(x_{k1}, x_{l1})}$. Further, equation (B.7) tells us that $\frac{x_{i1}^*}{x_{j1}^*} = F^*(c_{\mathrm{S}}, c_{\mathrm{W}})$. Consequently, using the assumption that $c_{\mathrm{W}} \geq c_{\mathrm{S}} = 1$, it must hold that

$$F^*(1, c_{\mathrm{W}}) = c_{\mathrm{W}} \frac{P_i(x_{k1}, x_{l1})}{P_j(x_{k1}, x_{l1})} = \frac{4c_{\mathrm{W}}^3 + c_{\mathrm{W}}(1 + c_{\mathrm{W}})^2 \times \frac{x_{k1}}{x_{l1}}}{(1 + c_{\mathrm{W}})^2 + 4 \times \frac{x_{k1}}{x_{l1}}}.$$

Note that

$$\frac{\partial F^*(1, c_{\mathrm{W}})}{\partial \frac{x_{k1}}{x_{l1}}} = \frac{(1 + c_{\mathrm{W}})^4 - 16c_{\mathrm{W}}^2}{[(1 + c_{\mathrm{W}})^2 + 4 \times \frac{x_{k1}}{x_{l1}}]^2} > 0$$

24

if $c_{\text{W}} > 1$. Further, recall that player l has both higher cost ($c_{\text{W}} > 1$) and a lower continuation value ($P_k > P_l$) than player k, such that $x_{k1} > x_{l1}$ does hold. Therefore, assuming $x_{k1} = x_{l1}$ underestimates $F^*(1, c_{\text{W}})$. Since

$$f(c_{\text{W}}) = \frac{5c_{\text{W}}^3 + 2c_{\text{W}}^2 + c_{\text{W}}}{c_{\text{W}}^2 + 2c_{\text{W}} + 5}$$

is the expression we derive from $F^*(1, c_{\text{W}})$ under this assumption, we have proven $F^*(1, c_{\text{W}}) > f(c_{\text{W}})$. If we assume $c_{\text{W}} = 1$, all players are perfectly symmetric, such that $x_{k1} = x_{l1}$ does hold. Consequently, the relation $F^*(1, c_{\text{W}}) = f(c_{\text{W}})$ does hold for $c_{\text{W}} = 1$. $\square$

Lemma 2. *Assume without loss of generality that $c_{\text{W}} \geq c_{\text{S}} = 1$ and define $f_{\text{low}}(c_{\text{W}}) = 2c_{\text{W}} - 1$. Then, the relation $F^*(1, c_{\text{W}}) < f_{\text{low}}(c_{\text{W}})$ does hold for all $c_{\text{W}} > 1$. Furthermore, for $c_{\text{W}} = 1$, it holds that $f(c_{\text{W}}) = f_{\text{low}}(c_{\text{W}})$.*

Proof. We start with the relation that we want to prove, namely:

$$
\begin{aligned}
f(c_{\text{W}}) \quad &> \quad f_{\text{low}}(c_{\text{W}}) \\
\Leftrightarrow 5c_{\text{W}}^3 + 2c_{\text{W}}^2 + c_{\text{W}} \quad &> \quad (2c_{\text{W}} - 1)(c_{\text{W}}^2 + 2c_{\text{W}} + 5) \\
\Leftrightarrow 3c_{\text{W}}^3 - c_{\text{W}}^2 - 7c_{\text{W}} + 5 \quad &> \quad 0
\end{aligned}
$$

We now have to prove that $\phi(c_{\text{W}}) \equiv 3c_{\text{W}}^3 - c_{\text{W}}^2 - 7c_{\text{W}} + 5 > 0$ does always hold for $c_{\text{W}} > 1$. To see this, note that $\phi(\cdot)$ is a cubic function that has a local minimum at $c_{\text{W}} = 1$, and a local maximum at $c_{\text{W}} = -7/9$. Furthermore, $\phi(1) = 0$, which implies that $\phi(c_{\text{W}}) > 0$ for all $c_{\text{W}} > 1$. $\square$

Lemma 3. *Assume without loss of generality that $c_{\text{W}} \geq c_{\text{S}} = 1$ and define $f_{\text{high}}(c_{\text{W}}) = \frac{c_{\text{W}}^3 + 2c_{\text{W}}^2 + c_{\text{W}}}{4}$. Then, the relation $F^*(1, c_{\text{W}}) < f_{\text{high}}(c_{\text{W}})$ does hold for all $c_{\text{W}} > 1$. Furthermore, for $c_{\text{W}} = 1$, it holds that $F^*(1, c_{\text{W}}) = f_{\text{high}}(c_{\text{W}})$.*

