Kuosmanen, Timo; Nguyen, Tuan

Capital bias in the Nordic revenue cap regulation

Published in:
Energy Policy

DOI:
10.1016/j.enpol.2020.111355

Published: 01/04/2020

Document Version
Peer reviewed version

Published under the following license:
CC BY-NC-ND

Please cite the original version:
Capital bias in the Nordic revenue cap regulation: Averch-Johnson critique revisited

Timo Kuosmanen* and Tuan Nguyen
Aalto University School of Business, 02150 Espoo, Finland.

Abstract
Nordic energy market reform and the regulation of local monopolies in electricity transmission and distribution sectors have served as role models for many other countries worldwide. The first contribution of this paper is to clarify the conceptual distinction between the Nordic revenue cap approach and the British revenue cap regulation. Our second contribution is to show that the Nordic revenue cap is similar to the U.S.-style rate of return regulation in that both are subject to capital bias, known as the Averch-Johnson effect. The third contribution of this paper is to examine the magnitude of the capital bias and its welfare effects by means of numerical simulations. We show that the Nordic revenue cap generally decreases the monopoly profit, increases the output, decreases the price, and hence increases consumer surplus compared to the unregulated monopoly. The simulation results prove robust to changes in the parameter values and the functional form of the production function. Our numerical simulations reveal that relatively light handed regulation suffices to yield the main benefits.

Key words: electricity transmission and distribution, incentive regulation, monopoly, price-cap, revenue-cap.
JEL classification: D24, D42, L43

* Corresponding author: E-mail timo.kuosmanen@aalto.fi, Tel +358943131.
1. Introduction

Electricity transmission and distribution services are classic examples of a natural monopoly. Several approaches to protect the customers from abuse of market power and avoid the inefficiency loss of the monopoly are known in the literature (e.g., Laffont and Tirole, 1994; Armstrong et al., 1994; Newbery, 1999). Based on the experiences from the regulation of other network industries such as U.S. railroads in the late 19th century, the rate of return regulation was widely used in the government regulation of monopolies in the first half of the 20th century (e.g., Baumol and Klevorick, 1970; Kahn, 1971). In this regime, the regulated company can recover the incurred costs while the customers are protected from paying monopoly profits. However, the rate of return regulation incentivizes the regulated monopoly to over-invest in capital, as first shown by Averch and Johnson (1962). Together with the works of Milton Friedman, Averch and Johnson (1962) is one of the key intellectual roots of the deregulation movement that started in the 1980s (e.g., Joskow and Schmalensee, 1988), and still continues in many countries today.1

In the U.K., deregulation of public utilities in the 1980s was combined with privatization. To avoid price increase due to privatization, the standard approach in the U.K. is referred to as price cap regulation (Littlechild, 1983). In this regime, the regulated monopoly can increase its prices along the general inflation rate, measured by the Retail Prices Index (RPI; in the U.K.) or the Consumer Price Index (CPI), minus the expected productivity growth denoted by X (see Bernstein and Sappington, 1999, for further discussion). Therefore, the British price cap regulation is also referred to as RPI-X or CPI-X approach (e.g., Mayer and Vickers, 1996; Jamison, 2007). In its purest form, the price cap is completely divorced from the firm’s realized production costs and earnings, which makes it immune to the Averch-Johnson effect (Bernstein and Sappington, 1999). The RPI-X or CPI-X can also be used to restrict the rate of revenue growth. For the sake of clarity, we will henceforth refer to such exogenous restrictions to the revenue growth as the British revenue cap.

In the continental Europe, the Nord Pool energy exchange is often cited as the most successful example of a liberalized competitive electricity market (e.g., Oseni and Pollitt, 2016). First launched in Norway in 1993, Nord Pool gradually expanded to Sweden, Finland and Denmark by the end of 1990s, and currently operates in 15 countries in the Northern and Western Europe.2 Clearly, a competitive electricity market cannot function without reliable transmission and distribution networks. Despite the market reforms in the electricity generation and retail markets,

1 Despite the early lead in the 1980s, deregulation of the U.S. energy sector ground to a halt as a result of the California electricity crisis, Enron’s bankruptcy, and fraudulent accounting practices (e.g., Joskow, 2006; Fabrizio et al. 2007).
2 For further information, see: https://www.nordpoolgroup.com/About-us/History/
the transmission and distribution services remain natural monopolies subject to government regulation. The rapid market-based reform in the Nordic countries in the 1990s created pressure for the Nordic energy regulators to develop incentive regulation of their electricity transmission and distribution sectors, which is has been covered in a large number of scientific articles (see, e.g., Agrell et al., 2005; Agrell and Bogetoft, 2010, Kuosmanen, 2012; Kuosmanen et al., 2013; Saastamoinen et al., 2017, and references therein). The Nordic electricity sector has also served as an example for many other European countries, for example, Germany (see, e.g., Agrell and Bogetoft, 2007).

The Nordic energy regulators’ approach is also often referred to as a revenue cap both in policy papers and scientific articles (e.g., NordREG; 2011; Kuosmanen et al., 2013; Saastamoinen et al., 2017). The first contribution of this paper is to clarify the conceptual distinction between the Nordic and the British interpretations of the revenue cap. We show that the Nordic revenue cap actually shares some common features with the U.S.-style rate of return regulation. Therefore, the second contribution of this paper is to formally show that the Nordic revenue cap is subject to capital bias, known as the Averch-Johnson effect.

Despite the capital bias, in our experience, the Nordic revenue cap seems to work reasonably well in the real-world regulation practice, achieving most of its intended objectives. To gain further insight, the third contribution of this paper is to compare the magnitude of the Averch-Johnson effect and the desirable welfare effects of the Nordic revenue cap regulation by means of numerical simulations. We compare the optimal profit maximizing behavior of the regulated monopoly with that of the unregulated monopoly and the competitive market in the controlled environment of the classic textbook setting with a monopoly that produces output using the Cobb-Douglas production function and faces a linear inverse demand function. Our results confirm the theoretical results, but also shed new light on the magnitude of the effects. The simulation results prove surprisingly robust to changes in the underlying parameter values and the functional form of the production function. While the economic literature traditionally emphasizes the distortionary effects of regulation, in defense of the Nordic revenue cap approach, we find that tolerating certain level of capital bias may be better than no regulation at all, especially if the opportunity cost of capital is very low.

The rest of this paper is organized as follows. To put the theoretical developments to an appropriate empirical context, Section 2 presents a brief overview of the Nordic revenue cap approach and its economic rationale. Section 3 presents a theoretical discussion of the inefficiency loss of the monopoly and the impacts of regulatory constraints following the classic model by Averch and Johnson (1962). Section 4 presents our numerical simulations. Robustness of the simulation results
is examined in Section 5. Section 6 discusses the policy implications of our results and some avenue for the future research. Further details of the numerical simulation are presented in the online supplement.

2. Nordic revenue cap

NordREG is an organization of the Nordic energy regulators with the stated aims to promote the development of efficient electricity markets in the Nordic area, consistent with the developments in the European Union. In their comprehensive methodological report NordREG (2011), the Nordic regulators refer to several regulation methods, including the rate of return, cost-of-service, price cap, revenue cap, yardstick regulation, performance standards, and earnings-sharing. Importantly, the Nordic regulators recognize that the real-world regulation schemes typically involve features from different approaches: “It is common to use combinations of these, for example a revenue cap regulation with inflation index and yardstick analysis in combination with bottom (minimum) and ceiling (maximum) rules on rate of return” NordREG (2011). Therefore, it is often very difficult, and potentially misleading, to try classify an incentive scheme implemented in a specific industry in a specific country under some academic label of a regulatory regime used in the literature.

