Collusion under Yardstick Competition

Potters, J.J.M.; Rockenbach, B.; Sadrieh, A.; van Damme, E.E.C.

COLLUSION UNDER YARDSTICK COMPETITION
AN EXPERIMENTAL STUDY

By J.J.M. Potters, B. Rockenbach, A. Sadrieh,
E.E.C. van Damme

TILBURG ◆ UNIVERSITY

Collusion under Yardstick Competition
An Experimental Study

Jan Potters*

Bettina Rockenbach**

Abdolkarim Sadrieh*

and

Eric van Damme*

* Tilburg University

** University of Erfurt

January 2003

Abstract

The effectiveness of relative performance evaluation schemes, such as yardstick competition, can be undermined by collusion. The degree to which the regulated agents manage to collude will be affected by the particulars of the scheme. We hypothesize that in a repeated game setting schemes will be more prone to collusion the smaller are the rents to the agents in case they behave non-cooperatively. We illustrate the relevance of this hypothesis by means of an economic experiment in which we compare the efficiency of two performance evaluation schemes.

Corresponding author

Jan Potters, Department of Economics, Tilburg University, PO Box 90153, 5000 LE Tilburg, The Netherlands. Phone +31 13 466 8204. Fax +31 13 466 3042. Email J.J.M.Potters@uvt.nl

1. Introduction

The benefits of yardstick competition[1] in particular, and relative performance evaluation in general, can only be reaped if the agents act non-cooperatively. As Shleifer (1985) notes "an important limitation of yardstick competition is its susceptibility to collusive manipulation" (p. 327).[2] By resisting the temptation of short term gains from the relative performance competition, colluding agents can prevent being played out against each other. This will usually allow them to get away with inefficiently low effort levels and a larger than first-best share of the rents. The attractiveness of a particular yardstick scheme to the principal or regulator will thus depend, at least partly, on the possibility to prevent collusive conduct.

Different schemes have different incentive properties. These properties will not only affect the outcome under non-cooperative behavior, they may also affect the likelihood and the degree of collusion between agents. The possibility that we explore in this paper is that there is a trade-off between the static efficiency properties of a particular yardstick scheme and the degree to which it prevents collusion. Possibly, a scheme that has very desirable properties if we assume that the agents act non-cooperatively will do much worse if we consider the potential for collusion.

We explore this possibility in an environment in which two agents interact repeatedly under yardstick competition.[3] Two different schemes are considered: a *discriminatory* scheme, in which each agent has a different yardstick, and a *uniform* scheme with the same yardstick for both agents. The non-cooperative equilibrium of the discriminatory scheme attains the first-best outcome, while the non-cooperative equilibrium of the uniform scheme involves substantial efficiency losses. At the same time, the agents earn a much smaller rent under the discriminatory scheme than under the uniform scheme. Relative to the non-cooperative outcome they have more to gain from colluding in the former than in the latter scheme. This

---

[1] The term *yardstick competition* was introduced by Shleifer (1985). Related models and ideas on relative performance evaluation are by Baiman and Demski (1980), Lazear and Rosen (1981), Holmstrom (1982), Nalebuff and Stiglitz (1983).

[2] The scope for collusion under yardstick competition and related incentive schemes is widely acknowledged (e.g., Dye 1984; Gibbons and Murphy 1990; Mookherjee 1984).

[3] Laffont and Martimort (1997, 1998) and Tangeras (1999) study collusion in a static model, and assume that the agents can write and commit to a collusive side contract. We focus on "tacit collusion" which is supported not by contracts but by repeated game strategies.

may, but does not necessarily, increase the scope for collusion under the discriminatory scheme relative to the uniform scheme.

The two proposed schemes are more than a theoretical curiosity and the possibility of a significant difference in their incentive properties has been recognized by regulatory agencies. For example, the Netherlands' Office of Energy Regulation (Dte) compares several yardstick schemes which mainly vary in the degree to which a firm's own performance affects its own yardstick. The uniform and the discriminative scheme are among the alternatives considered. The regulator acknowledges that the incentive properties are improved if a firm's own performance does not affect its own yardstick. "This system removes the suppliers' individual incentives for manipulation or misuse" (Dte, 2001, p. 71). At the same time, the report acknowledges that "as a group, the suppliers still have incentives for manipulation or misuse".

Collusion is a pertinent possibility in environments in which there is limited number of agents and repeated interaction. Such repeated game settings are often characterized by multiple equilibria of which some may be highly collusive. In fact, it is well-known that under certain conditions collusion can be supported by a subgame perfect equilibrium of the repeated game, even if no binding agreements are possible (Fudenberg and Tirole, 1991). Hence, it seems difficult to assess the scope for collusion purely on the basis of theoretical arguments.

In the present paper, we follow a long tradition of searching for "behaviorally relevant" outcomes of a repeated game by conducting a laboratory experiment (see e.g. Roth 1995). Interestingly, the repeated game effects in settings that allow both competitive and collusive outcomes have neither uniformly supported collusion nor competition. In Cournot oligopolies, for example, repeated interaction has been found to lead both to collusive outcomes (e.g. Fouraker and Siegel 1963; Selten, Mitzkewitz, and Uhlich 1997) and to competitive outcomes (e.g. Huck, Normann, and Oechssler 1999 and 2000; Offerman, Potters, and Sonnemans 2002). Important factors are the number of oligopolists and the type of information feedback (see Holt 1995 for an overview). Similarly, in finitely repeated prisoner's dilemma games, cooperation frequently dominates initial play, but breaks down towards the end (Selten and Stoecker 1986; Roth 1995). Also, repeated interaction has a positive influence on cooperation in public goods games, but the extent of cooperation generally decreases towards the end of the repeated game (see Croson 1996 and Keser and van Winden 2000 and the literature cited therein). Finally, the results of Clark and Sefton (2001) show that repeated game effects can effectively help subjects to coordinate on an efficient, but risky equilibrium.

2

Although our experimental design shares some features with a number of earlier studies, it does not fit well into any of the categories of experiments reported so far. We know of no previous experiment in which independent decision makers are linked together via a "virtual competition" scheme such as the yardstick competition suggested by Shliefer (1985). Our decision makers can be thought of as symmetric local monopolists who independently choose their cost and price levels while facing separate local demand.[4] Their payoffs are "regulated" and linked together via one of the two yardstick competition schemes discussed above.

Our experimental results indicate that the discriminatory yardstick is much more prone to collusion than the uniform scheme. We find that cost choices under the discriminatory regime are significantly further away from equilibrium than in the uniform case. Moreover, the discriminatory scheme does not only fare worse than the uniform scheme relative to equilibrium, but also in absolute terms. We find significantly more cases of perfect collusion and significantly lower total welfare under the discriminatory than under the uniform yardstick regulation.

In the next section we present the theoretical model, which is an adapted version of Shleifer's (1985) model of yardstick competition. Sections 3 and 4 outline our experimental design. The results are presented in Section 5. Finally, Section 6 concludes.