Proof. From equation (B.5), we know that $\frac{x_{i1}}{x_{j1}} = \frac{c_{\text{W}}}{c_{\text{S}}} \frac{P_i(x_{k1}, x_{l1})}{P_j(x_{k1}, x_{l1})}$. Further, equation (B.7) tells us that $\frac{x_{i1}^*}{x_{j1}^*} = F^*(c_{\text{S}}, c_{\text{W}})$. Consequently, using the assumption that $c_{\text{W}} \geq c_{\text{S}} = 1$, it must hold that

$$F^*(1, c_{\text{W}}) = c_{\text{W}} \frac{P_i(x_{k1}, x_{l1})}{P_j(x_{k1}, x_{l1})} = \frac{4c_{\text{W}}^3 \times \frac{x_{l1}}{x_{k1}} + c_{\text{W}}(1 + c_{\text{W}})^2}{(1 + c_{\text{W}})^2 \times \frac{x_{l1}}{x_{k1}} + 4}.$$

Note that

$$\frac{\partial F^*(1, c_{\text{W}})}{\partial \frac{x_{l1}}{x_{k1}}} = -\frac{(c_{\text{W}} - 1)^2 c_{\text{W}}(c_{\text{W}}^2 + 6c_{\text{W}} + 1)}{[(1 + c_{\text{W}})^2 \times \frac{x_{l1}}{x_{k1}} + 4]^2} < 0$$

if $c_{\text{W}} > 1$. Further, recall from the main text that player l will never drop out in a pairwise competition for any finite degree of heterogeneity in terms of costs and continuation value, such

25

that $x_{l1} > 0$ does hold. Therefore, assuming $x_{l1} = 0$ (which implies $\frac{x_{l1}}{x_{k1}} = 0$) overestimates $F^*(1, c_{\text{W}})$, since this expression is decreasing in $\frac{x_{l1}}{x_{k1}}$. Since

$$f_{\text{high}}(c_{\text{W}}) = \frac{c_{\text{W}}^3 + 2c_{\text{W}}^2 + c_{\text{W}}}{4}$$

is the expression we derive from $F^*(1, c_{\text{W}})$ under this assumption, we have proven $F^*(1, c_{\text{W}}) < f_{\text{high}}(c_{\text{W}})$. If we assume $c_{\text{W}} = 1$, all players are perfectly symmetric, such that $x_{l1} = x_{j1}$ does hold. When inserting this relation in $F^*(1, c_{\text{W}})$, we see that the relation $F^*(1, c_{\text{W}}) = f(c_{\text{W}})$ does hold for $c_{\text{W}} = 1$. $\qquad\square$

Proof of Proposition 1

(a): To prove the relation $\mathcal{E}(\text{II}) > \mathcal{E}(\text{I})$ for all $c_{\text{W}} > c_{\text{S}}$, we assume without loss of generality that $c_{\text{W}} > c_{\text{S}} = 1$. Recall from (3) that $\mathcal{E}(\text{I})$ is defined stepwise, i.e., $\mathcal{E}(\text{I}) = \max\{\frac{3}{2+2c_{\text{W}}}P, \frac{1}{2}P\}$. First, we will consider the range $1 < c_{\text{W}} \leq 2$, where $\mathcal{E}(\text{I}) = \frac{3}{2+2c_{\text{W}}}P$. In the second part of this proof, we will devote attention to $c_{\text{W}} > 2$ and $\mathcal{E}(\text{I}) = \frac{1}{2}P$.

(i) We consider the range $1 < c_{\text{W}} \leq 2$ and want to prove that

$$
\begin{aligned}
\mathcal{E}(\text{I}) &< \mathcal{E}(\text{II}) \\
\Leftrightarrow \mathcal{E}(\text{I}) &< \frac{2}{3}\mathcal{E}(\text{SWSW}) + \frac{1}{3}\mathcal{E}(\text{SSWW}).
\end{aligned}
$$

Recall from the proof of Proposition 1 that the formal expression for $\mathcal{E}(\text{SWSW})$ is fairly complicated, in particular due to the $F^*(1, c_{\text{W}})$-function. To simplify the subsequent analysis, we will therefore make use again of Lemmata 1/2 and replace $F^*(1, c_{\text{W}})$ by $f_{\text{low}}(c_{\text{W}}) = 2c_{\text{W}} - 1$. This is without loss of generality, since $\mathcal{E}(\text{SWSW})$ is strictly increasing in $F^*(1, c_{\text{W}})$:

$$\frac{\partial \mathcal{E}(\text{SWSW})}{\partial F^*(1, c_{\text{W}})} = \frac{(2c_{\text{W}}^3 - c_{\text{W}}^2 - 4c_{\text{W}} + 7)F^*(1, c_{\text{W}}) + 3c_{\text{W}}^2 + 2c_{\text{W}} - 1}{2c_{\text{W}}(1 + c_{\text{W}})^2(1 + F^*(1, c_{\text{W}}))^3} > 0.$$