NordREG (2011) report uses the term revenue cap to describe the incentive schemes applied in Denmark, Finland, Norway and Sweden. Similar use of the term can be found in academic journal articles (see, e.g., Kuosmanen et al., 2013; Saastamoinen et al., 2017 in this journal). Since the use of the term differs considerably from its usual meaning in the Anglo-American literature, it is useful to provide a stylized presentation of the main features of the Nordic revenue cap approach that is common to all four countries mentioned above. Practical details of implementation differ across countries and tend to change over time in each country (e.g., Agrell and Bogetoft, 2010), which makes it difficult to compare the regimes at a highly detailed level. In the following we present our interpretation of the economic rationale of the Nordic revenue cap approach.

As Bernstein and Sappington (1999) insightfully note, the purpose of regulation “is to replicate the discipline that market forces would impose on the regulated firm if they were present.” Therefore, the competitive market equilibrium serves as a useful analogue. It is well established in the microeconomic theory that market competition will eliminate any excess profits in the long run, driving the economic profit down to zero for all firms in the market equilibrium. Therefore in the competitive market equilibrium, we have

\[ \text{total revenue} = \text{variable cost} + \text{fixed cost}. \]  (1)
That is, the revenue is exactly equal to the variable cost (including wages, materials, services etc.) plus the fixed cost (i.e., the opportunity to cost of capital, including both equity and debt) in the competitive market equilibrium. Note that zero economic profit does not mean zero accounting profit: a normal return on equity is included as a part of the fixed component.

The economic rationale of the Nordic revenue cap approach is to mimic the competitive market equilibrium (1) by setting the revenue cap, defined here as the maximum level of revenue that a monopoly can generate, equal to the total cost that is considered acceptable. That is:

\[
\text{revenue cap} = \text{acceptable variable cost} + \text{acceptable fixed cost}
\] (2)

By controlling the total revenue of the monopoly firm, the regulation has direct effect on consumer prices that the monopoly firm can charge for its services. To set the appropriate level of the revenue cap, the regulator can use information and data from the input side to evaluate the firm’s total cost.

An important feature of the Nordic revenue cap approach is that a clear distinction between the fixed and variable costs are drawn in equation (2). In the terminology of regulators, the variable cost is often referred to as operational expenditures (OPEX), and the fixed cost is referred to as capital expenditures (CAPEX), but the basic idea is the same. In practice, the OPEX and CAPEX components are assessed separately.

To evaluate the acceptable level of fixed cost (CAPEX), the regulator needs to estimate the opportunity cost of the capital base. In finance, the capital asset pricing model (CAPM) is the most standard approach to determine a theoretically appropriate required rate of return of a risky asset. The Nordic energy regulators estimate the weighted average cost of capital (WACC), which builds on the basic idea of CAPM but also draws a distinction between equity and debt. To summarize, the acceptable fixed cost is generally calculated as the product of the regulatory capital base and WACC percentage. While there are some differences in the specification of the capital based and WACC across Nordic countries, in broad terms, a similar approach is applied in all four countries mentioned above, both in transmission and distribution industries.

To evaluate the acceptable level of variable cost (OPEX), most Nordic countries apply best-practice benchmarking in one form or another to provide incentives for efficient operation (e.g., Agrell and Bogetoft, 2011), but the methods and incentives different considerably both across countries and between the transmission and distribution sectors. We next illustrate the possibilities

---

3 It is worth noting that uncontrollable part of OPEX is often treated as a fixed cost. In addition to OPEX and CAPEX, the Nordic revenue cap typically includes additional components such as investment allowances and quality incentives. See NordREG (2011) for further discussion.

4 Benchmarking is also applied by many other regulators around the world (see, e.g., Jamasb and Pollitt, 2000; Bogetoft, and Otto, 2011).
by briefly reviewing the methods applied by the Finnish energy regulator in the electricity distribution sector, which is the industry that the authors are most familiar with, the development of which is documented in numerous journal articles to be cited below.

In the first regulation period (2005-2007), the acceptable level of OPEX was based on the benchmarking approach developed by Korhonen and Syrjänen (2003). The approach was built upon input-oriented data envelopment analysis (DEA) with a single input (total expenditure, TOTEX) and three outputs. The output variables introduced by Korhonen and Syrjänen are still in use today. The firm-specific DEA efficiency score computed based on TOTEX, which was subsequently translated to a monetary OPEX target. For the second regulation period (2008-2011), stochastic frontier analysis (SFA) based estimation of the efficiency score was used in parallel with DEA (Syrjänen, Agrell, and Bogetoft, 2007). The OPEX target was calculated based on the average of DEA and SFA efficiency percentages estimated based on TOTEX. The third regulation period (2012-2015) introduced major methodological reforms based on Kuosmanen (2012). The OPEX target was directly estimated from OPEX without dubious transformations. The OPEX target was estimated using stochastic nonparametric envelopment of data (StoNED) method (Kuosmanen and Kortelainen, 2012), which combines the attractive properties of DEA and SFA to a unified estimation framework. In the fourth regulation period (2016-2019), the OPEX target was estimated conditional on the capital stock in order to bridge the gap between OPEX incentives and CAPEX incentives, which had been treated separately in the prior regulation periods. The StoNED estimation was also developed further to incorporate panel data and model interruptions as an undesirable output (see Kuosmanen, 2018, for details). The same benchmarking approach will be applied in the fifth regulation period (2020-2023), only the parameters of the OPEX norm will be updated using the latest data from the fourth regulation period. To summarize the main point of the above discussion, energy regulators can systematically collect data of costs and outputs of all distribution firms, which offers the regulator some information advantage over a single firm that has only access to its own information. With the help of competent econometricians and operations researchers, such a rich panel data allows the regulator to set OPEX targets.

Now, what exactly are the key differences to the British revenue cap? In contrast to the Nordic revenue cap that defines the maximum level of revenue, the British revenue cap restricts the maximum rate of change in revenue (e.g. Mayer and Vickers, 1996; Jamison, 2007), similar to the price cap regulation (Littlechild, 1983). It might be tempting to conclude that the main difference concerns regulating the level or the growth rate of revenue, but such interpretation is too simplistic. In fact, imposing an ex ante cap for the level of revenue can be equivalently stated as a constraint for
the rate of revenue growth (see Agrell and Bogetoft, 2010, for further discussion about the *ex ante* versus *ex post* targets in the Nordic countries). While the stylized description of the Nordic revenue cap presented above was phrased in the static setting for the sake of simplicity, in fact, the implementation of the revenue cap is usually a dynamic process, which differs from one country to another. For example, the Finnish regulator allows the regulated monopolies temporarily exceed their revenue cap, but the surplus generated during a four-year regulation period must be compensated with interest during the next year regulation period.

From the strategic perspective, the key question concerns exogeneity of the revenue cap: whether the regulated monopoly can influence the cap by its own choices or not. We would argue that this is the key difference between the Nordic and the British revenue cap approaches. The Nordic revenue cap regulation is explicitly based on the CAPEX and OPEX of the regulated monopoly, whereas the British revenue cap, at least in its purest form, is exogenously given. To gain further insight, the next section revisits the Averch-Johnson critique and demonstrates that the Nordic revenue cap is certainly not immune to the capital bias, while some variants of the British revenue cap approach may also be affected.

3. **Theoretical model of monopoly**

3.1 *Inefficiency loss*

Following Averch and Johnson (1962), consider a monopoly firm that produces a single homogenous product \( y \) using a capital input \( x_1 \) and a variable input \( x_2 \) and production function \( y = f(x_1, x_2) \). We assume that the production function is monotonic increasing and strictly quasi-concave, which rules out the possibility of perfectly substitutable inputs \( x_1 \) and \( x_2 \). Facing a strictly decreasing inverse demand function \( p(y) \), the objective of the monopoly is to maximize profit

\[
\pi(x_1, x_2) = p(y) \cdot y - r_1 x_1 - r_2 x_2
\]

subject to

\[
y = f(x_1, x_2)
\]

where \( r_i, i = 1, 2 \), are the market prices of inputs, assumed as constant. Note that problem (3) is the long-run profit maximization problem of the monopoly where capital \( x_1 \) is treated as a variable input.