2. The Model

The model we implement in the experiment is an adjusted version of Shleifer's (1985) model of yardstick competition. There are two symmetric agents (firms). Each agent acts as a local monopolist facing a downward-sloping local demand function $q(.)$, which is assumed to be identical for both. Each agent i decides on a marginal cost c_i and a price p_i. Firm i's profit is $(p_i - c_i)q(p_i)$. In addition to profits the agent receives a *benefit of slack, $B(c_i)$,* which is assumed to be increasing in c_i at a decreasing rate ($B'>0$ and $B''<0$). This benefit for slack

---

[4] A number of different interpretations of our model are also possible. For example, the decision makers can also be thought of as managers running independent but comparable business units within a company, or lawyers working on independent but comparable cases for a company, and so forth.

represents the disutility to the agent of keeping costs low. It is equivalent to agents' cost of cost reduction in Shleifer's (1985) original model.[5] Thus, the total payoff V_i to agent i is:

$$V_i(c_i, p_i) = (p_i - c_i)q(p_i) + B(c_i) \qquad (1)$$

Social welfare S_i in the local market of the agent i is defined as the sum of the consumer surplus and agent's payoff:

$$S_i(c_i, p_i) = \int_{p_i}^{\infty} q(x)dx + (p_i - c_i)q(p_i) + B(c_i) \qquad (2)$$

The social optimum in each market i is specified by the first order conditions: [6]

$$p_i = c_i \qquad (3)$$

$$B'(c_i) = q(p_i) \qquad (4)$$

Let $p^* = c^*$ denote the optimal level of price and cost. It is assumed that the principal (regulator) is not informed about the functions $q(.)$ and $B(.)$. Hence, simply commanding the optimum defined by (3) and (4) is not an option. The regulator can observe the realized levels of c_i, however, and set restrictions on prices that are conditioned on the cost levels. The restriction for the price p_i of firm i is called the *price-cap* $\hat{p}_i$ and defines the maximum price that firm i may charge.

There are two basic "non-competitive" regulatory schemes. Under the *cost-of-service regulation*, the regulator sets a price-cap for each firm i that is strictly increasing in the firm's own cost level. The simplest form is $\hat{p}_i(c_i) = c_i$. The profits to the firm are thus restricted. However, the incentives for the firm to reduce its cost of production are undermined, because of the negative effect of cost reduction on the price-cap. This typically leads to cost levels above the social optimum. Under *price-cap regulation*, the price-cap $\hat{p}_i$ is a constant which is independent of the realized cost levels. This type of regulation leads to inefficient outcomes, unless the regulator just happens to fix the price cap at the socially optimal price, i.e. $\hat{p}_i =$

---

[5] This is the only difference between Shleifer's original model and our adjusted model. The reason for our adjustment is that a model with a "benefit of slack" enables us to track the agents' incentives to collude under yardstick competition. This is more difficult in Shleifer's original model. The important feature in both models though is that the agents have an incentive to be cost inefficient.

[6] Second order conditions are $q''(p^*)p^* > q'(p^*)$ and $B''(c^*) < q'(p^*)$.

p^*. If the price-cap is set too low, there are strong incentives to reduce cost but production levels will be suboptimal. If the price-cap is too high, firms reap above-optimal profits at above-optimal cost levels.

As Shleifer (1985) shows, creating virtual competition between the local monopolies can alleviate the incentive problems. The *discriminatory yardstick scheme* proposed by Shleifer is a price-cap regulation that induces the socially optimum price and cost levels in the unique non-cooperative equilibrium of the game. In its simplest form, the regulator sets the price-cap of each firm equal to the marginal cost of the other firm:[7]

$$\hat{p}_1(c_1, c_2) = c_2 \text{ and } \hat{p}_2(c_1, c_2) = c_1 \tag{5}$$

The regulation is discriminatory because the price caps are firm-specific. Firm 1 is allowed to charge a price at most equal to the cost level realized by firm 2 and vice versa. If a firm reduces its production cost, it does not affect its own price-cap (negatively), but does reduce the price-cap of the competing firm. This gives both firms the proper incentives to reduce costs. At the same time, every cost reduction by a firm reduces the profit levels of the other firm. Thus, under discriminatory yardstick regulation (5) both firms' profits are at socially optimal level. Specifically, with symmetric firms the price-cap and the prices charged in equilibrium are equal to the optimal marginal cost, i.e. $\hat{p}_i = p^* = c^*$. The equilibrium payoff V_i^N of each agent i consists only of the benefit of slack at the socially optimal cost level c^*:

$$V_i^N \equiv V_i(c^*, c^*) = B_i(c^*) \tag{6}$$

An alternative to the discriminatory yardstick scheme is the *uniform yardstick scheme*. In this type of regulation all firms face the same price-cap, equal to the average of firms' marginal cost levels:

$$\hat{p}_i(c_1, c_2) = \tfrac{1}{2}(c_1 + c_2), \text{ for } i = 1,2 \tag{7}$$

It is easy to see that the non-cooperative equilibrium under the uniform scheme does not coincide with the social optimum. Remember that one of the first-order conditions for a social optimum (4) is that the agent's marginal benefit of an increase in the cost level $B'(c_i)$ should

---

[7] Shleifer shows that the results on the discriminatory price-cap regulation basically go through for any price-cap scheme that only depends on the cost level of the other firm and additionally satisfies the conditions $\partial \hat{p}_i / \partial c_j \geq 0$ and $\hat{p}_i(c_j) = c_j$ if $c_j = c^*$.

be equal to the consumers' marginal benefit of a decrease in the price $q(p_i)$. This condition holds true in the equilibrium of the game with discriminatory yardstick regulation, because (4) is identical to the first order condition for agent i when $\hat{p}_i = c_j$. Condition (4) does not hold in the equilibrium of a game with uniform yardstick regulation. At the social optimum, $p_i = c_i = c*$, for $i = 1, 2$, agent i's payoff is increasing in c_i:

$$\left.\frac{\partial V_i}{\partial c_i}\right|_{c_i = c_j = c*} = -\tfrac{1}{2}q(c*) + B'(c*) > 0 \tag{8}$$

The agents have an incentive to increase cost levels above the socially optimal level. The important difference between the uniform scheme and the discriminatory scheme is that in the former an agent's cost positively affects the own price-cap while it does not in the latter. Although the uniform yardstick scheme induces some virtual competition and enhances efficiency compared to the cost-of-service regulation, it does not completely solve the incentive problem as the discriminatory yardstick scheme does.

In many of the empirically relevant cases agents are interacting repeatedly. The question we briefly consider now is whether there are theoretical reasons to expect that the incentives to collude will be different under the discriminatory and the uniform yardstick schemes. The most frequently used theoretical measure for the incentives to collude in a repeated game is based on the idea of *grim trigger strategies* introduced by Friedman (1971). The measure is the threshold discount rate $\bar{r}$ (i.e. the maximum discount rate that can support a subgame perfect equilibrium with trigger strategies in which all players cooperate in all stages):

$$\bar{r} = \frac{V^C - V^N}{V^D - V^C} \tag{9}$$

where V^N is the stage game payoff in the non-cooperative equilibrium, V^C is the stage game payoff under collusion (joint payoff maximization), and V^D is the stage game payoff of the best reply to collusion (defect).

Under quite general assumptions, it can be shown that the numerator in (9) is greater for threshold discount rate of the discriminatory scheme $\bar{r}^{DIS}$ than for threshold discount rate of the uniform scheme $\bar{r}^{UNI}$. The cooperative payoff V^C is the same in both cases, but the Nash equilibrium payoff under the uniform scheme V^{N-UNI} is generally greater than that under the

discriminatory scheme V^{N-DIS}, because the equilibrium cost level in the former case c^{N-UNI} is greater than the equilibrium cost level in the latter case, i.e., $c^{N-UNI} > c^{N-DIS} = c*$.

However, the denominator in (9) is smaller for $\bar{r}^{DIS}$ than for $\bar{r}^{UNI}$. The reason is that the defection payoff under the uniform scheme V^{D-UNI} is generally smaller than the defection payoff under the discriminatory scheme V^{D-DIS}, because reducing the cost (i.e. defecting from collusion) has a negative influence on the own price-cap in the former case while does not in the latter case.