Note that the denominator is always greater than zero, since we know from Lemma 1 that (a) $\frac{\partial F^*(1, c_{\text{W}})}{\partial c_{\text{W}}} > 0$ and (b) $F^*(1) = 1$; this implies that the sign of the derivative is determined by the numerator, which is also greater than zero for all $1 < c_{\text{W}} \leq 2$. Consequently, effort $\mathcal{E}(\text{SWSW})$ is underestimated through the replacement of $F^*(1, c_{\text{W}})$ by $f_{\text{low}}(c_{\text{W}})$. Inserting $f_{\text{low}}(c_{\text{W}})$

and simplifying leaves us with the sufficient condition

$$Q(c_{\mathtt{W}}) \equiv \frac{(c_{\mathtt{W}} - 1)^2(6c_{\mathtt{W}}^3 + 2c_{\mathtt{W}}^2 - 9c_{\mathtt{W}} + 4)}{12c_{\mathtt{W}}^3(1 + c_{\mathtt{W}})^2} = \frac{(c_{\mathtt{W}} - 1)^2 q(c_{\mathtt{W}})}{12c_{\mathtt{W}}^3(1 + c_{\mathtt{W}})^2} > 0.$$

In the relevant range $1 < c_{\mathtt{W}} \leq 2$, the expression $(c_{\mathtt{W}} - 1)^2$ in the numerator as well as the denominator $12c_{\mathtt{W}}^3(1 + c_{\mathtt{W}})^2$ are always greater than zero, such that the sign of $Q(c_{\mathtt{W}})$ is determined by the expression $q(c_{\mathtt{W}}) \equiv 6c_{\mathtt{W}}^3 + 2c_{\mathtt{W}}^2 - 9c_{\mathtt{W}} + 4$. Note that $q(1) = 3$ and $q(2) = 42$. Since

$$\frac{\partial q(c_{\mathtt{W}})}{\partial c_{\mathtt{W}}} = 18c_{\mathtt{W}}^2 + 4c_{\mathtt{W}} - 9 > 0$$

for all $c_{\mathtt{W}} > 1$, it holds that $q(c_{\mathtt{W}}) > 0$ for all $1 < c_{\mathtt{W}} \leq 2$, which immediately implies that $Q(c_{\mathtt{W}}) > 0$. This completes the first part of the proof.

(ii) When $c_{\mathtt{W}} > 2$, it holds that $\mathcal{E}(\mathtt{I}) = \frac{1}{2}P$. We have to prove that the relation $\mathcal{E}(\mathtt{II}) > \mathcal{E}(\mathtt{I})$ is satisfied. Inserting the respective expressions for $\mathcal{E}(\mathtt{II})$ and $\mathcal{E}(\mathtt{I})$ gives the condition:

$$\frac{(3c_{\mathtt{W}}^3 + 6c_{\mathtt{W}}^2 + 4c_{\mathtt{W}} + 9)F^*(1, c_{\mathtt{W}})^2 + (2c_{\mathtt{W}}^3 + 14c_{\mathtt{W}}^2 + 16c_{\mathtt{W}} + 4)F^*(1, c_{\mathtt{W}}) + c_{\mathtt{W}}^3 + 4c_{\mathtt{W}}^2 + 6c_{\mathtt{W}} + 3}{3c_{\mathtt{W}}(c_{\mathtt{W}} + 1)^2(1 + F^*(1, c_{\mathtt{W}}))^2} > 1$$

$$\Leftrightarrow \underbrace{(c_{\mathtt{W}} + 9)F^*(1, c_{\mathtt{W}})^2 - (4c_{\mathtt{W}}^3 - 2c_{\mathtt{W}}^2 - 10c_{\mathtt{W}} - 4)F^*(1, c_{\mathtt{W}}) - (2c_{\mathtt{W}}^2 - 3)(c_{\mathtt{W}} + 1)}_{\equiv B(F^*(1,c_{\mathtt{W}}),c_{\mathtt{W}})} > 0$$

Note that $B(\cdot)$ is minimized for $F^*(1, c_{\mathtt{W}})_{\min} = \frac{2c_{\mathtt{W}}^3 - c_{\mathtt{W}}^2 - 5c_{\mathtt{W}} - 2}{9 + c_{\mathtt{W}}}$, since

$$\frac{\partial B(\cdot)}{\partial F^*(1, c_{\mathtt{W}})} = (18 + 2c_{\mathtt{W}})F^*(1, c_{\mathtt{W}}) - 4c_{\mathtt{W}}^3 + 2c_{\mathtt{W}}^2 + 10c_{\mathtt{W}} + 4 \quad \text{and} \quad \frac{\partial^2 B(\cdot)}{\partial[F^*(1, c_{\mathtt{W}})]^2} = 18 + 2c_{\mathtt{W}} > 0.$$