For the unregulated monopoly, the optimal input demands \( (x_1^u, x_2^u) \) satisfy

\[
\frac{\partial f(x_1^u, x_2^u)}{\partial x_i} = \frac{r_i}{p(y) + p'(y)y}, \quad i = 1, 2.
\]

That is, inputs are used until their marginal products are equal to their marginal cost. Note that in the case of a competitive firm taking price \( p \) as given, the marginal cost of input \( i \) on the right-hand
side of (4) would be \( r_i / p \). We refer to this hypothetical situation of market equilibrium with competitive price taking firms as the social optimum. In contrast, the unregulated monopoly takes the inverse demand into account in its optimization, and hence demands smaller amounts of inputs than an industry consisting of competitive firms. We assume that the inverse demand function is strictly decreasing (i.e., the law of demand holds), and hence the critical level of marginal product is higher than that of the social optimum, resulting as a smaller input use, a smaller output \( y \), and a higher price \( p \) than in the social optimum. This is the well-known *inefficiency loss of the monopoly* in welfare economics.

While the unregulated monopoly demands smaller amounts of inputs than a competitive industry in the market equilibrium, at least, it produces its output using the efficient input mix. To see this, note that the marginal rate of technical substitution (MRTS) is given by

\[
\frac{\partial f(x_1^*, x_2^*)}{\partial x_1} / \frac{\partial f(x_1^*, x_2^*)}{\partial x_2} = \frac{r_1 / p(y) + p'(y)y}{r_2 / p(y) + p'(y)y} = r_1 / r_2.
\]

We will henceforth refer to the input demands where the MRTS is equal to the relative market price of inputs as the socially optimal input mix.

### 3.2 Rate of return regulation and Averch-Johnson critique

In the rate of return regime, the regulator caps the rate of return on capital to avoid excessive monopoly profit. Formally, the regulated rate of return is denoted by \( s_1 \), and the regulatory constraint can be stated as

\[
\frac{p(y) \cdot y - r_i x_2}{x_1} \leq s_1.
\]

This constraint prevents the monopoly from overcharging its customers while still allowing the monopoly to operate at fair rates of return after subtracting all the costs and expenses (e.g., Baumol and Klevorick, 1970).

Suppose the regulated monopoly maximizes profit (problem (3)) subject to the rate of return constraint (5). Unfortunately, the true opportunity cost of capital \( r_1 \) is not directly observable. In practice, the regulators must resort to empirical estimates based on WACC, as discussed in Section 2. Note that if the regulated rate of return is set below the market price \( (s_1 < r_i) \) then the optimal response of the regulated monopoly is to shut down the operation. In practice, the regulators are afraid to drive the regulated monopoly to bankruptcy, so it is standard to assume that the regulated rate of return exceeds the market price, that is, \( s_1 > r_i \).
Examining the Kuhn-Tucker conditions of the resulting constrained optimization problem, Averch and Johnson (1962) show that the regulated monopoly responds to the revenue cap constraint by increasing the use of capital input relative to that of the variable input. That is, if the regulation constraint is binding, the optimal input demands of the regulated monopoly, denoted by \((x'_1, x'_2)\), satisfy the following inequality,

\[
\frac{x'_1}{x'_2} > \frac{x''_1}{x''_2}.
\]

Based on this result, the tendency of regulated companies to engage in excessive amounts of capital accumulation in order to expand the volume of their profits is known as the Averch-Johnson effect.

### 3.3 Averch-Johnson effect in the Nordic revenue cap

Since the Averch-Johnson effect is strongly associated with the rate of return regulation, it is worth to examine whether and to what extent the Nordic revenue cap approach discussed in Section 2 is subject to a similar capital bias. To this end, note that the Nordic revenue cap (2) can be stated in terms of the Averch-Johnson model as the following inequality constraint

\[
p(y) \cdot y \leq \bar{R},
\]

where \(\bar{R}\) is the maximum acceptable revenue imposed by the regulator. Importantly, if the revenue cap \(\bar{R}\) is not an exogenously given constant, but a function \(\bar{R}(x_1)\) that depends on the capital input \(x_1\) of the monopoly, then the Averch-Johnson critique applies equally well to revenue cap regulation.

**Lemma 1:** If the regulator applies the Nordic revenue cap specifying the acceptable total cost as

\[
\bar{R} = s_1 x_1 + r_2 x_2,
\]

then the revenue cap (6) is directly equivalent to the rate of return constraint (5). In this case, the Nordic revenue cap regulation is subject to the Averch-Johnson effect.

Proof. Note that inequality (5) can be reorganized in the form of the revenue cap as follows

\[
p(y) \cdot y \leq s_1 x_1 + r_2 x_2.
\]

(7)

If the revenue cap \(\bar{R}\) depends on the capital input \(x_1\) of the monopoly, as stated in Lemma 1, then the revenue cap is equivalent to the rate of return regulation. \(\square\)
This result demonstrates that the Nordic revenue cap regulation is by no means immune to the capital bias known as the Averch-Johnson effect. As discussed in Section 2, it is a standard approach in the Nordic revenue cap approach to specify the revenue cap $\bar{R}$ based on the observed asset base $x_1$ and the WACC rate $\delta_1$. To alleviate the capital bias, the regulator can draw a distinction between the accounting assets of the monopoly firm and the regulatory asset base. While the monopoly controls its accounting assets, the regulatory asset base can be controlled by so-called “command & control” measures either ex ante or ex post, possibly both. An example of ex ante control mechanism is to require the regulated monopoly firm to apply for regulator’s approval before making significant new investments to the power grid. Any investment that was not approved by the regulator would not be counted in the regulatory asset base. Another example of ex ante controls is to specify in advance which type of network investments are acceptable to the regulatory asset base. If the regulator suspects manipulation of the asset base (referred to as gold-plating), it could also apply ex post revisions to the regulatory asset base. If the regulator can prove that the monopoly has invested in gold-plating to take advantage of the regulation premium, it could remove such investment from the regulatory asset base, and hence leave those investments uncompensated. Clearly communicated ex ante controls and a threat of ex post revisions to the regulatory asset base can at least dampen if not completely eliminate the Averch-Johnson effect.

While the preceding discussion focuses on the fixed cost component, the Nordic revenue cap also includes the variable cost component $r_2 x_2$. In Lemma 1, we implicitly assumed that the regulator has perfect information about the variable cost. Of course, in reality the variable cost must also be estimated from empirical data, as discussed in Section 2. Interestingly, underestimation of the variable cost part would further enhance the Averch-Johnson effect, while overestimation of the variable cost would have the opposite effect. Since the regulators usually prefer a conservative approach to estimate the variable cost, allowing the regulated monopoly a certain degree of the benefit of doubt, the variable cost estimates are likely upward biased, which would help to offset the capital bias at least to some extent.

