Energy Economics 66 (2017) 261–271

Contents lists available at ScienceDirect

Energy Economics

journal homepage: www.elsevier.com/locate/eneco

Design of yardstick competition and consumer prices: Experimental evidence☆

Peter T. Dijkstra[a,*,1], Marco A. Haan[b], Machiel Mulder[b]

[a] Netherlands Authority for Consumers & Markets (ACM), The Hague, The Netherlands
[b] Department of Economics, Econometrics and Finance, University of Groningen, P.O. Box 800, Groningen 9700 AV, The Netherlands

ARTICLE INFO

Article history:
Received 18 December 2015
Received in revised form 12 January 2017
Accepted 24 June 2017
Available online 8 July 2017

JEL classification:
C73
C92
L13
L51

Keywords:
Collusion
Industry structure
Yardstick competition
Experiment

ABSTRACT

In this paper we analyze the effect of the design of yardstick competition on consumer prices, by means of a theoretical analysis as well as an economic experiment. We compare four different designs: the uniform yardstick, the unweighted uniform yardstick, the discriminatory yardstick, and the best-practice yardstick. The effect of a specific design on prices depends on two separate mechanisms, one which affects the incentive power to increase productive efficiency and another which affects the risk of collusion. We show theoretically that for the best-practice yardstick, which is widely applied in several industries in a number of countries, these two mechanisms point in the same direction (high prices), which is confirmed by the experiment. Both the theoretical analysis as well as the economic experiment show that the discriminatory yardstick results in lower prices than the unweighted uniform yardstick. The theory, however, does not give a clear answer on the relative performance of the discriminatory versus the weighted uniform yardstick. In the experimental analysis, we find that the advantage of the discriminatory yardstick in terms of giving incentives to improve productive efficiency exceeds the disadvantage of a relatively higher risk of collusion. This conclusion appears to be robust for different degrees of heterogeneity of the industry. Hence the discriminatory yardstick yields the lowest prices for consumers.

1. Introduction

The revenues of firms subject to tariff regulation are often based on yardstick competition. With this type of regulation, the (maximum) price for consumers is based on the costs of a group of similar firms. The idea behind yardstick competition is that this gives each firm an incentive to produce efficiently, while the benefits of these efficiency improvements are transferred to consumers. In the US, for example, the payment a hospital received for a treatment under *Medicare*, depended on the average cost of that treatment in other hospitals (Shleifer, 1985). Yardstick competition is used in industries as diverse as water supply and sewerage,[2] electricity networks,[3] railways[4] and bus transport services,[5] to name but a few. In several cases, the introduction of yardstick regulation improved productive efficiency or resulted in lower consumer prices. In other cases, however, the experience appeared to be less successful.

There are many ways to implement yardstick competition. With a uniform yardstick, prices that firms can charge depend on the average costs of all firms. That average can be weighted or unweighted average costs. With a discriminatory yardstick, the price that an individual firm can charge depends on the average costs of all *other* firms. Each firm then faces a different price cap. With a best-practice yardstick, prices depend on the costs of the most efficient firm.

If firms behave in a noncooperative fashion, the incentive to raise productive efficiency is higher when a firm's regulated price bears less relation to its own costs. Indeed, Shleifer (1985) shows that a discriminatory yardstick then leads to the socially efficient level of cost

☆ We thank Editor Richard Tol and three anonymous referees for their valuable comments. We are also indebted to participants of CRESSE 2015 (Crete) and seminar participants at the University of Groningen (RUG) and the Netherlands Authority for Consumers & Markets (ACM). Financial support of the Netherlands Authority for Consumers & Markets (ACM) is gratefully acknowledged. The authors are fully responsible for the contents of the paper.

* Corresponding author.
E-mail address: Peter.Dijkstra@acm.nl (P.T. Dijkstra).
[1] Previous affiliation: Department of Economics, Econometrics and Finance, University of Groningen, P.O. Box 800, 9700 AV Groningen, The Netherlands.

[2] Dassler et al. (2006).
[3] Haffner et al. (2010), Blázques-Gómez and Grifell-Tatjé (2011), Jamasb and Pollitt (2007).
[4] Mizutani et al. (2009).
[5] Dalen and Gomez-Lobo (2002).

http://dx.doi.org/10.1016/j.eneco.2017.06.016
0140-9883/

reduction. But a yardstick that performs well under the assumption that firms behave non-cooperatively may still backfire if it facilitates collusion. Firms may agree, perhaps tacitly, not to reduce costs or to jointly manipulate productivity reports (Tangerås, 2002). To evaluate the effectiveness of a yardstick we have to take into account the incentive to increase efficiency as well as the incentive to collude; some types of yardstick may make collusion easier than others. A priori, it is not always clear which of these two possibly countervailing effects is strongest. The risk of collusion in case of yardstick competition was already mentioned by Shleifer (1985), but to what extent that risk could mitigate the benefits of yardstick competition has hardly been analyzed up to now.

Hence, given the fact that regulated firms may be able to collude, it is not clear what type of yardstick would be most successful in achieving low consumer prices. To try to establish the potential impact of collusion, we conduct a laboratory experiment. First, we evaluate the relative performance of a weighted uniform, an unweighted uniform, a discriminatory, and a best-practice yardstick. We do so in an environment with two firms that have different sizes, and study the effect of yardstick type on average prices and on the incidence of collusion. In a second experiment, we zoom in on the weighted uniform, and the discriminatory yardstick. These two yardsticks perform well in our first experiment, but both theory and that first experiment do not provide a clear prediction in terms of their market performance. We compare these two yardsticks for a variety of industry structures as the distribution of firms regarding market size may have an effect on the effectiveness of yardstick competition.