Moreover, note that $F^*(1, c_{\mathtt{W}})_{\min} > 0$ for all $c_{\mathtt{W}} > 2$, which implies that $B(\cdot)$ is increasing in $F^*(1, c_{\mathtt{W}})$ in the range which is relevant for this proof. Consequently, when solving the relation $B(\cdot) > 0$ for $F^*(1, c_{\mathtt{W}})$, we know that $F^*(1, c_{\mathtt{W}})$ must not be in the range between the two roots, as $B(\cdot)$ is negative here. We obtain

$$B(F^*(1, c_{\mathtt{W}}), c_{\mathtt{W}}) > 0$$

$$\Leftrightarrow F^*(1, c_{\mathtt{W}})^2 - \frac{(4c_{\mathtt{W}}^3 - 2c_{\mathtt{W}}^2 - 10c_{\mathtt{W}} - 4)}{(c_{\mathtt{W}} + 9)}F^*(1, c_{\mathtt{W}}) - \frac{(2c_{\mathtt{W}}^2 - 3)(c_{\mathtt{W}} + 1)}{(c_{\mathtt{W}} + 9)} > 0$$

$$\Leftrightarrow F^*(1, c_{\mathtt{W}}) < \frac{2c_{\mathtt{W}}^3 - c_{\mathtt{W}}^2 - 5c_{\mathtt{W}} - 2 - \sqrt{K(c_{\mathtt{W}})}}{9 + c_{\mathtt{W}}} \quad \vee \quad F^*(1, c_{\mathtt{W}}) > \frac{2c_{\mathtt{W}}^3 - c_{\mathtt{W}}^2 - 5c_{\mathtt{W}} - 2 + \sqrt{K(c_{\mathtt{W}})}}{9 + c_{\mathtt{W}}},$$

where

$$K(c_\mathtt{W}) = 4c_\mathtt{W}^6 - 4c_\mathtt{W}^5 - 17c_\mathtt{W}^4 + 22c_\mathtt{W}^3 + 44c_\mathtt{W}^2 - 10c_\mathtt{W} - 23.$$

We do only have to consider the second relation, since the first one is always below one for $c_\mathtt{W} > 2$, while $F^*(1, c_\mathtt{W}) \geq 1$ for all $c_\mathtt{W} \geq 1$.[28] To complete the proof, we have to show that

$$F^*(1, c_\mathtt{W}) > \frac{2c_\mathtt{W}^3 - c_\mathtt{W}^2 - 5c_\mathtt{W} - 2 + \sqrt{K(c_\mathtt{W})}}{9 + c_\mathtt{W}}$$

for all $c_\mathtt{W} > 2$. Inserting the equilibrium relation $F^*(1, c_\mathtt{W})$ from (B.7) gives:

$$\frac{(c_\mathtt{W} - 1)(1 + c_\mathtt{W})^2 + \sqrt{64c_\mathtt{W}^3 + (1 - c_\mathtt{W})^2(1 + c_\mathtt{W})^4}}{8} > \frac{2c_\mathtt{W}^3 - c_\mathtt{W}^2 - 5c_\mathtt{W} - 2 + \sqrt{K(c_\mathtt{W})}}{9 + c_\mathtt{W}}$$

Rearranging and simplifying gives the condition

$$H(c_\mathtt{W}) \equiv \underbrace{c_\mathtt{W}^4 - 6c_\mathtt{W}^3 + 16c_\mathtt{W}^2 + 30c_\mathtt{W} - 7}_{\mu(c_\mathtt{W})} + \underbrace{(9 + c_\mathtt{W})\sqrt{64c_\mathtt{W}^3 + (1 - c_\mathtt{W})^2(1 + c_\mathtt{W})^4}}_{\gamma(c_\mathtt{W})} - \underbrace{8\sqrt{K(c_\mathtt{W})}}_{\zeta(c_\mathtt{W})} > 0.$$