### 3.4 Averch-Johnson effect in the British revenue cap?

While the main focus of this paper is on the Nordic revenue cap, we conclude this section with a brief discussion about the possibility of the Averch-Johnson effect in the British revenue cap. The immunity of the RPI-X style price and revenue caps to capital bias is frequently cited as one of the main advantages of the approach (see, e.g., Mayer and Vickers, 1996; Jamison, 2007). But even if the inflation component is beyond the control of the regulated monopoly, but it is often far from self-
evident how the X-factor, often understood as total factor productivity (TFP) of the industry, should be measured in practice (see, e.g., Bernstein and Sappington, 1999, for an insightful discussion). In theory, if the X-factor does not depend on the choices of the monopoly firm, then the British revenue cap is indeed immune to the Averch-Johnson effect. In practice, however, exogenous specification of the TFP target without any connection to the historical performance of the regulated monopoly can be challenging. For example, if the industry consists of a single monopoly firm, as is often the case in the electricity transmission, then the TFP of the industry is clearly dependent on the production decisions of the monopoly. In practice, some regulators resort to the WACC approach to set the X-factor in the RPI-X regime.5

The British price cap and revenue cap regimes are historically geared towards protecting customers from price increase, originally developed in the context of privatizing the public utilities in the U.K (Littlechild, 1983). However, if the firm extracted monopoly profit already before the introduction of price controls, then protecting the customer from price increase only maintains the status quo, but does not suffice to eliminate the market failure. If the purpose of the X-factor is to gradually eliminate the monopoly profit, then the specification of the X-factor must somehow dependent on efficiency of the monopoly relative to some well-defined benchmark, whatever definition of efficiency and benchmark might be used. But efficiency assessment and benchmarking cannot be conducted in isolation of the operational decisions of the monopoly. The point is, if the X-factor is conditional on observed performance of the regulated monopoly, the conceptual distinctions between the Nordic and British revenue caps and the U.S.-style rate of return regulation become blurred, and the Averch-Johnson effect can creep in to the British price cap and revenue cap regimes.

4. Numerical simulations
4.1. Motivation
Having established that the Averch-Johnson effect remains relevant to the Nordic revenue cap and possibly also to the British revenue cap and price cap regimes if the X-factor is endogenous, it is worth to ask how much harm does the capital bias cause in the case of a divested transmission or distribution monopoly that is by law not allowed to operate in retail energy or other competitive markets?6 Is the regulation of transmission and distribution monopolies actually causing more

---

5 One example is the UK water sector where Ofwat applies the WACC approach in setting the X-factor (see, e.g., Ofwat, 2019).
6 Averch and Johnson (1962) extensively discuss the case of a multi-product firm that operates as a regulated monopoly in one market but also sells other commodity in a competitive market. In this paper we focus on a single product firm, noting that the competition law can prevent the regulated monopoly from distortionary participation in other markets.
harmful than leaving these sectors completely unregulated? While Averch and Johnson (1962) focus on examining the undesirable impact of regulation on the input mix of the regulated firm, they also note in passing that the regulation increases output $y$ in their model, except for the special case of a linear production function. They discuss the impact of production function, particularly the substitutability of inputs, on how effectively the regulator can force the monopoly to increase its output as follows:

“The extent to which regulation affects output depends upon the nature of the production function. If it involves fixed proportions, the regulated firm is constrained to the efficient expansion path... If the production function is linear and if the iso-output curves have a slope equal to $-r_1/r_2$, the firm could substitute $x_1$ for $x_2$ and, with no change in marginal rate of substitution, hold output constant. In this case it could remain at the unregulated monopoly output.” (Averch and Johnson, 1962, p. 1057)

While the subsequent literature emphasizes the distortionary effects of the rate of return regulation, a closer examination of the Averch-Johnson model reveals that such theoretically imperfect regulation can nevertheless help to achieve many of its intended goals. While the Averch-Johnson effect has had a major impact on economic thinking in general and the regulation practice specifically, there is surprisingly little quantitative evidence concerning the magnitude of this effect. One reason is that production data of monopoly firms, whether regulated or not, are usually not available for research purposes. To shed further light on the magnitude of the Averch-Johnson effect and the possible benefits of the rate of return regulation, in this study we resort to numerical simulations.

4.2. Simulation model
In this section we examine the behavior of a regulated monopoly subject to the Nordic revenue cap (or equivalently, rate of return regulation), an unregulated monopoly, and the competitive market in the controlled environment of numerical simulations. We assume the most standard textbook case of the linear demand function and the Cobb-Douglas production function to analyze the impacts of regulation parameter $s_1$ (regulated rate of return), demand elasticity (slope of demand function), substitution elasticity (production function) on the following outcomes:

- output $y$,
- market price $p(y)$,
- total revenue $p(y)y$,
- monopoly profit $p(y)y - s_1x_1 - r_2x_2$,
- consumer surplus (CS) $y(A - p(y))/2$, and
capital intensity $x_1/x_2$.

More specifically, the Cobb-Douglas production function is defined as:

$$y = f^{CD}(x_1, x_2) = x_1^{\beta_1} x_2^{\beta_2}$$

Where $\beta_1, \beta_2$ are the output elasticities of inputs. In the baseline scenario, we set $\beta_1 = \beta_2 = 0.5$, which implies constant returns to scale in production. We examine the sensitivity of results on changing the output elasticities as well as the use of the more general constant elasticity of substitution (CES) functional form in Section 5.

The linear inverse demand function is defined as

$$p(y) = A - by,$$

where parameters $A$ and $b$ are constant. In the baseline scenario, we set $A = 20$, and $b = 1$, which implies that the price elasticity of demand is simply $-p/y$. The baseline scenario hence approximates the unit elastic demand. The sensitivity of results on changing the slope $b$ that governs the price elasticity of demand will be examined in Section 5.

In our simulations, the competitive market equilibrium is computed by maximizing the output subject to the constraint that profit must be non-negative: note that positive profit and free entry would attract new firms to the market, which increases the total output until the economic profit approaches to zero and the equilibrium is reached. The unregulated firm solves the profit maximizing problem (3), taking the input prices $r_1 = r_2 = 1$ as given.\(^7\) The regulated firm solves the same profit maximization problem subject to constraint (7), which is equivalent to the rate of return constraint (5): the profit maximizing regulated monopoly behaves exactly the same way irrespective of the regulator’s formulation of the constraint as the Nordic revenue cap or the rate of return constraint. Different values of the regulated rate of return $\kappa_1$ are considered, ranging between 1.02 and 10.

In all three cases, analytical solution of the resulting nonlinear programming problem proves rather tedious (see Averch and Johnson, 1962, for the explicit presentation of the Kuhn-Tucker conditions). Since the problem involves only two unknowns $(x_1, x_2)$, the numerical solution is computationally rather trivial. To facilitate broader applicability of the numerical simulations reported in this paper among practitioners (including management consultants and regulators), in this study, the profit maximization problem of the regulated firm, the unregulated firm, and the

\(^7\) Large scale capital bias in one industry would likely affect both capital and labor markets, and hence influence other industries as well. Extending the welfare analysis to the general equilibrium setting seems an interesting avenue for future research, but it would also require a large number of additional assumptions far beyond the original Averch and Johnson (1962) model. Since the original argument was presented in the partial equilibrium setting, we consider it important and relevant to examine the welfare effects within the same setting.
competitive firm has been set up in an Excel spreadsheet, and solved using the standard Excel solver; see the online supplement for further details.

4.3. Results

We first examine how a profit maximizing regulated monopoly subject to the Nordic revenue cap responds to changes in the maximum rate of return $s_1$. Table 1 reports the results on the outcomes of the regulated monopoly. In the left-most column of Table 1, the regulation parameter $s_1$ increases from 1.02 to 10: the smallest value of $s_1$ corresponds to heavy handed regulation where the regulator enforces the acceptable rate of return very close to the true opportunity cost of capital, and the largest value of $s_1$ correspond to cosmetic regulation where the regulation constraint is not binding. That is, the case of $s_1 = 10$ is equivalent to the optimal solution of the unregulated firm: regulation constraint becomes redundant when $s_1$ is large enough.

Table 1: Simulation results for the regulated monopoly: the effect of regulation parameter $s_1$ on output, price, total revenue, monopoly profit, consumer surplus, and capital intensity.