Taking the two effects together does not allow a definite prediction on the relationship of $\bar{r}^{DIS}$ and $\bar{r}^{UNI}$. Since V^C is the same under both schemes, ceteris paribus, $V^{N-UNI} > V^{N-DIS}$ results in $\bar{r}^{UNI} < \bar{r}^{DIS}$, while $V^{D-UNI} < V^{D-DIS}$ results in $\bar{r}^{UNI} > \bar{r}^{DIS}$. It can be shown, however, that in a linear version of the model (i.e., $q(.)$ linear and $B'(.)$ linear) the first effect always dominates, meaning that there are theoretical reasons to expect more collusive behavior under discriminatory than under uniform yardstick regulation. In the next section, we illustrate this for the parameterization of the model used in our experiment.

3. Parameterization Used in the Experiment

The demand function for each monopoly i in our experiment is $q_i(p_i) = 34 - \frac{1}{2}p_i$, for $i = 1, 2$. The benefit of slack is $B(c_i) = 40c_i - c_i^2$. This parameterization leads to the theoretical predictions that are summarized in table 1.

First, consider the social optimum. Using the first order conditions (3) and (4), we have $c_i* = c* = 4$, $p_i* = p* = 4$ and $V_i* = V* = 144$ for each monopoly i in the social optimum. Obviously, the social optimum is not affected by the regulation scheme used.

As explained, the social optimum coincides with the non-cooperative equilibrium of the game with the discriminatory pricing rule (5). Inserting the experimental parameters into the first order condition (4) gives the symmetric best response functions $c_i^{N-DIS} = 3 + \frac{1}{4}c_j^{N-DIS}$. It is easily verified in equilibrium we have $c_i^{N-DIS} = c* = 4$, $p_i^{N-DIS} = p* = 4$, and $V_i^{N-DIS} = V* = 144$, for $i = 1, 2$.

The non-cooperative equilibrium of the game with a uniform yardstick scheme has cost levels above the social optimum. From the first order condition in (8) and the parameters of the

experiment we can derive the best response functions under a uniform yardstick scheme: $c_i^{N-UNI} = 12\frac{20}{29} + \frac{1}{29}c_j^{N-UNI}$. Solving for equilibrium, we find that the firms produce at higher marginal cost levels, $c_i^{N-Uni} = 13.1$, and make higher profits, $V_i^{N-UNI} = 353$, than is socially optimal.

The fourth column of table 1 displays choices and outcomes for the case that the monopolists engage in collusive symmetric joint payoff maximization. If both firms choose the marginal cost level of $c_i^C = 20$, they achieve their maximum symmetric payoffs $V_i^C = 400$. This can be achieved no matter which yardstick scheme is implemented, because collusion completely overrides the yardstick competition in either case.

Given that the monopolist j plays the collusive cost choice $c_j^C = 20$, the other monopolist i has an incentive to defect and play a best-reply to this cost choice. The fifth column of table 1 shows the choices and outcomes for the defecting firm. Note that unlike the other cases presented in the table, the defection outcomes are based on an asymmetric strategy combination, in which one player chooses the collusion cost level, while the other chooses the best response to it. This means that the price chosen by the defecting monopolist is greater than the marginal cost, because the price-cap is influenced by the other monopolist's choice of cost. In the case of the discriminatory yardstick regulation, the defecting monopolist chooses a price equal to the price-cap which in turn is exactly equal to the colluding monopolist's marginal cost, i.e. $p_i^{D-DIS} = c_j^C = 20$. Given this price, the defecting monopolist sets the marginal cost to $c_i^{D-DIS} = 8$ and receives the maximal payoff of $V_i^{D-DIS} = 544$.

The payoff of defection in the uniform yardstick regime is smaller than in the discriminatory regime ($V_i^{D-UNI} = 441$). The reason that defection in the uniform regime has a negative payoff effect that defection in the discriminatory regime does not have. Lowering the cost below the collusive level leads to an immediate downwards adjustment of the price-cap and, thus, to a loss of revenue due to the lower price. As shown in table 1, the optimal trade-off between the positive cost reduction effect and the negative revenue reduction effect leads to a defector's cost choice of $c_i^{D-UNI} = 13.4$. This in turn means that the price-cap and the price are adjusted downwards to $p_i^{D-UNI} = \frac{1}{2}(c_i^{D-UNI} + c_j^C) = 16.7$.

Finally, the last column of table 1 displays the collusion threshold discount rates for the two different yardstick schemes. Using the three payoffs, V^N, V^C, and V^D, these thresholds can be

calculated according to (9). Since $\bar{r}^{UNI} < \bar{r}^{DIS}$, we hypothesize that collusion is more likely to be observed under the discriminatory scheme than under the uniform scheme. Although the temptation to defect is higher under the discriminatory scheme than under the uniform scheme ($V^{D-DIS} - V^C > V^{D-UNI} - V^C$), this is more than offset by the much severer punishment of defection under the discriminatory scheme ($V^C - V^{N-DIS} > V^C - V^{N-UNI}$).

Table 1. Theoretical predictions for the parameters used in the experiment

price-cap scheme	Social Optimum	Non-cooperative symmetric Nash equilibrium	Collusion joint payoff maximization	Defection best response to collusion	Collusion Threshold Discount Rate
discriminatory yardstick	$p_i = p* = 4$	$p_i^{N-DIS} = 4$	$p_i^C = 20$	$p_i^{D-DIS} = 20$	
$\hat{p}_i = c_j$	$c_i = c* = 4$	$c_i^{N-DIS} = 4$	$c_i^C = 20$	$c_i^{D-DIS} = 8$	$\bar{r}^{DIS} = 1.78$
with $j \neq i$	$V_i = V* = 144$	$V_i^{N-DIS} = 144$	$V_i^C = 400$	$V_i^{D-DIS} = 544$	
uniform yardstick	$p_i = p* = 4$	$p_i^{N-UNI} = 13.1$	$p_i^C = 20$	$p_i^{D-UNI} = 16.7$	
$\hat{p}_i = \frac{1}{2}(c_i + c_j)$	$c_i = c* = 4$	$c_i^{N-UNI} = 13.1$	$c_i^C = 20$	$c_i^{D-UNI} = 13.4$	$\bar{r}^{UNI} = 1.14$
	$V_i = V* = 144$	$V_i^{N-UNI} = 353$	$V_i^C = 400$	$V_i^{D-UNI} = 441$	

The example of our parameter setup illustrates that the unambiguous efficiency ranking of the two schemes in the one-shot game is lost in the repeated game version of the model. It is possible that the scheme that performs better in the one-shot game, performs worse in the repeated game setting. Whether or not this is a relevant possibility is difficult to decide on the basis of theoretical reasoning. As is usually the case, the repeated game is characterized by a multiplicity of equilibria. There is a larger range of discount rates for which collusion is an equilibrium under the discriminatory scheme than under the uniform scheme (at least in the linear version of the model). This does not imply, however, that collusion will actually occur more often. After all, a repetition of the non-cooperative outcome is also an equilibrium of the repeated game. Even outcomes in between the non-cooperative and the cooperative outcome are equilibria of the repeated game.

4. Experimental Design

The experiment was conducted at the CentER*lab* at the Tilburg University. The subjects were recruited for voluntary participation on campus. All subjects were undergraduate students, most of them studying business or economics. Each subject was allowed to participate only once and none had participated in a similar experiment before. The experiment was computerized using the experimental software toolbox RatImage (Abbink and Sadrieh 1995).