$H(c_\mathtt{W})$ consists of three parts $\mu(c_\mathtt{W})$, $\gamma(c_\mathtt{W})$, and $\zeta(c_\mathtt{W})$. Close inspection of $\mu(c_\mathtt{W})$ reveals that $\mu(c_\mathtt{W})$ is strictly increasing and greater than zero for all $c_\mathtt{W} > 2$.[29] Consequently, it is a sufficient condition for $H(c_\mathtt{W}) > 0$ to show that $\gamma(c_\mathtt{W}) > \zeta(c_\mathtt{W})$ in the range $c_\mathtt{W} > 2$:

$$\begin{aligned}
(9 + c_\mathtt{W})\sqrt{64c_\mathtt{W}^3 + (1 - c_\mathtt{W})^2(1 + c_\mathtt{W})^4} &> 8\sqrt{K(c_\mathtt{W})} \\
\Leftrightarrow (9 + c_\mathtt{W})^2[64c_\mathtt{W}^3 + (1 - c_\mathtt{W})^2(1 + c_\mathtt{W})^4] &> 64K(c_\mathtt{W}) \\
\Leftrightarrow c_\mathtt{W}^8 + 20c_\mathtt{W}^7 - 140c_\mathtt{W}^6 + 460c_\mathtt{W}^5 + 2086c_\mathtt{W}^4 + 3436c_\mathtt{W}^3 - 2860c_\mathtt{W}^2 + 820c_\mathtt{W} + 1553 &> 0
\end{aligned}$$

A sufficient condition for the above relation to hold is

$$\begin{aligned}
20c_\mathtt{W}^7 - 140c_\mathtt{W}^6 + 460c_\mathtt{W}^5 + 2086c_\mathtt{W}^4 + 3436c_\mathtt{W}^3 - 2860c_\mathtt{W}^2 &> 0 \\
\Leftrightarrow c_\mathtt{W}^2[20c_\mathtt{W}^5 - 140c_\mathtt{W}^4 + 460c_\mathtt{W}^3 + 2086c_\mathtt{W}^2 - 2860] &> 0.
\end{aligned}$$

Since $c_\mathtt{W} > 2$ by assumption, we are left with

$$20c_\mathtt{W}^5 - 140c_\mathtt{W}^4 + 460c_\mathtt{W}^3 + 2086c_\mathtt{W}^2 - 2860 > 0.$$

---

[28]Note that $F^*(1,1) = 1$; also, we know from Lemma 1 that $\frac{\partial F^*(1,c_\mathtt{W})}{\partial c_\mathtt{W}} > 0$. Therefore, $F^*(1, c_\mathtt{W}) \geq 1$ for all $c_\mathtt{W} \geq 1$.

[29]$\mu(2) = 85$, and $\mu'(c_\mathtt{W}) = 4c_\mathtt{W}^3 - 18c_\mathtt{W}^2 + 32c_\mathtt{W} + 30$.

For $c_W > 2$, it must hold that $2086c_W^2 - 2860 > 0$, such that we can drop those two expressions without loss of generality. We get

$$
\begin{aligned}
20c_W^5 - 140c_W^4 + 460c_W^3 &> 0 \\
\Leftrightarrow c_W^3[20c_W^2 - 140c_W + 460] &> 0 \\
\Leftrightarrow c_W^2 - 7c_W + 23 &> 0,
\end{aligned}
$$

which is greater than zero for all $c_W > 2$. This completes part (a) of the proof.

(b): Part (b) of Proposition 2 establishes that selection in setting SSWW is always dominated by selection in SSWW (the proof follows further below). Consequently, it is sufficient to show that $\mathcal{S}(\text{I}) > \mathcal{S}(\text{SWSW})$ to prove part (b) of Proposition 1, since $\mathcal{S}(\text{II})$ is a composite measure of $\mathcal{S}(\text{SWSW})$ and $\mathcal{S}(\text{SSWW})$. We start with the relation which we want to prove:

$$
\begin{aligned}
\mathcal{S}(\text{I}) &> \mathcal{S}(\text{SWSW}) \\
\Leftrightarrow \min\left\{\frac{2c_W - c_S}{c_S + c_W}, 1\right\} &> \frac{(c_S + c_W)F^*(1, c_W)^2 + 2c_W F^*(1, c_W)}{(c_S + c_W)[1 + F^*(1, c_W)]^2}
\end{aligned}
$$

Since $\mathcal{S}(\text{I})$ is defined stepwise, we have to proceed in two steps. First, we start with the case where $1 < c_W \leq 2$ such that $\mathcal{S}(\text{I}) = \frac{2c_W - c_S}{c_S + c_W}$, before we consider $c_W > 2$ and $\mathcal{S}(\text{I}) = 1$.