<table>
<thead>
<tr>
<th>$s_1$</th>
<th>output</th>
<th>price</th>
<th>total revenue</th>
<th>monopoly profit</th>
<th>consumer surplus</th>
<th>capital intensity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.02</td>
<td>9.90</td>
<td>10.10</td>
<td>99.99</td>
<td>1.94</td>
<td>48.99</td>
<td>96.11</td>
</tr>
<tr>
<td>1.05</td>
<td>9.90</td>
<td>10.10</td>
<td>99.99</td>
<td>4.71</td>
<td>48.97</td>
<td>90.68</td>
</tr>
<tr>
<td>1.25</td>
<td>9.88</td>
<td>10.12</td>
<td>99.99</td>
<td>19.75</td>
<td>48.83</td>
<td>63.90</td>
</tr>
<tr>
<td>1.50</td>
<td>9.85</td>
<td>10.15</td>
<td>99.98</td>
<td>32.83</td>
<td>48.51</td>
<td>44.44</td>
</tr>
<tr>
<td>1.75</td>
<td>9.83</td>
<td>10.17</td>
<td>99.97</td>
<td>42.11</td>
<td>48.28</td>
<td>32.64</td>
</tr>
<tr>
<td>2.00</td>
<td>9.80</td>
<td>10.20</td>
<td>99.96</td>
<td>49.00</td>
<td>47.98</td>
<td>25.02</td>
</tr>
<tr>
<td>2.25</td>
<td>9.78</td>
<td>10.22</td>
<td>99.95</td>
<td>54.31</td>
<td>47.79</td>
<td>19.75</td>
</tr>
<tr>
<td>10.00</td>
<td>9.01</td>
<td>10.99</td>
<td>99.01</td>
<td>81.00</td>
<td>40.56</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Our baseline scenario confirms that the output increases and the price decreases as the rate of return regulation becomes more stringent. As a result, the consumer surplus increases. While the total revenue increases modestly, the monopoly profit decreases. However, the simulation also confirms the Averch-Johnson effect: the capital intensity dramatically increases as the monopoly responds to the regulation.

It might be tempting to compare the large decrease of monopoly profit with the modest increase in consumer surplus, but such a comparison is misleading from the welfare point of view. Note that the monopoly profit here refers to $(s_1 - r_1)x_1$, which is the additional return on capital exceeding the opportunity cost. In our terminology, the monopoly profit is the same as the economic profit, whereas
the accounting profit is the sum of monopoly profit and the market return on capital. From the welfare point of view, the total revenue is a more relevant measure of producer surplus in these simulations because, in the absence of intermediate inputs, the total revenue is here equal to the total wage expenditure plus accounting profit paid to the investor.  

Table 2 presents the same information in relative terms, as percentage relative to the optimal solution of the unregulated monopoly. In the case of \( s_1 = 10 \), the regulation has no effect, and the outcomes of the regulated monopoly are exactly the same as those of the unregulated monopoly. In the case of heavy handed regulation, setting \( s_1 = 1.02 \), the output of the regulated monopoly is 10 percent higher than that of the unregulated monopoly, and the price is 8 percent lower. As a result, the consumer surplus is 21 percent higher than in the absence of regulation.

Table 2: Comparison of the regulated vs unregulated monopoly: the percentage of the regulated monopoly’s outcomes relative to that of the unregulated monopoly as a function of parameter \( s_1 \).

<table>
<thead>
<tr>
<th>( s_1 )</th>
<th>output</th>
<th>price</th>
<th>total revenue</th>
<th>monopoly profit</th>
<th>consumer surplus</th>
<th>capital intensity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.02</td>
<td>110 %</td>
<td>92 %</td>
<td>101 %</td>
<td>2 %</td>
<td>121 %</td>
<td>9611 %</td>
</tr>
<tr>
<td>1.05</td>
<td>110 %</td>
<td>92 %</td>
<td>101 %</td>
<td>6 %</td>
<td>121 %</td>
<td>9068 %</td>
</tr>
<tr>
<td>1.25</td>
<td>110 %</td>
<td>92 %</td>
<td>101 %</td>
<td>24 %</td>
<td>121 %</td>
<td>6390 %</td>
</tr>
<tr>
<td>1.50</td>
<td>109 %</td>
<td>92 %</td>
<td>101 %</td>
<td>41 %</td>
<td>120 %</td>
<td>4444 %</td>
</tr>
<tr>
<td>1.75</td>
<td>109 %</td>
<td>92 %</td>
<td>101 %</td>
<td>52 %</td>
<td>119 %</td>
<td>3264 %</td>
</tr>
<tr>
<td>2.00</td>
<td>109 %</td>
<td>93 %</td>
<td>101 %</td>
<td>60 %</td>
<td>118 %</td>
<td>2502 %</td>
</tr>
<tr>
<td>2.25</td>
<td>109 %</td>
<td>93 %</td>
<td>101 %</td>
<td>67 %</td>
<td>118 %</td>
<td>1975 %</td>
</tr>
<tr>
<td>10.00</td>
<td>100 %</td>
<td>100 %</td>
<td>100 %</td>
<td>100 %</td>
<td>100 %</td>
<td>100 %</td>
</tr>
</tbody>
</table>

Interestingly, Tables 1 and 2 demonstrate that the consumer benefits from regulation even if the implementation is relatively light handed. Even if the regulated rate of return is two times higher than the opportunity cost of capital, as in the case of \( s_1 = 2 \), the consumer is much better off than in the absence of regulation. Decreasing \( s_1 \) below this level affects the monopoly profit and the capital intensity, but provides only rather marginal benefits to the consumer in our baseline scenario. In practice, the regulator does not know the true opportunity cost \( r_1 \): setting \( s_1 \) too low (below \( r_1 \)) involves a risk that the regulated monopoly exits the market. Therefore, we consider it a rather

---

8 Mishan (1968) presents sharp critique of the Marshallian notion of producer’s surplus, recommending that this term “be struck from economist’s vocabulary.” In line with Mishan, we exclude the producer’s surplus from our analysis, but suggest that the welfare comparisons of our simulation results could be based on the sum of consumer surplus and the total revenue. Note that if the variable input consists of labor and does not include any intermediate inputs, then the total revenue of the monopoly is equal to the value added of the industry.
encouraging result that the regulator does not need to drive the monopoly to bankruptcy in order to benefit the consumer.

To put the results of Tables 1 and 2 in perspective, we also compared the outcomes of the regulated monopoly with those achieved in an ideal case of the perfectly competitive market. Table 3 presents the outcomes of the regulated monopoly in relative terms against the benchmark of the competitive market. Note: the comparison of profit is excluded as the economic profit is equal to zero in the competitive market equilibrium.

Table 3: Comparison of the regulated competitive market: the percentage of the regulated monopoly’s outcomes relative to that of the unregulated monopoly as a function of parameter $s_1$.

<table>
<thead>
<tr>
<th>$s_1$</th>
<th>output</th>
<th>price</th>
<th>total revenue</th>
<th>consumer surplus</th>
<th>capital intensity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.02</td>
<td>55 %</td>
<td>505 %</td>
<td>278 %</td>
<td>30 %</td>
<td>9611 %</td>
</tr>
<tr>
<td>1.05</td>
<td>55 %</td>
<td>505 %</td>
<td>278 %</td>
<td>30 %</td>
<td>9068 %</td>
</tr>
<tr>
<td>1.25</td>
<td>55 %</td>
<td>506 %</td>
<td>278 %</td>
<td>30 %</td>
<td>6390 %</td>
</tr>
<tr>
<td>1.50</td>
<td>55 %</td>
<td>507 %</td>
<td>278 %</td>
<td>30 %</td>
<td>4444 %</td>
</tr>
<tr>
<td>1.75</td>
<td>55 %</td>
<td>509 %</td>
<td>278 %</td>
<td>30 %</td>
<td>3264 %</td>
</tr>
<tr>
<td>2.00</td>
<td>54 %</td>
<td>510 %</td>
<td>278 %</td>
<td>30 %</td>
<td>2502 %</td>
</tr>
<tr>
<td>2.25</td>
<td>54 %</td>
<td>511 %</td>
<td>278 %</td>
<td>30 %</td>
<td>1975 %</td>
</tr>
<tr>
<td>10.00</td>
<td>50 %</td>
<td>550 %</td>
<td>275 %</td>
<td>25 %</td>
<td>100 %</td>
</tr>
</tbody>
</table>

We consider the results of Table 3 as a sobering reminder that the regulated monopoly remains inefficient compared to the competitive market. In the absence of regulation, the monopoly produces only 50 percent of the output in the competitive market. By imposing rate of return regulation, the output increases up to 55 percent. As a result, the price decreases, and the consumer surplus increases. Still, the results of a regulated monopoly are far from the ideal case of the competitive market. Table 3 aptly illustrates that even if the regulator sets parameter $s_1$ very close to the true opportunity cost, the performance of the regulated monopoly does not converge to that of the competitive market.