The experiment consisted of two treatments, a discriminatory yardstick treatment (DIS) and a uniform yardstick (UNI) treatment. In both cases, the subjects were randomly matched in pairs at the beginning of the experimental session and then played for 50 consecutive rounds. Subject were informed about this matching protocol, and about the number of rounds that would be played.[8] Each of the two subjects represented a local monopoly regulated by yardstick competition. In each round, the two firms simultaneously decided on a cost level between 0 and 25. Then, the price-caps (resulting from the yardstick scheme of the corresponding treatment) were announced to the firms. Next, the firms specified their prices. Each firm's price had to be smaller or equal to the firm's price-cap. Finally, each firm's payoff was calculated and communicated to the corresponding subject. With this the round ended and the next round commenced. The same yardstick scheme was used for each pair in all 50 rounds.

The demand and cost functions were presented to the subjects both as formulas and in tables. Furthermore, an "as-if calculator" was provided on screen. This allowed them to see the price-caps resulting from any combination of hypothetical cost choices they typed in. By typing a hypothetical price choice, they could see the corresponding own payoff (profit and management benefit).

At the end of the experimental session, the subjects were paid the equivalent of their accumulated earnings. The experimental earnings were exchanged at the rate of 3 Dutch Guilders per 1000 points. Subjects earned between 34 and 63 Dutch Guilders, with an average of about 55 (which at the time corresponded to about 23 US dollars). Since the average

---

[8] Given the finite number of rounds played, and common knowledge about this, collusion is not a subgame perfect equilibrium of the repeated game. Previous experiments have shown though that even finitely repeated games sometimes lead to cooperative outcomes, at least in the intitial stages of the game (e.g., Selten and Stoecker, 1986). Note furthermore that collusion can still be supported as a Nash equilibrium even though it is not subgame perfect.

duration of the experimental session was somewhat less than two hours, the average per hour payment was well over the typical student wage of about 15 Dutch Guilders per hour.

We collected 16 independent observations in the DIS treatment and 13 independent observations in the UNI treatment. These independent observations are the basis for the statistical evaluation of the data in the next section.

5. Experimental Results

Our main interest is in the cost levels. First, we will discuss the development of the average cost levels. Second, we analyze the frequency distribution and variance of the cost levels. Third, we analyze the extent to which the cost levels for each pair of players are coordinated. Finally, we discuss prices and welfare levels.

Figure 1 shows the development of the cost levels over the 50 rounds of the experiment, averaged over all players in each of the two treatments. We see that average cost choices in the early rounds start at a level of about 13 in both treatments. From round 10 on we see a gradual increase in the average cost levels. This increase is somewhat more pronounced in the DIS treatment than in the UNI treatment. From round 15 onwards, the average cost level in the groups under the discriminatory yardstick exceeds the level in the uniform yardstick. By round 40 the average cost choice in the DIS treatment reaches a level at 18.4, which is remarkably close to the full collusion level of 20. In the UNI treatment, the average cost choice reaches its peak at a level of 16.8 in round 47. Finally, towards the end of the session, an end-effect kicks in.[9] This end-effect is more pronounced under the discriminatory than under the uniform scheme.

A look at the frequency distributions of cost choices in Figure 2, reveals that the distributions for both treatments have a peak at the full collusive cost choice $c_i^C = 20$. In DIS, almost 60 percent of the cost choices are equal to 20. The rest of the cost choices are distributed widely with most weight around the cost level of 8, which is equal to the optimal defection level (c_i^{D-DIS}). Remarkably, the non-cooperative equilibrium, $c_i^{N-DIS} = 4$, has almost no drawing power in this treatment. In UNI, only about 25 percent of the cost choices are equal to 20.

---

[9] Such end effects are frequently observed in finitely repeated games. Most notably, Selten and Stoecker (1986) have documented that the end effect can be observed regularly and follows certain behavioral patterns.

There is a second distinct peak in the distribution around the non-cooperative equilibrium $c_i^{N-UNI} = 13.1$ (which in this case is not distinguishable from the optimal defection level $c_i^{D-UNI} = 13.4$). In sum, these frequency distributions confirm that (full) collusion is much more prominent in the discriminatory than in the uniform treatment.
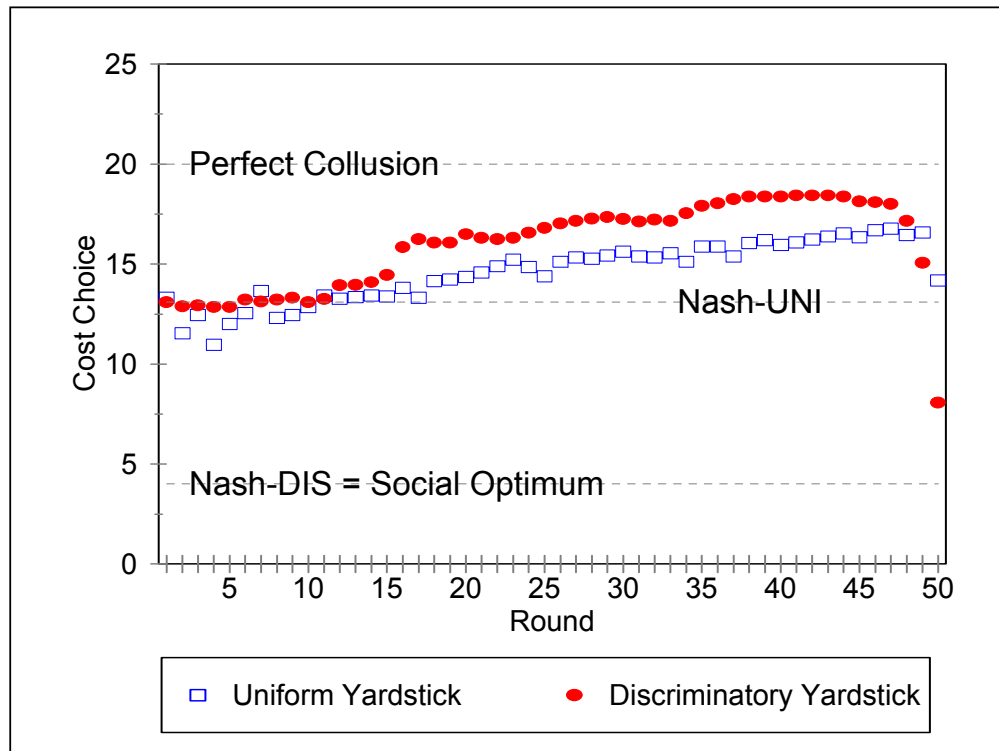