(i) We assume without loss of generality that $c_S = 1$ and consider the range $1 < c_W \leq 2$. Then, we get

$$
\begin{aligned}
\mathcal{S}(\text{I}) &> \mathcal{S}(\text{SWSW}) \\
\Leftrightarrow \frac{2c_W - 1}{1 + c_W} &> \frac{(1 + c_W)F^*(1, c_W)^2 + 2c_W F^*(1, c_W)}{(1 + c_W)[1 + F^*(1, c_W)]^2} \\
\Leftrightarrow (2c_W - 1)[1 + F^*(1, c_W)]^2 &> (1 + c_W)F^*(1, c_W)^2 + 2c_W F^*(1, c_W)
\end{aligned}
$$

Rearranging gives the condition

$$
N(c_W) = (c_W - 2)[F^*(1, c_W)]^2 + 2(c_W - 1)F^*(1, c_W) + 2c_W - 1 > 0
$$

Recall from equation (B.7) that the expression for $F^*(1, c_W)$ is fairly complicated. To simplify the subsequent analysis, we make use of Lemma 3, where we established that $F^*(1, c_W) < f_{\text{high}}(c_W)$ for all $c_W > 1$. Since $\mathcal{S}(\text{SWSW})$ is strictly increasing in $F^*(1, c_W)$, it is sufficient for the proof if we use the much simpler expression $f_{\text{high}}(c_W)$, as this tends to reduce the difference between the

one-stage and the two-stage contest in terms of selection:

$$\frac{\partial \mathcal{S}(\text{SWSW})}{\partial F^*(1, c_{\text{W}})} = \frac{2[c_{\text{W}} + F^*(1, c_{\text{W}})]}{(1 + c_{\text{W}})[1 + F^*(1, c_{\text{W}})]^3} > 0.$$

This leaves us with

$$
\begin{aligned}
\bar{N}(c_{\text{W}}) &= (c_{\text{W}} - 2)\left[\frac{c_{\text{W}}^3 + 2c_{\text{W}}^2 + c_{\text{W}}}{4}\right]^2 + 2(c_{\text{W}} - 1)\frac{c_{\text{W}}^3 + 2c_{\text{W}}^2 + c_{\text{W}}}{4} + 2c_{\text{W}} - 1 \\
&= \frac{(c_{\text{W}} - 2)[c_{\text{W}}^3 + 2c_{\text{W}}^2 + c_{\text{W}}]^2 + 8(c_{\text{W}} - 1)(c_{\text{W}}^3 + 2c_{\text{W}}^2 + c_{\text{W}}) + 32c_{\text{W}} - 16}{16} \\
&= \frac{(c_{\text{W}} - 1)[(c_{\text{W}} - 1)(c_{\text{W}} + 2)(c_{\text{W}}(c_{\text{W}} + 1)^2 + 4)c_{\text{W}} + 16]}{16}.
\end{aligned}
$$

Recall that we must show that $\bar{N}(c_{\text{W}}) > 0$ holds for all $1 < c_{\text{W}} \leq 2$. Note that $\bar{N}(1) = 0$ and $\bar{N}(2) = 12$. Therefore, the proof is complete if we can show that $\bar{N}(c_{\text{W}})$ is strictly increasing in the relevant range. Since $c_{\text{W}} > 1$, the factor $(c_{\text{W}} - 1)$ in the expression of $\bar{N}(c_{\text{W}})$ is always positive and can be disregarded in the subsequent analysis of the slope. Subsequently, we use the simpler expression

$$\hat{N}(c_{\text{W}}) = \frac{(c_{\text{W}} - 1)(c_{\text{W}} + 2)(c_{\text{W}}(c_{\text{W}} + 1)^2 + 4)c_{\text{W}} + 16}{16}.$$

When computing the first derivative of $\bar{N}(c_{\text{W}})$ with respect to c_{W}, we obtain

$$\frac{\partial \hat{N}(c_{\text{W}})}{\partial c_{\text{W}}} = \frac{6c_{\text{W}}^5 + 15c_{\text{W}}^4 + 4c_{\text{W}}^3 + 3c_{\text{W}}^2 + 4c_{\text{W}} - 8}{16},$$

which is clearly positive for all values in the range $1 < c_{\text{W}} \leq 2$. This proves the first part of the Proposition.

(ii) We assume without loss of generality that $c_{\text{S}} = 1$. Then, a comparison of $\mathcal{S}(\text{I})$ and $\mathcal{S}(\text{SWSW})$ in the range $c_{\text{W}} > 2$ gives

$$
\begin{aligned}
\mathcal{S}(\text{I}) &> \mathcal{S}(\text{SWSW}) \\
\Leftrightarrow \quad 1 &> \frac{(1 + c_{\text{W}})F^*(1, c_{\text{W}})^2 + 2c_{\text{W}}F^*(1, c_{\text{W}})}{(1 + c_{\text{W}})[1 + F^*(1, c_{\text{W}})]^2} \\
\Leftrightarrow (1 + c_{\text{W}})[1 + F^*(1, c_{\text{W}})]^2 &> (1 + c_{\text{W}})F^*(1, c_{\text{W}})^2 + 2c_{\text{W}}F^*(1, c_{\text{W}})
\end{aligned}
$$