In conclusion, the results of the numerical simulation presented in Tables 1-3 demonstrate that the Nordic revenue cap and the rate of return regulation can achieve most of their goals: the output increases, the price decreases, the consumer surplus increases, and the monopoly profit decreases. In the absence of better regulatory instruments, imperfect regulation is far better than no regulation at all. Interestingly, most benefits to the consumer are achieved already with relatively light handed regulation, it is not necessary to force the monopoly close to bankruptcy. However, the simulations
also clearly reveal that the result of regulation is very far from the social optimum. The magnitude of the Averch-Johnson effect is rather dramatic in our baseline scenario.

The simulations were conducted using the most standard textbook case of the linear demand function and the Cobb-Douglas production function. The robustness of these findings to different parameter values and the CES production function will be explored in the next section.

5. Sensitivity analysis

5.1. Changes in demand and production parameters

Since the specific parameter values used in the baseline scenario are rather arbitrary, it is worth to examine how robust the insights gained in Section 4 are to changes in the parameter values. Obviously, the specific parameter values influence the absolute level of outcomes, but we are more interested in the relative performance of the regulated monopoly compared to the unregulated monopoly and the competitive market. The relative performance turns out to be highly robust to changes in the parameter values.

Let us first examine the impact of changes in the key parameters of the inverse demand function and the production function in the case of no government interference. Table 4 reports the relative outcomes of the unregulated monopoly as percentage of those of the competitive market. The parameter values of the baseline scenario reported in Section 3 are highlighted in bold font.

On the first seven rows of Table 4, we increase the slope \( b \) of the inverse demand function from 0.25 to 2 (the baseline value is 1), keeping all other parameter values at their baseline levels. Interestingly, such changes in the slope of the demand function have no impact whatsoever on the inefficiency loss of the unregulated monopoly: the monopoly produces only half of the output of the competitive market, charging 5.5 times higher price.

In the bottom part of Table 4 we adjust the output elasticity parameters of the Cobb-Douglas production function, again, keeping all other parameters at their baseline levels. We find that the price and the total revenue are critically dependent on the production function. Note that in the absence of regulation the effects are symmetric between the fixed and variable input. Interestingly, the output elasticity parameters have virtually no impact on the output, the consumer surplus, or the capital intensity of the unregulated monopoly in comparison to the competitive market.

Table 4: Comparison of the unregulated monopoly vs competitive market: the percentage of the unregulated monopoly’s outcomes compared to that of the competitive market as a function of demand parameter \( b \), output elasticity of capital \( \beta_1 \), and output elasticity of labor \( \beta_2 \).
Tables 5 and 6 present analogous results for the relative comparison of the regulated monopoly and the unregulated monopoly. The percentages indicate the outcome of the regulated monopoly relative to the unregulated monopoly. In Table 5, the regulation parameter is set as $s_1 = 1.25$, while in Table 6 we set $s_1 = 1.5$. Otherwise Tables 5 and 6 are constructed in the same way as Table 4.

The top rows of Tables 5 and 6 report the impact of adjusting the slope $b$ of the inverse demand function. We find that changes in the slope of the demand function have no impact whatsoever on the relative efficiency of regulation. These findings illustrate that the insights of Section 3 are surprisingly robust to the changes in the demand function. While we here restrict attention to the linear demand function, the robustness of outcomes to major changes in the slope $b$ suggests that nonlinear demand function would not dramatically affect the results.

Table 5: Comparison of the regulated monopoly vs unregulated monopoly: the percentage of the regulated monopoly’s outcomes compared to that of the unregulated monopoly as a function of demand parameter $b$, output elasticity of capital $\beta_1$, and output elasticity of labor $\beta_2$. The regulation parameter is set as $s_1 = 1.25$.

<table>
<thead>
<tr>
<th>parameter</th>
<th>output</th>
<th>price</th>
<th>total revenue</th>
<th>monopoly profit</th>
<th>consumer surplus</th>
<th>capital intensity</th>
</tr>
</thead>
<tbody>
<tr>
<td>$b$</td>
<td>0.25</td>
<td>110 %</td>
<td>92 %</td>
<td>24 %</td>
<td>120 %</td>
<td>64 %</td>
</tr>
<tr>
<td></td>
<td>0.50</td>
<td>101 %</td>
<td>92 %</td>
<td>24 %</td>
<td>120 %</td>
<td>64 %</td>
</tr>
<tr>
<td></td>
<td>0.75</td>
<td>101 %</td>
<td>92 %</td>
<td>24 %</td>
<td>120 %</td>
<td>64 %</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>0.10</td>
<td>49 %</td>
<td>174 %</td>
<td>85 %</td>
<td>24 %</td>
<td>100 %</td>
</tr>
<tr>
<td></td>
<td>0.25</td>
<td>49 %</td>
<td>266 %</td>
<td>131 %</td>
<td>24 %</td>
<td>100 %</td>
</tr>
<tr>
<td></td>
<td>0.50</td>
<td>50 %</td>
<td>550 %</td>
<td>275 %</td>
<td>25 %</td>
<td>100 %</td>
</tr>
<tr>
<td></td>
<td>0.75</td>
<td>50 %</td>
<td>964 %</td>
<td>484 %</td>
<td>25 %</td>
<td>101 %</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>0.10</td>
<td>49 %</td>
<td>174 %</td>
<td>85 %</td>
<td>24 %</td>
<td>100 %</td>
</tr>
<tr>
<td></td>
<td>0.25</td>
<td>49 %</td>
<td>266 %</td>
<td>131 %</td>
<td>24 %</td>
<td>100 %</td>
</tr>
<tr>
<td></td>
<td>0.50</td>
<td>50 %</td>
<td>550 %</td>
<td>275 %</td>
<td>25 %</td>
<td>100 %</td>
</tr>
<tr>
<td></td>
<td>0.75</td>
<td>50 %</td>
<td>964 %</td>
<td>484 %</td>
<td>25 %</td>
<td>99 %</td>
</tr>
</tbody>
</table>
The bottom parts of Tables 5 and 6 report the impacts of adjusting output elasticities of the capital input and the variable input, respectively. Recall from Table 4 that the effects were symmetrical under no regulation, but this is no longer the case under the rate of return regulation, as evident from Tables 5 and 6. Indeed, the impacts of regulation critically depend on the production function. In the case of the Cobb-Douglas production function, the scale elasticity is equal to the sum of the output elasticities, that is, $\beta_1 + \beta_2$. Another property of the Cobb-Douglas production function is that the elasticity of substitution between the two inputs is always equal to one, irrespective of the parameter values. For these two reasons, the changes in parameters $\beta_1, \beta_2$ reported in Tables 5 and 6 reflect more the ability of the monopoly to utilize economies of scale than its input substitution possibilities.