Figure 1: Development of the average cost choices

Figure 2 also suggests that the cost choices in DIS are more extreme than in UNI with both more collusive and more competitive choices. We can test for this phenomenon by showing that the distance between the average cost choice and the overall median is smaller for the data from the UNI treatment than for the data from the DIS treatment. Applying the Mann-Whitney-U test, we find statistical significance allowing the rejection of the "no-difference" null hypothesis in favor of the alternative hypothesis (5 percent level, one-tailed).
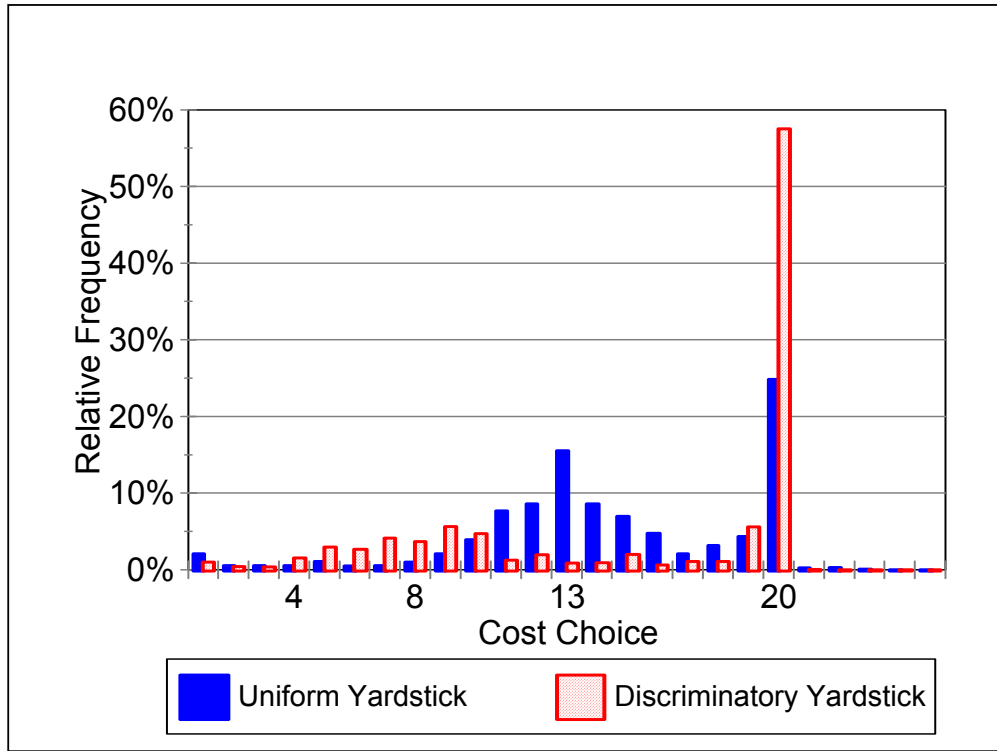
Figure 2: Distribution of cost choices

The frequency of collusive cost choices indicates that the game with a discriminatory yardstick is more prone to collusion than the game with a uniform yardstick. The next question we examine is whether the players actually managed to (tacitly) coordinate their collusive actions. Figure 3 shows the distribution of the frequency of coordinated full collusive outcomes in each regulated market (i.e. for each pair of monopolies). The categories on the horizontal axis indicate the number of rounds (out of 50) in which a pair of players colluded perfectly (i.e. both players chose the cost level $c_i^C = 20$). Again we see a clear treatment difference. While over 60 percent of the monopoly pairs in the UNI treatment never manage to coordinate on perfectly collusive play, this is true only 19 percent of the pairs in the DIS treatment. On the other hand, almost 40 percent of the pairs in the DIS treatment exhibit perfectly collusive play in more than 40 of the 50 rounds, whereas this occurs in less than 10 percent of the pairs in the UNI treatment. The average number of perfect cooperation rounds in the UNI treatment is 11.3, while it is 27.3 in the DIS treatment. A Mann-Whitney U-test applied to each pair's number of perfectly collusive rounds supports the treatment difference significantly (5 percent level, one-tailed).
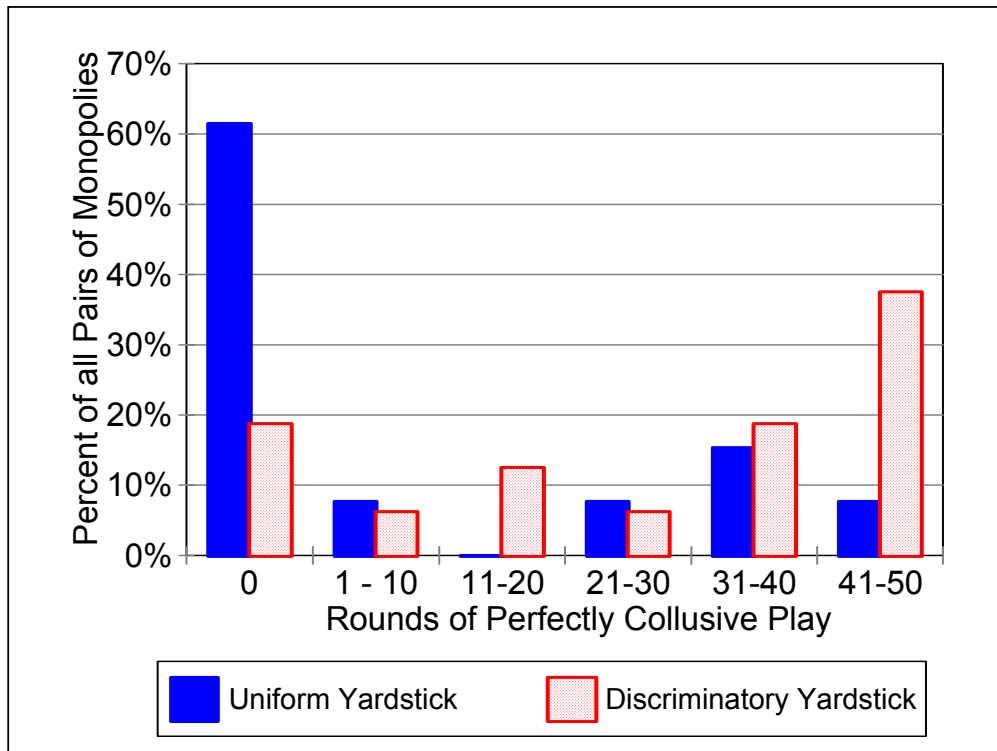
13

Figure 3: Frequency of perfectly collusive play

It should be noted that it is not just at the collusive outcome that the cost choices in the DIS treatment are better coordinated than in the UNI treatment. The average absolute difference in the cost levels between the two players in a market (averaged over the 50 rounds), is 1.4 in DIS and 2.2 in UNI. This difference between the two treatments is also significant (Mann-Whitney U-test, 5 percent level, one-tailed). In the DIS treatment, players are clearly better able to coordinate their cost levels.

Apart from the cost choices, we can also look at the price choices. Since the price-caps under both regimes are binding, rational payoff maximizing agents would always choose their price equal to the given price-caps, i.e. $p_i = \hat{p}_i$. In fact, we find that in about 99 percent of all cases subjects choose prices that are equal to the price-cap (99.2 percent in DIS and 98.2 percent in UNI, with no statistical difference between treatments). This result seems to confirm that our subjects did understand the nature of the game and were actually maximizing payoffs.

The analysis above shows that the players in the DIS treatment behave significantly more collusive than in the UNI treatment. Theoretically, we would expect that this also leads to lower social welfare (efficiency) levels. Table 2 presents welfare levels achieved theoretically and empirically under the different yardstick regimes. Our measure of welfare is the ratio of

actually realized surplus (consumer plus producer surplus) to the optimal surplus in the social optimum. The last column of the table shows the average welfare observed in the two treatments of our experiment. It is obvious that the high incidence of collusion in the DIS treatment leads to substantial welfare losses. Moreover, the table shows that the theoretical ordering of the two types of yardstick competition under the assumption of non-cooperative behavior is empirically reversed. Realized welfare levels are higher on average under the uniform than under the discriminatory regime. Figure 4 gives graphical support to the observation. The frequency distribution of welfare levels in UNI is clearly to the right of the distribution of welfare in DIS. The difference furthermore is statistically significance (Mann-Whitney U-test, 5 percent level, one-tailed).