Rearranging gives the condition

$$M(c_\mathtt{w}) = 2F^*(1, c_\mathtt{w}) + c_\mathtt{w} + 1 > 0.$$

As in the first part of this proof, we substitute $f_{\text{high}}(c_\mathtt{w})$ for $F^*(1, c_\mathtt{w})$, which gives

$$
\begin{aligned}
\bar{M}(c_\mathtt{w}) &= 2\frac{c_\mathtt{w}^3 + 2c_\mathtt{w}^2 + c_\mathtt{w}}{4} + c_\mathtt{w} + 1 \\
&= \frac{c_\mathtt{w}^3 + 2c_\mathtt{w}^2 + 3c_\mathtt{w} + 2}{2}.
\end{aligned}
$$

$\bar{M}(c_\mathtt{w})$ is clearly positive for all $c_\mathtt{w} > 2$, which proves the second part of the Proposition.

Proof of Proposition 2

We will separately prove parts (a) and (b) of Proposition 2. We start with part (a) below.

(a): To prove the relation $\mathcal{E}(\mathtt{SSWW}) > \mathcal{E}(\mathtt{SWSW})$ for all $c_\mathtt{w} > c_\mathtt{s}$, we assume without loss of generality that $c_\mathtt{w} > c_\mathtt{s} = 1$. In the proof, we will proceed in two steps. First, we derive a necessary and sufficient condition in terms of the function $F^*(1, c_\mathtt{w})$ for the relation $\mathcal{E}(\mathtt{SSWW}) > \mathcal{E}(\mathtt{SWSW})$ to hold. Second, we prove that the equilibrium function $F^*(1, c_\mathtt{w})$, which was derived in (B.7), indeed satisfies this condition. We start with the relation which we want to prove:

$$\mathcal{E}(\mathtt{SSWW}) > \mathcal{E}(\mathtt{SWSW})$$

$$\Leftrightarrow \frac{c_\mathtt{w}^3 + 2c_\mathtt{w}(1 + c_\mathtt{w}) + 1}{2c_\mathtt{w}(1 + c_\mathtt{w})^2} > \frac{(1 + c_\mathtt{w})^2[1 + [1 + F^*(\cdot)]c_\mathtt{w}F^*(\cdot)] + 4c_\mathtt{w}[c_\mathtt{w}^2 + (1 + c_\mathtt{w})F^*(\cdot)]}{2c_\mathtt{w}(1 + c_\mathtt{w})^2[1 + F^*(\cdot)]^2}$$

Multiplying both sides by $2c_\mathtt{w}(1 + c_\mathtt{w})^2[1 + F^*(1, c_\mathtt{w})]^2$ and rearranging delivers

$$F^*(1, c_\mathtt{w})^2 + \frac{c_\mathtt{w}^3 - 2c_\mathtt{w}^2 - c_\mathtt{w} + 2}{c_\mathtt{w} + 1}F^*(1, c_\mathtt{w}) - \frac{3c_\mathtt{w}^3 - c_\mathtt{w}^2}{c_\mathtt{w} + 1} > 0$$

Solving for $F^*(1, c_\mathtt{w})$ gives us two conditions:

$$F^*(1, c_\mathtt{w}) < \frac{-c_\mathtt{w}^3 + 2c_\mathtt{w}^2 + c_\mathtt{w} - 2 - R(c_\mathtt{w})}{2c_\mathtt{w} + 2} \quad \vee \quad F^*(1, c_\mathtt{w}) > \frac{-c_\mathtt{w}^3 + 2c_\mathtt{w}^2 + c_\mathtt{w} - 2 + R(c_\mathtt{w})}{2c_\mathtt{w} + 2} \equiv Z(c_\mathtt{w}),$$

where

$$R(c_\mathtt{w}) = \sqrt{c_\mathtt{w}^6 - 4c_\mathtt{w}^5 + 14c_\mathtt{w}^4 + 16c_\mathtt{w}^3 - 11c_\mathtt{w}^2 - 4c_\mathtt{w} + 4}.$$

We do only have to consider the second relation, since the first one is below one for some values of $c_{\mathtt{W}}$, while $F^*(1, c_{\mathtt{W}}) \geq 1$ for all $c_{\mathtt{W}} \geq 1$.[30] This completes the first part of the proof. We now have to prove that

$$F^*(1, c_{\mathtt{W}}) > Z(c_{\mathtt{W}}) \equiv \frac{-c_{\mathtt{W}}^3 + 2c_{\mathtt{W}}^2 + c_{\mathtt{W}} - 2 + R(c_{\mathtt{W}})}{2c_{\mathtt{W}} + 2} \tag{C.1}$$