The results of Tables 5 and 6 suggest that the rate of return regulation is most effective if the technology exhibits decreasing returns to scale, and its effect deteriorates under increasing returns to scale. Of course, presence of increasing returns may be the reason why a natural monopoly exists in the first place.

Table 6: Comparison of the regulated monopoly vs unregulated monopoly the percentage of the regulated monopoly’s outcomes compared to that of the unregulated monopoly as a function of demand parameter $b$, output elasticity of capital $\beta_1$, and output elasticity of labor $\beta_2$. The regulation parameter is set as $s_1 = 1.5$. 

<table>
<thead>
<tr>
<th>parameter</th>
<th>output price</th>
<th>total revenue</th>
<th>monopoly profit</th>
<th>consumer surplus</th>
<th>capital intensity</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_1$</td>
<td>0.10</td>
<td>119 % 110 % 92 % 26 % 142 % 13 %</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.25</td>
<td>119 % 105 % 89 % 27 % 141 % 15 %</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>0.50</strong></td>
<td><strong>110 % 101 % 92 % 24 % 121 % 64 %</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.75</td>
<td>105 % 100 % 95 % 23 % 110 % 381 %</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>0.10</td>
<td>159 % 121 % 76 % 37 % 253 % 11 %</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.25</td>
<td>129 % 106 % 83 % 30 % 166 % 28 %</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>0.50</strong></td>
<td><strong>110 % 101 % 92 % 24 % 120 % 64 %</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.75</td>
<td>104 % 100 % 96 % 22 % 109 % 102 %</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
5.2. CES production function

The capital bias of regulation critically depends on the ability of the monopoly to substitute labor by capital. As Averch and Johnson (1962) note, the regulation would increase output without distorting capital intensity in the special case of the fixed proportions technology, known as the Leontief production function. By contrast, the regulated monopoly could eliminate variable input completely by capital, keeping the output at the unregulated level, if the production function is linear. Since the effectiveness of regulation critically depend on substitutability of inputs, and because the Cobb-Douglas production function has the substitution elasticity equal to one at all parameter values, we next extend the simulations to the more general CES production function that allows us to control for the input substitutability.

In the following simulations, the CES production function is specified as

\[ y = f^{CES}(x_1, x_2) = (0.5x_1^\gamma + 0.5x_2^\gamma)^{1/\gamma} \]

where parameter \( \gamma \) determines the substitution elasticity (\( \sigma \)) between capital and the variable input. Specifically,

\[ \sigma = \frac{1}{1-\gamma}. \]

Note that the CES function converges to the Cobb-Douglas functional form used in our previous simulations as parameter \( \gamma \) approaches to zero.
Directly analogous to Table 2, in Tables 7 – 9 we compare the outcomes of the regulated monopoly relative to the unregulated monopoly as the regulation parameter $s_1$ increases, except now we use the CES production function instead of the Cobb-Douglas. While the Cobb-Douglas production function has the substitution elasticity equal to one by construction, in Table 7 the substitution elasticity is set at 0.5, in Table 8 the substitution elasticity is further decreased to 0.25, and in Table 9 the substitution elasticity goes down to 0.02.

Table 7: Comparison of the regulated vs unregulated monopoly: the percentage of the regulated monopoly’s outcomes relative to that of the unregulated monopoly as a function of parameter $s_1$. The CES production function specified as $\gamma = -1, \sigma = 0.5$.

<table>
<thead>
<tr>
<th>$s_1$</th>
<th>output</th>
<th>price</th>
<th>total revenue</th>
<th>monopoly profit</th>
<th>consumer surplus</th>
<th>capital intensity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.02</td>
<td>108 %</td>
<td>93 %</td>
<td>101 %</td>
<td>2 %</td>
<td>117 %</td>
<td>1812 %</td>
</tr>
<tr>
<td>1.10</td>
<td>108 %</td>
<td>93 %</td>
<td>101 %</td>
<td>11 %</td>
<td>117 %</td>
<td>1672 %</td>
</tr>
<tr>
<td>1.20</td>
<td>108 %</td>
<td>93 %</td>
<td>101 %</td>
<td>19 %</td>
<td>117 %</td>
<td>1525 %</td>
</tr>
<tr>
<td>1.30</td>
<td>108 %</td>
<td>94 %</td>
<td>101 %</td>
<td>27 %</td>
<td>116 %</td>
<td>1400 %</td>
</tr>
<tr>
<td>1.40</td>
<td>108 %</td>
<td>94 %</td>
<td>101 %</td>
<td>33 %</td>
<td>116 %</td>
<td>1293 %</td>
</tr>
<tr>
<td>1.50</td>
<td>108 %</td>
<td>94 %</td>
<td>101 %</td>
<td>39 %</td>
<td>116 %</td>
<td>1200 %</td>
</tr>
<tr>
<td>2.00</td>
<td>108 %</td>
<td>94 %</td>
<td>101 %</td>
<td>58 %</td>
<td>116 %</td>
<td>875 %</td>
</tr>
<tr>
<td>5.00</td>
<td>106 %</td>
<td>95 %</td>
<td>101 %</td>
<td>92 %</td>
<td>113 %</td>
<td>291 %</td>
</tr>
<tr>
<td>10.00</td>
<td>100 %</td>
<td>100 %</td>
<td>100 %</td>
<td>100 %</td>
<td>100 %</td>
<td>100 %</td>
</tr>
</tbody>
</table>

By comparing the results of Tables 7 – 9 with those of Table 2, we find that the substitution elasticity has surprisingly little effect on behavior of the regulated monopoly. One would expect that the Averch-Johnson effect on capital intensity diminishes as the substitution elasticity decreases, and this is indeed confirmed by our simulations. But somewhat unexpectedly, the desirable impacts on output, price, and consumer surplus are also affected by the substitution elasticity. Based on the extreme special cases of the linear and the Leontief production functions, one might be tempted to assume that rate of return regulation becomes more effective as the substitution elasticity of inputs decreases, but our simulations show that this assumption is not generally true. By comparing Tables 7 – 9 and Table 2, we see that the regulation decreases price and increases output and consumer surplus most effectively when the substitution elasticity is high, and the welfare impacts decrease as the substitution elasticity goes down.
Table 8: Comparison of the regulated vs unregulated monopoly: the percentage of the regulated monopoly’s outcomes relative to that of the unregulated monopoly as a function of parameter $s_1$. The CES production function specified as $\gamma = -3, \sigma = 0.25$.

<table>
<thead>
<tr>
<th>$s_1$</th>
<th>output</th>
<th>price</th>
<th>total revenue</th>
<th>monopoly profit</th>
<th>consumer surplus</th>
<th>capital intensity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.02</td>
<td>107 %</td>
<td>95 %</td>
<td>101 %</td>
<td>2 %</td>
<td>114 %</td>
<td>1186 %</td>
</tr>
<tr>
<td>1.10</td>
<td>107 %</td>
<td>95 %</td>
<td>101 %</td>
<td>10 %</td>
<td>114 %</td>
<td>1100 %</td>
</tr>
<tr>
<td>1.20</td>
<td>107 %</td>
<td>95 %</td>
<td>101 %</td>
<td>19 %</td>
<td>114 %</td>
<td>1008 %</td>
</tr>
<tr>
<td>1.30</td>
<td>107 %</td>
<td>95 %</td>
<td>101 %</td>
<td>26 %</td>
<td>114 %</td>
<td>930 %</td>
</tr>
<tr>
<td>1.40</td>
<td>107 %</td>
<td>95 %</td>
<td>101 %</td>
<td>33 %</td>
<td>114 %</td>
<td>864 %</td>
</tr>
<tr>
<td>1.50</td>
<td>107 %</td>
<td>95 %</td>
<td>101 %</td>
<td>38 %</td>
<td>114 %</td>
<td>806 %</td>
</tr>
<tr>
<td>2.00</td>
<td>107 %</td>
<td>95 %</td>
<td>101 %</td>
<td>57 %</td>
<td>114 %</td>
<td>604 %</td>
</tr>
<tr>
<td>5.00</td>
<td>106 %</td>
<td>95 %</td>
<td>101 %</td>
<td>91 %</td>
<td>113 %</td>
<td>237 %</td>
</tr>
<tr>
<td>10.00</td>
<td>100 %</td>
<td>100 %</td>
<td>100 %</td>
<td>100 %</td>
<td>100 %</td>
<td>100 %</td>
</tr>
</tbody>
</table>

Table 9: Comparison of the regulated vs unregulated monopoly: the percentage of the regulated monopoly’s outcomes relative to that of the unregulated monopoly as a function of parameter $s_1$. The CES production function specified as $\gamma = -50, \sigma = 0.02$.