Table 2. Theoretical and realized welfare levels

	social optimum	non-cooperative equilibrium	full collusion	realized average in experiment
discriminatory yardstick	100%	100%	73%	76%
uniform yardstick	100%	91%	73%	88%

Finally, we briefly turn to the question: How are players able to coordinate on collusion and why are they better able to do so under the discriminatory regime than under the uniform regime? We find no evidence that the subjects used trigger strategies in the strict game-theoretical sense to enforce collusion. At the same time, however, there is evidence for strategic behavior. For one thing, we observe a clear end-effect. This suggests that players are well aware that collusion hinges on the repeated nature of the interaction. Furthermore, the subjects react to each others' cost choices. The numbers in table 3 indicate how players respond to their relative cost position in the previous round. For example in the UNI treatment, a player who has a strictly higher cost level than the opponent in a round, goes up with the own cost level in the next round in 20 percent of the relevant cases and goes down in 42 percent of the cases. The player who has the lower cost, however, goes up with the own cost level in 54 percent of the cases and goes down in only 19 percent of the cases. These percentages are similar in the DIS treatment. Hence, players respond to their relative cost

15

positions in a manner that is similar to the "measure-for-measure" pattern that was found to support collusion in the duopoly experiment by Selten, Mitzkewitz, and Uhlich (1997).
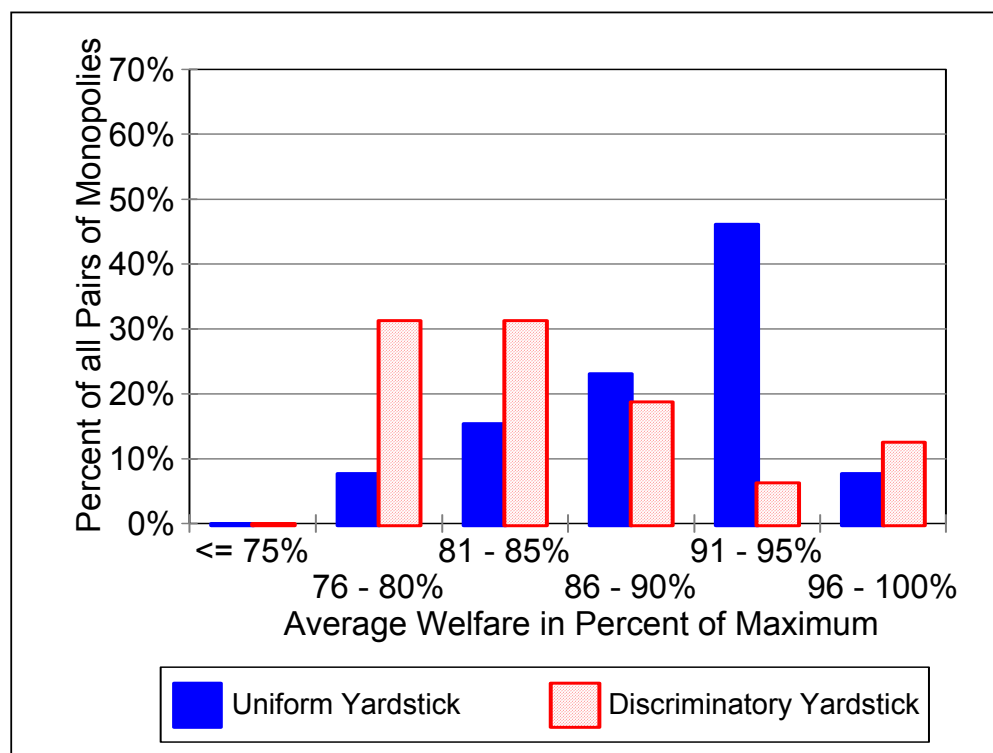

Figure 4: Distribution of welfare

Interestingly, these strategic responses are somewhat stronger in the DIS treatment than in the UNI treatment. For example, the player with the lower cost level in a round, on average increases the own cost in the next round by 2.7 in the DIS treatment, but only by 1.3 in the UNI treatment. Similarly, the player with the higher cost level in a round, on average decreases the own cost level in the next round by 2.1 in the DIS treatment, but only by 1.2 in the UNI treatment.

It is not obvious why this is the case. One possibility is that signals of cooperation and defection are clearer under a discriminatory scheme than under the uniform scheme. Under the discriminatory pricing rule the own cost choice does not influence the own price cap. This means that choosing a high cost level only has a positive payoff effect for the other player. This turns the choice of a high cost level into a clear cooperative signal to the other player.

Under the uniform yardstick the choice of a high cost level is a more ambiguous signal, because a high cost level not only benefits the other player, it also increases the own price cap. To the extend that clearer signals lead to stronger reciprocation, collusion under the discriminatory yardstick scheme may be facilitated relative to the uniform scheme. As Katzenbach, Kottman, and Krueger (1995) put it: "Once the considerable relevance of implicit collusion is taken into account, great importance attaches especially to market components which facilitate signaling and which therefore pave the way for covert coordination."

Table 3. Reactions to previous round's relative cost position

cost level in the previous round was	higher than opponent		lower than opponent	
	uniform	discriminatory	uniform	discriminatory
increases cost level	20%	13%	54%	58%
decreases cost level	42%	43%	19%	9%

7. Conclusion

We find that the discriminatory yardstick scheme leads to one of two types of behavior. In some case, the two firms engage in "virtual competition," in which case they choose cost levels close to the social optimum, as predicted by the non-cooperative equilibrium of the game. In other cases, the two firms engage in almost perfect collusion, choosing high cost levels and maximizing joint payoffs. In contrast, the uniform yardstick regulation leads to more moderate outcomes. The cost level is clearly above to the social optimum, but there is also significantly less collusive action than in the discriminatory yardstick regime.

These observations are roughly in line with the theoretical predictions. Non-cooperative theory predicts that the discriminatory yardstick scheme should lead to lower cost levels and higher welfare levels than the uniform yardstick scheme. In fact, we observe lower cost levels more frequently in the treatment with the discriminatory yardstick than in the treatment with the uniform yardstick. On the other hand, because the players receive a smaller part of the rent under the discriminatory regulation, the repeated game analysis suggests that players in that regime have a higher incentive to collude. In fact, we observe that there is significantly more

collusion in the treatment with the discriminatory yardstick than in the treatment with the uniform yardstick.

Our results suggest that principals and regulators should be careful when setting up relative performance evaluation schemes. High-powered schemes which leave few rents to the regulated agents may be particularly prone to collusion. Exclusive focus on the non-cooperative equilibria of an incentive scheme is likely to be short-sighted for environments in which the regulated agents interact repeatedly.

References

Abbink, Klaus, and Abdolkarim Sadrieh (1995): "RatImage – Research Assistance Toolbox for Computer-Aided Human Behavior Experiments." *SFB Discussion Paper B-325*, University of Bonn.

Clark, Kenneth and Martin Sefton (2001): "Repetition and signalling: experimental evidence from games with efficient equilibria," Economics Letters 70, 357-362.

Croson, Rachel (1996): "Partners and strangers revisited," *Economics Letters* 53, 25-32.

Dte (Office for Energy Regulation of the Netherlands) (2001): *Market-Based Regulation Method for Electricity Tariffs for Captive Consumers*, Amsterdam: SKM Energy Consulting.

Dye, Ronald A. (1984): "The Trouble with Tournaments," *Economic Inquiry* 22 (1), 147-149.

Fouraker, Lawrence E. and Sidney Siegel (1963): *Bargaining Behavior*, New York: McGraw-Hill.