This paper is not the first to conduct such experiments. Potters et al. (2004) find that the extent of collusion as well as consumer prices are higher with a discriminatory rather than a uniform yardstick. Apparently, the authors conclude, with a discriminatory yardstick the stronger incentives to collude outweigh the higher incentives to raise productive efficiency. Yet, the authors only show this for two equally-sized firms, prompting the question whether a more heterogeneous firm-size distribution would give the same result. In a companion paper (Dijkstra et al., 2017), we study the effect of industry structure on the incentive to collude with a (weighted) uniform yardstick. We find that in more heterogeneous industries, firms appear to collude less than in homogeneous industries.[6]

In our experiment subjects decide upon marginal cost levels, knowing that their output prices are set by a regulator using a specific form of yardstick competition. Our design is the same as that in Dijkstra et al. (2017). We use a simple model, loosely based on Shleifer (1985), where managers have to exert costly effort to lower their costs. Managers aim to maximize the sum of profits and managerial benefits, and do so in a repeated game. For simplicity, managerial benefits are a concave and quadratic function of marginal costs.

In our first experiment we compare the four different yardsticks for an industry with two firms with different sizes. We find that the price for consumers is highest for a best-practice yardstick, while a discriminatory yardstick yields prices that are lower (i.e. closer to the social optimum) than an unweighted uniform yardstick. The discriminatory yardstick seems to result in lower prices but more collusion than a weighted uniform yardstick, but not significantly so. In our second experiment, we compare the weighted uniform and the discriminatory yardstick for different industry structures. We find that the discriminatory yardstick yields the lowest consumer prices overall. From this we conclude that the discriminatory yardstick is the preferred yardstick in order to raise productive efficiency of monopolistic firms and to pass on these benefits to consumers. Note that this conclusion runs counter to that in Potters et al. (2004)

As noted, in a companion paper (Dijkstra et al., 2017) we use a similar set-up to study a different question. Therefore, parts of the description of the theoretical model, the experimental set-up and the literature survey in this paper closely follows that in our companion paper.

The structure of this paper is as follows. Section 2 gives some background information, including a discussion on policy practice, as well as related experiments. Section 3 describes our theoretical model. The experimental design is presented in Section 4. Section 5 discusses the result of our experiment with 4 yardsticks, while Section 6 discusses the experiment that compares the discriminatory and (weighted) uniform yardsticks. Section 7 concludes.

2. Background

2.1. Yardstick competition[7]

The last decades saw the liberalization of several network industries that used to be state-owned vertically-integrated monopolies. Examples include telecommunications, electricity, gas, water and sewerage. Parts of these industries are natural monopolies characterized by subadditivity of costs, which calls for regulatory supervision (Viscusi et al., 2005). A key component of such supervision is the regulation of tariffs. Historically, regulated tariffs were either based on actual costs or an allowed rate of return. However, such schemes give little incentive to increase productive efficiency as lower costs directly translate to lower revenues. Hence, the incentive power is low, as a change in costs hardly affect firms' profits. Introducing price-cap regulation solves this problem as any cost decrease then fully benefits the firm. But this method also implies the risk of significant rents or losses for regulated firms. To overcome both problems, Shleifer (1985) proposed to impose tariffs based on the actual costs of a group of similar (benchmark) firms. This type of tariff regulation is called yardstick regulation. As tariffs are based on the relative performance of a firm, yardstick regulation is generally seen as a powerful tool both for giving incentives for productive efficiency as well as for rent extraction (Tangerås, 2002; Burns et al., 2005).

The incentive power of a yardstick partly depends on the exact manner in which costs of the benchmark firms determine tariffs. There are various types of yardsticks. With a *uniform yardstick* every firm faces the same cap on the tariffs it may charge, based on some average of the cost information of all firms in the benchmark group. With a *discriminatory yardstick*, every firm faces a specific cap based on the costs of all *other* firms in the benchmark group. With a *best-practice yardstick*, every firm faces the same cap, based on the cost information of the most efficient firm.

2.2. Practical implementations and experience

Yardstick regulation has been implemented in a number of industries in several countries. In the United Kingdom, yardstick competition is often based on some type of best practice. Dassler et al. (2006) show how it is applied in telecommunication, gas, electricity, water and sewerage. For electricity, the yardstick was based on the average performance in the group of similar firms, though the exact implementation differs.[8] A reason for this different treatment is the degree

[6] The result that asymmetries might lead to less collusion is also observed in unregulated markets, see e.g. Mason et al. (1992), Fonseca and Normann (2008), Dugar and Mitra (2009), Phillips et al. (2011), and Dugar and Mitra (2016).

[7] This section partly follows our companion paper, Dijkstra et al. (2017).

[8] In their assessment of the UK regulatory scheme for electricity distribution networks, for example, Jamasb and Pollitt (2007) state that the regulatory framework included different treatments of operational costs and capital costs. For operational costs (OPEX), the price cap was based on a form of best-practice benchmarking. For capital costs (CAPEX) standardized parameters were used regarding costs of capital and depreciation methods.