<table>
<thead>
<tr>
<th>$s_1$</th>
<th>output</th>
<th>price</th>
<th>total revenue</th>
<th>monopoly profit</th>
<th>consumer surplus</th>
<th>capital intensity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.02</td>
<td>107 %</td>
<td>94 %</td>
<td>101 %</td>
<td>2 %</td>
<td>115 %</td>
<td>932 %</td>
</tr>
<tr>
<td>1.10</td>
<td>106 %</td>
<td>95 %</td>
<td>101 %</td>
<td>10 %</td>
<td>112 %</td>
<td>876 %</td>
</tr>
<tr>
<td>1.20</td>
<td>106 %</td>
<td>95 %</td>
<td>101 %</td>
<td>19 %</td>
<td>112 %</td>
<td>803 %</td>
</tr>
<tr>
<td>1.30</td>
<td>106 %</td>
<td>95 %</td>
<td>101 %</td>
<td>26 %</td>
<td>112 %</td>
<td>742 %</td>
</tr>
<tr>
<td>1.40</td>
<td>106 %</td>
<td>95 %</td>
<td>101 %</td>
<td>32 %</td>
<td>112 %</td>
<td>689 %</td>
</tr>
<tr>
<td>1.50</td>
<td>106 %</td>
<td>95 %</td>
<td>101 %</td>
<td>37 %</td>
<td>112 %</td>
<td>643 %</td>
</tr>
<tr>
<td>2.00</td>
<td>106 %</td>
<td>95 %</td>
<td>101 %</td>
<td>56 %</td>
<td>112 %</td>
<td>482 %</td>
</tr>
<tr>
<td>5.00</td>
<td>106 %</td>
<td>95 %</td>
<td>101 %</td>
<td>89 %</td>
<td>112 %</td>
<td>193 %</td>
</tr>
<tr>
<td>10.00</td>
<td>100 %</td>
<td>100 %</td>
<td>100 %</td>
<td>100 %</td>
<td>100 %</td>
<td>100 %</td>
</tr>
</tbody>
</table>

In conclusion, the additional simulations presented in this section demonstrate that the main results of Section 3 are surprisingly robust, and continue to hold even if one makes major changes to the underlying parameter values or uses a different functional form for the production function. The impact of the substitution elasticity between inputs proves surprisingly small, and unexpectedly, the decrease of substitution elasticity does not necessarily help to improve effectiveness of rate of return regulation.
6. Conclusion and Policy Implications

The contributions of the present paper to the regulation of natural monopolies were threefold. First, we clarified the conceptual distinction between the two alternative types of revenue cap known in the literature: the British revenue cap that restricts the revenue growth using the RPI-X approach, and the Nordic revenue cap that constrains the level of revenue. We noted that, from the regulated monopoly’s point of view, the Nordic variant of the revenue cap is equivalent to the classic rate of return regulation. Therefore, the second contribution of the paper was to demonstrate formally that the Nordic revenue cap regime is vulnerable to the Averch-Johnson effect. The third contribution of this paper was to examine the magnitude of the capital bias and the welfare effects by means of numerical simulations, applying the most standard textbook case of a linear demand function and the Cobb-Douglas production function.

Our simulation experiment demonstrates that the Nordic revenue cap regulation has several desirable effects, in spite of the capital bias. Most importantly, regulation generally decreases the consumer price and the monopoly profit while increasing the output and the consumer surplus compared to the unregulated monopoly. In the absence of superior regulatory instruments, the Nordic revenue cap or the simple rate of return regulation can offer substantial benefits compared to an unregulated monopoly. In comparison with the welfare benefits, the distortionary effect on capital bias appears to be of secondary importance.

Our findings provide several practical lessons for the real-world regulation practice. Firstly, our conceptual discussion demonstrates that one should not let some superficial labels on regulation regimes blur the view on the incentives and constraints that influence the economic behavior of the regulated monopoly. Indeed, the real world regulation practice often involves a combination of several approaches, as recognized by NordREG (2011). By drawing attention to the potential capital bias, we would argue that the Averch-Johnson effect should be considered when designing incentive schemes based on the Nordic revenue cap regime.

Secondly, our numerical simulations show that the regulator does not need to force the monopoly to operate at zero profit (at the risk of bankruptcy), but relatively light handed regulation suffices to achieve the main benefits. Even if the regulator applies a stringent constraint to virtually eliminate the monopoly profit, the outcome will still fall short from the competitive market equilibrium. The rate of return regulation can alleviate the symptoms of the market failure, but it cannot cure the underlying disease, which is the lack of competition.
Thirdly, the main lessons from numerical simulations prove surprisingly robust to changes in the underlying parameter values and even the functional form of the production function. Changes in the price elasticity of demand or the substitution elasticity of inputs have only negligible impacts on the effectiveness of regulation.

One important practical caveat in the Averch-Johnson model is worth noting, which directly affects the results of this paper. In the tradition of the microeconomic theory, the model does not consider the possibility of inefficient production. Therefore, capital investment will immediately increase the output of the monopoly, as the possibility of inefficient investment is not recognized by the model. Clearly, it would be in the monopoly’s interest to invest in “gold-plating” such that its asset base increases while the output remains unaffected. Indeed, addressing productive efficiency of monopolies in the context of regulation is an active area of research (see, e.g., Agrell and Bogetoft, 2011; Kuosmanen, 2012; Kuosmanen et al. 2013; Saastamoinen et al., 2017). As noted in Section 2, the regulator could incorporate efficiency incentives to rate of return regulation by making the acceptable return on capital conditional on efficiency.

Finally, we hope that our study could stimulate further application of computer simulations to test the impacts of regulation instruments before their application in the real world regulation. Simulations are a relatively inexpensive tool to assess the performance and feasibility of the regulatory instruments, which could be utilized much more extensively by the energy regulators. To keep a close contact with the theoretical discussion of the Averch-Johnson effect, the simulations conducted in this study are based on a highly stylized economic model that ignores several important features of the real-world regulation. Developing more realistic simulations that are calibrated to match the essential features and the regulation environment of a specific industry in a specific country offers a fascinating avenue for future research.9

Acknowledgements
The authors would like to thank the three anonymous reviewers for their helpful comments. The usual disclaimer applies.

References

---

9 Kuosmanen et al. (2013) present simulations calibrated to the Finnish electricity distribution industry to examine performance of alternative OPEX benchmarking methods.


Online appendix: Numerical simulations

Solving the constrained profit maximization problem of the regulated firm and the corresponding unconstrained problems of the unregulated firm and the competitive firm are computationally trivial for modern mathematical programming solvers. To facilitate broader applicability of the numerical simulations reported in this paper among practitioners (including management consultants and regulators), we have set up an Excel spread sheet where the profit maximizing problems of the monopoly solved using the standard Excel solver tool. An Excel file containing the baseline scenario and the results of most alternative scenarios is attached as a supplementary file to this paper.