Friedman, James W. (1971): "A Noncooperative View of Oligopoly," *International Economic Review* 12(1), 106-122.

Fudenberg, Drew, and Jean Tirole (1991): *Game Theory*, Cambridge, MA: MIT Press.

Gibbons, Robert and Murphy, Kevin J. (1990): "Relative Performance Evaluation for Chief Executive Officers," *Industrial and Labor Relations Review* 43, 30-51.

Holmstrom, Bengt (1982): "Moral Hazard in Teams," *Bell Journal of Economics* 13 (2), 324-340.

Holt, Charles A. (1995): "Industrial Organization: A Survey of Laboratory Research," *Handbook of Experimental Economics*, John H. Kagel and Alvin E. Roth (eds.), Princeton, NJ: Princeton University Press, 349-444.

Huck, Steffen, Hans-Theo Normann, and Jörg Oechssler (1999): "Learning in Cournot Oligopoly – An Experiment," *Economic Journal* 109, C80-C95.

Huck, Steffen, Hans-Theo Normann, and Jörg Oechssler (2000): "Does information about competitors' actions increase or decrease competition in experimental oligopoly markets?" *International Journal of Industrial Organization* 18, 39-57.

Katzenbach, Erhard, Elke Kottman, and Reinald Krueger (1995): "New Industrial Economics and Experiences from European Merger Control - New Lessons about Collective Dominance?" Study prepared for the European Commission, Luxembourg: European Commission.

Keser, Claudia and Frans van Winden (2000): "Conditional Cooperation and Voluntary Contributions to Public Goods," Scandinavian Journal of Economics 102(1), 23-40.

Laffont, Jean-Jacques and David Martimort (1997): "Collusion under Asymmetric Information," *Econometrica* 65, 875-911.

Laffont, Jean-Jacques and David Martimort (1998): "Collusion and Delegation," *Rand Journal of Economics* 29, 280-305.

Mookherjee, Dilip (1984): "Optimal Incentive Schemes with Many Agents," *Review of Economic Studies* 51 (3), 433-46.

Offerman, Theo, Jan Potters, and Joep Sonnemans (2002): "Imitation and adaptation in an oligopoly experiment," *Review of Economic Studies*, forthcoming.

Roth, Alvin E. (1995): "Introduction to Experimental Economics," *Handbook of Experimental Economics*, John H. Kagel and Alvin E. Roth (eds.), Princeton, NJ: Princeton University Press, 3-110.

Shleifer, Andrei (1985): "A Theory of Yardstick Competition," *Rand Journal of Economics* 16, 319-327.

Selten, Reinhard, Michael Mitzkewitz, and G. Uhlich (1997): "Duopoly Strategies Programmed by Experienced Players," *Econometrica* 65, 517-555.

Selten, Reinhard and Rolf Stoecker (1986): "End Behavior in Sequences of Finite Prisoner's Dilema Supergames - A Learning Theory Approach," *Journal of Behavior and Organization* 7, 47-70.

Sobel, Joel (1999): "A Reexamination of Yardstick Competition," *Journal of Economics and Management Strategy* 8, 33-60.

Tangeras, Thomas P. (2002): "Collusion-proof yardstick competition," *Journal of Public Economics* 83, 231–254.

# Appendix. Instructions for the Experiment

The Experimental Situation

- The world is divided in the two **regions** A and B.

- The two **firms** A and B offer their service to the customers of their region (firm A in region A and firm B in region B). Each firm is the only service provider in its region, meaning that no firm has a competitor in its region.

- Each firm has to make two decisions: **technology** (c) and **price** (p).

- In the experiment, you will make the decisions for one of the firms. Your earnings (in Guilders), are determined by the **management payoff** of that firm (in points). For each point you will receive 0.3 Cents.

- The **management payoff** is to sum of two components: **profits** and **management benefit**.

Technology, Unit Costs, and Management Benefit

- Each firm chooses a technology c out of 25 different technologies: 1, ..., 25.

- The choice of the technology influences the **unit cost** of the firm and the **management benefit**:

  - A firm with technology c has a **unit cost** of c.

    The higher the chosen technology, the higher is the unit cost for providing the service.

The Unit Cost of Service

  - The management of a firm with a technology c receives a **management benefit** of $40c - c^2$.

    The management benefit is highest for technology $c = 20$. For technologies lower than 20 the management benefit increases with an increase in the technology and for technologies higher than 20 the management benefit decreases with an increase in the technology.


The Management Benefit
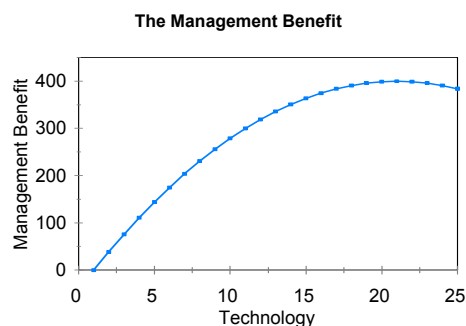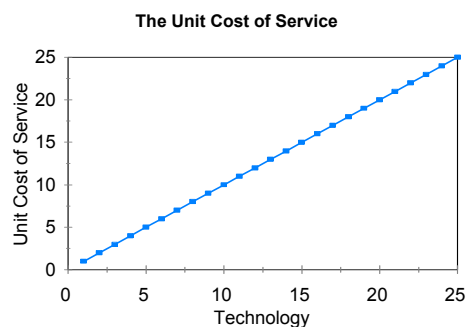
Price, Maximum Price, and Demand

- Each firm chooses the **price** p of its service.

- A firm's price must be between 1 and the **maximum price**. The maximum price is equal to the [**unit cost of the other firm**]^{DIS} [**average of the unit costs of both firms**]^{UNI}.

- The demand for a firm's service depends on its price. A firm with price p has a demand $34 - \frac{1}{2}p$.

  The higher the price of the service, the lower is the demand for the service.

The Demand Function
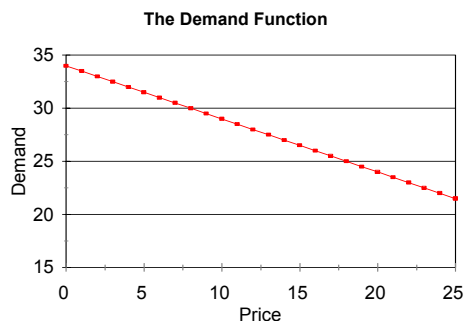
Profit and Payoff

- A firm with the technology (= unit cost) c and the price p has **profits** of $(p-c) \cdot (34 - \frac{1}{2}p)$.

  If the price is higher than the unit costs, the profit is positive. If the price equals the unit costs, the profit is zero. If the price is lower than the unit costs, the profit is negative. The attached table shows the profit for every possible combination of a technology choice c and an integer valued price choice p.

- The **payoff to the firm's management** is the sum of the profits plus the management benefit: $(p - c) \cdot (34 - \frac{1}{2}p) + 40c - c^2$. This payoff determines your earnings in the experiment.

Course of the Experiment

- The experiment consists of 50 consecutive rounds.

- In the beginning of the experiment you are randomly matched with another participant. Each of you plays the part of one of the two firms. The matching of subjects remains unchanged throughout the entire experiment.

- Each round consists of the following steps:

  1. The two firms A and B choose their technology c. The choices are made simultaneously, so when making its choice a firm does not know the technologies chosen by the other firm.