for all $c_{\mathtt{W}} > 1$. From Lemmata 1 and 2 we know that $F^*(1, c_{\mathtt{W}}) > f_{\text{low}}(c_{\mathtt{W}})$. Consequently, a sufficient condition for (C.1) is given by $f_{\text{low}}(c_{\mathtt{W}}) > Z(c_{\mathtt{W}})$. Rearranging this condition gives

$$c_{\mathtt{W}}^3 + 2c_{\mathtt{W}}^2 + c_{\mathtt{W}} > R(c_{\mathtt{W}}).$$

Squaring both sides leaves us with[31]

$$2c_{\mathtt{W}}^5 - 2c_{\mathtt{W}}^4 - 3c_{\mathtt{W}}^3 + 3c_{\mathtt{W}}^2 + c_{\mathtt{W}} - 1 > 0$$
$$\Leftrightarrow 2(c_{\mathtt{W}} - 1)^2 (c_{\mathtt{W}} + 1)(c_{\mathtt{W}} - \frac{1}{\sqrt{2}})(c_{\mathtt{W}} + \frac{1}{\sqrt{2}}) > 0.$$

This relation is always satisfied if $c_{\mathtt{W}} > 1$, which completes part (a) of this proof.

(b): In part (b) of this proof, we first derive a necessary and sufficient condition which assures that the relation $\mathcal{S}(\mathtt{SSWW}) < \mathcal{S}(\mathtt{SWSW})$ does hold in terms of the function $F^*(1, c_{\mathtt{W}})$. Then, we prove that the equilibrium function $F^*(1, c_{\mathtt{W}})$ satisfies this condition.

(i) As previously, we assume that $c_{\mathtt{W}} > c_{\mathtt{S}} = 1$ does hold without loss of generality. Consequently, we can use the expressions in equations (6) and (8) in what follows. We start with the relation which we want to prove:

$$\mathcal{S}(\mathtt{SWSW}) > \mathcal{S}(\mathtt{SSWW})$$
$$\Leftrightarrow (1 + c_{\mathtt{W}})F^*(1, c_{\mathtt{W}})^2 + 2c_{\mathtt{W}}F^*(1, c_{\mathtt{W}}) > c_{\mathtt{W}}F^*(1, c_{\mathtt{W}})^2 + 2c_{\mathtt{W}}F^*(1, c_{\mathtt{W}}) + c_{\mathtt{W}}$$
$$\Leftrightarrow F^*(1, c_{\mathtt{W}})^2 > c_{\mathtt{W}}$$
$$\Leftrightarrow F^*(1, c_{\mathtt{W}}) < -\sqrt{c_{\mathtt{W}}} \quad \vee \quad F^*(1, c_{\mathtt{W}}) > \sqrt{c_{\mathtt{W}}}$$

Note that it is sufficient to show that $F^*(1, c_{\mathtt{W}}) > c_{\mathtt{W}}$, since $c_{\mathtt{W}} > \sqrt{c_{\mathtt{W}}}$ for $c_{\mathtt{W}} > 1$.

(ii) From Lemma 1, we know that $F^*(1, c_{\mathtt{W}}) > f(c_{\mathtt{W}})$. We will now prove that $f(c_{\mathtt{W}}) > c_{\mathtt{W}}$ for

---

[30]Note that $F^*(1,1) = 1$; also, we know from Lemma 1 that $\frac{\partial F^*(1, c_{\mathtt{W}})}{\partial c_{\mathtt{W}}} > 0$. Therefore, $F^*(1, c_{\mathtt{W}}) \geq 1$ for all $c_{\mathtt{W}} \geq 1$.

[31]Note that squaring is without loss of generality here, since we are only interested in solutions for $c_{\mathtt{W}} > 1$.

$c_{\texttt{W}} > 1$ to complete the proof. $f(c_{\texttt{W}}) > c_{\texttt{W}}$ implies that

$$\frac{5c_{\texttt{W}}^3 + 2c_{\texttt{W}}^2 + c_{\texttt{W}}}{c_{\texttt{W}}^2 + 2c_{\texttt{W}} + 5} > c_{\texttt{W}}$$

does hold. Rearranging gives

$$
\begin{aligned}
5c_{\texttt{W}}^3 + 2c_{\texttt{W}}^2 + c_{\texttt{W}} \quad &> \quad c_{\texttt{W}}^3 + 2c_{\texttt{W}}^2 + 5c_{\texttt{W}} \\
\Leftrightarrow c_{\texttt{W}}(c_{\texttt{W}}^2 - 1) \quad &> \quad 0 \\
\Leftrightarrow c_{\texttt{W}} > 1 \quad &\vee \quad -1 < c_{\texttt{W}} < 0
\end{aligned}
$$

This proves the claim $\mathcal{S}(\texttt{SWSW}) > \mathcal{S}(\texttt{SSWW})$ for all $c_{\texttt{W}} > 1$.