  2. The maximum price is determined and revealed to each firm.

  3. The two firms A and B simultaneously choose their price p.

  4. Each firm is informed on its own payoff (profit and management benefit).

  5. The round ends.

21

Profits (depending on Technology *c* and Price *p*) and Management Benefit (depending on Technology *c*)

c \ p	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	management benefit
1	0.0	33.0	65.0	96.0	126.0	155.0	183.0	210.0	236.0	261.0	285.0	308.0	330.0	351.0	371.0	390.0	408.0	425.0	441.0	456.0	470.0	483.0	495.0	506.0	516.0	39.0
2	-33.5	0.0	32.5	64.0	94.5	124.0	152.5	180.0	206.5	232.0	256.5	280.0	302.5	324.0	344.5	364.0	382.5	400.0	416.5	432.0	446.5	460.0	472.5	484.0	494.5	76.0
3	-67.0	-33.0	0.0	32.0	63.0	93.0	122.0	150.0	177.0	203.0	228.0	252.0	275.0	297.0	318.0	338.0	357.0	375.0	392.0	408.0	423.0	437.0	450.0	462.0	473.0	111.0
4	-100.5	-66.0	-32.5	0.0	31.5	62.0	91.5	120.0	147.5	174.0	199.5	224.0	247.5	270.0	291.5	312.0	331.5	350.0	367.5	384.0	399.5	414.0	427.5	440.0	451.5	144.0
5	-134.0	-99.0	-65.0	-32.0	0.0	31.0	61.0	90.0	118.0	145.0	171.0	196.0	220.0	243.0	265.0	286.0	306.0	325.0	343.0	360.0	376.0	391.0	405.0	418.0	430.0	175.0
6	-167.5	-132.0	-97.5	-64.0	-31.5	0.0	30.5	60.0	88.5	116.0	142.5	168.0	192.5	216.0	238.5	260.0	280.5	300.0	318.5	336.0	352.5	368.0	382.5	396.0	408.5	204.0
7	-201.0	-165.0	-130.0	-96.0	-63.0	-31.0	0.0	30.0	59.0	87.0	114.0	140.0	165.0	189.0	212.0	234.0	255.0	275.0	294.0	312.0	329.0	345.0	360.0	374.0	387.0	231.0
8	-234.5	-198.0	-162.5	-128.0	-94.5	-62.0	-30.5	0.0	29.5	58.0	85.5	112.0	137.5	162.0	185.5	208.0	229.5	250.0	269.5	288.0	305.5	322.0	337.5	352.0	365.5	256.0
9	-268.0	-231.0	-195.0	-160.0	-126.0	-93.0	-61.0	-30.0	0.0	29.0	57.0	84.0	110.0	135.0	159.0	182.0	204.0	225.0	245.0	264.0	282.0	299.0	315.0	330.0	344.0	279.0
10	-301.5	-264.0	-227.5	-192.0	-157.5	-124.0	-91.5	-60.0	-29.5	0.0	28.5	56.0	82.5	108.0	132.5	156.0	178.5	200.0	220.5	240.0	258.5	276.0	292.5	308.0	322.5	300.0
11	-335.0	-297.0	-260.0	-224.0	-189.0	-155.0	-122.0	-90.0	-59.0	-29.0	0.0	28.0	55.0	81.0	106.0	130.0	153.0	175.0	196.0	216.0	235.0	253.0	270.0	286.0	301.0	319.0
12	-368.5	-330.0	-292.5	-256.0	-220.0	-186.0	-152.5	-120.0	-88.5	-58.0	-28.5	0.0	27.5	54.0	79.5	104.0	127.5	150.0	171.5	192.0	211.5	230.0	247.5	264.0	279.5	336.0
13	-402.0	-363.0	-325.0	-288.0	-252.0	-217.0	-183.0	-150.0	-118.0	-87.0	-57.0	-28.0	0.0	27.0	53.0	78.0	102.0	125.0	147.0	168.0	188.0	207.0	225.0	242.0	258.0	351.0
14	-435.5	-396.0	-357.5	-320.0	-283.5	-248.0	-213.5	-180.0	-147.5	-116.0	-85.5	-56.0	-27.5	0.0	26.5	52.0	76.5	100.0	122.5	144.0	164.5	184.0	202.5	220.0	236.5	364.0
15	-469.0	-429.0	-390.0	-352.0	-315.0	-279.0	-244.0	-210.0	-177.0	-145.0	-114.0	-84.0	-55.0	-27.0	0.0	26.0	51.0	75.0	98.0	120.0	141.0	161.0	180.0	198.0	215.0	375.0
16	-502.5	-462.0	-422.5	-384.0	-346.5	-310.0	-274.5	-240.0	-206.5	-174.0	-142.5	-112.0	-82.5	-54.0	-26.5	0.0	25.5	50.0	73.5	96.0	117.5	138.0	157.5	176.0	193.5	384.0
17	-536.0	-495.0	-455.0	-416.0	-378.0	-341.0	-305.0	-270.0	-236.0	-203.0	-171.0	-140.0	-110.0	-81.0	-53.0	-26.0	0.0	25.0	49.0	72.0	94.0	115.0	135.0	154.0	172.0	391.0
18	-569.5	-528.0	-487.5	-448.0	-409.5	-372.0	-335.5	-300.0	-265.5	-232.0	-199.5	-168.0	-137.5	-108.0	-79.5	-52.0	-25.5	0.0	24.5	48.0	70.5	92.0	112.5	132.0	150.5	396.0
19	-603.0	-561.0	-520.0	-480.0	-441.0	-403.0	-366.0	-330.0	-295.0	-261.0	-228.0	-196.0	-165.0	-135.0	-106.0	-78.0	-51.0	-25.0	0.0	24.0	47.0	69.0	90.0	110.0	129.0	399.0
20	-636.5	-594.0	-552.5	-512.0	-472.5	-434.0	-396.5	-360.0	-324.5	-290.0	-256.5	-224.0	-192.5	-162.0	-132.5	-104.0	-76.5	-50.0	-24.5	0.0	23.5	46.0	67.5	88.0	107.5	400.0
21	-670.0	-627.0	-585.0	-544.0	-504.0	-465.0	-427.0	-390.0	-354.0	-319.0	-285.0	-252.0	-220.0	-189.0	-159.0	-130.0	-102.0	-75.0	-49.0	-24.0	0.0	23.0	45.0	66.0	86.0	399.0
22	-703.5	-660.0	-617.5	-576.0	-535.5	-496.0	-457.5	-420.0	-383.5	-348.0	-313.5	-280.0	-247.5	-216.0	-185.5	-156.0	-127.5	-100.0	-73.5	-48.0	-23.5	0.0	22.5	44.0	64.5	396.0
23	-737.0	-693.0	-650.0	-608.0	-567.0	-527.0	-488.0	-450.0	-413.0	-377.0	-342.0	-308.0	-275.0	-243.0	-212.0	-182.0	-153.0	-125.0	-98.0	-72.0	-47.0	-23.0	0.0	22.0	43.0	391.0
24	-770.5	-726.0	-682.5	-640.0	-598.5	-558.0	-518.5	-480.0	-442.5	-406.0	-370.5	-336.0	-302.5	-270.0	-238.5	-208.0	-178.5	-150.0	-122.5	-96.0	-70.5	-46.0	-22.5	0.0	21.5	384.0
25	-804.0	-759.0	-715.0	-672.0	-630.0	-589.0	-549.0	-510.0	-472.0	-435.0	-399.0	-364.0	-330.0	-297.0	-265.0	-234.0	-204.0	-175.0	-147.0	-120.0	-94.0	-69.0	-45.0	-22.0	0.0	375.0