A PRINCIPAL-AGENT PROBLEM OF ITS OWN

Introductory Remark

In the Winter term 2001/02, I gave a course on "Institutional Economics" (*Institutionenökonomik)* at Kassel University. Among other topics I discussed the principal-agent approach, using the textbook by Richter/Furobotn (1996/1999; 1997). As I was unable to reproduce their result that flat marginal costs of effort would lead to a second-best solution close to the first-best solution, I was reading the underlying article by Holmstrom/Milgrom (1987). There I found that these authors came to the conclusion reported by Richter/Furobotn in a rather particular way, namely by using a measure of relative distance instead of an ordinary distance function, which in fact did not fulfil the properties of a mathematical distance function at all. When I reported this finding to one of the authors, Bengt Holmstrom, he replied to me that they were well aware of this problem, and they had deliberately chosen their distance measure to avoid the opposite result (that a high marginal cost of effort would lead to a solution close to the first-best which would have come out by using a normal distance function). Without regarding the mathematical properties of a distance function, he flatly declared that the whole stuff was "a matter of taste". After I sent my note to the editor of Econometrica, Glenn Ellison, I received a reply from him that they "did not find any flaws in the analysis of [my] paper", but that he did not regard it "as warranting publication as a correction of an error". For this reason I present it to the audience in the internet leaving it to the readers how they judge themselves the importance of my accidental finding.

Reference:
Richter, Rudolph/Furobotn, Eirik G.: *Neue Institutionenökonomik*. Tübingen: Mohr Siebeck, 1996, 2nd ed. 1999 [the 3rd ed. (2003) includes a hint to the problem raised in my note on p. 236]; English version: Furobotn, E. G./Richter, R.: *Institutions and Economic Theory*. Ann Arbor, U.S.A.: University of Michigan Press.

NOTE ON B. HOLMSTROM / P. MILGROM: AGGREGATION AND LINEARITY IN THE PROVISION OF INTERTEMPORAL INCENTIVES

In their path-breaking and often quoted article, B. Holmstrom and P. Milgrom illustrate their general theoretical reasoning by discussing some examples (section 5): They assume a quadratic function for the agent's costs of effort,

$$c(\mu) = (k/2)\mu^2 .$$

The agent is assumed risk averse, the principal risk neutral (cf. p. 323). The authors observe "that the first-best solution in this situation, attainable if the agent's choice of μ were costless observable, would entail instructing the agent to chose $\mu = k^{-1}$ and paying the agent a constant wage equal to the cost of his action, $(2k)^{-1}$ (assuming a zero certain equivalent). This yields the profit level: $\pi = (2k)^{-1}$." As a consequence of moral hazard and the agent's risk aversion, they come to the second-best solution

$$\mu^* = \left(1 + rk\sigma^2\right)^{-1} k^{-1} \text{[their equation (30)]},$$

and the principal's net profit becomes

$$\pi^* = \left(1 + rk\sigma^2\right)^{-1} (2k)^{-1} \text{ [their equation (31)]}$$

Referring to the equation

$$\mu^* / \mu = \pi^* / \pi = \left(1 + rk\sigma^2\right)^{-1}$$

the authors conclude: "Thus, a small degree of risk aversion or uncertainty, or a flat marginal cost of effort will allow a solution close to the first-best" (p. 323-4).

While this assertion is true for risk aversion measured by r or uncertainty measured by the variance σ^2, it is not true for the marginal cost of effort k.

Of course, a high k is something bad, and it should have unpleasant consequences which it has indeed: If we look at the first-best solution, $\mu = k^{-1}$ and $\pi = (2k)^{-1}$, we see immediately that the value of the solution decreases if k increases. So, the bad consequences of a high k are directly revealed by comparing the values of the first-best solution for alternative values of k.

By looking at equations (30) and (31), we see a similar relationship between μ^* and π^* on the one hand and k on the other hand; if we write equations (30) and (31) as fractions, we see that the negative impact of a high k is even strengthened as k enters the denominator twice.

But as a consequence of all this, we get the differences between the first- and the second-best solutions as follows

$$\mu - \mu^* = r\sigma^2\left(1 + rk\sigma^2\right)^{-1}$$

and

$$\pi - \pi^* = r\sigma^2\left(1 + rk\sigma^2\right)^{-1}2^{-1}.$$

That leads to the somewhat surprising and at a first glance counterintuitive result that the differences between the first- and second-best solutions become *smaller* as k *increases* (in contrast to the authors' statement quoted above).

Measuring the "closeness" of first- and second-best solutions in terms of μ^*/μ and π^*/π, as the authors do, makes no economic sense, as the welfare loss in this situation is measured by the difference π-π^* and not by π^*/π. If one applied Holmstrom's and Milgrom's relative measure π^*/π as a reasonable measure for the "closeness" of first- and second-best solutions, this would lead to the strange conclusion that welfare losses, due to second-best solutions, increase as first- and second-best solutions come closer and closer.

But why bother about welfare losses as long as the result of the relative measure "fits well with intuition" (p. 324)? Is it not just a matter of taste which measure we apply for "closeness"? Well, it is a matter of mathematics: By speaking of "close(ness)" of first and second best solutions, the authors refer to the notion of "distance". Measuring distances in mathematics is subject to the requirements of a *distance function* d(x,y). For any $x,y,z \in X$, the following properties must be given: (1) $d(x,x) = 0$ (2) $d(x,y) \geq 0$; furthermore (3) $d(x,y) = d(y,x)$ and (4) the triangle inequality $d(x,z) \leq d(x,y)+d(y,z)$ must be fulfilled. Holmstrom's and Milgrom's measure violates in general all these requirements, apart from the nonnegativity condition (2) and hence is not a permissible distance function. In contrast, the conventional measure as (absolute value of the) difference between first and second best solutions fulfills all mathematical properties of a distance function.

Therefore, the "unpleasant" consequences of this conventional measure with respect to k cannot be avoided by relying on another measure which leads to the desired results vis-à-vis k but does not measure distance and closeness, and hence does not allow to draw conclusions of a "flat k" for the closeness of first and second best solutions. It is better to accept and to explain the at first unexpected consequences of low marginal costs of effort for the distance between first and second best solutions than to give up the mathematical concept of a distance at all. What is generally true in economics, seems also to prevail in mathematics: There ain't such things as a free lunch.

Hans G. Nutzinger, Professor of Economics, University of Kassel, Department of Economics, 34109 Kassel, Germany; e-mail: nutzinger@uni-kassel.de

REFERENCE

HOLMSTROM, BENGT, AND MILGROM, PAUL (1987): "Aggregation and Linearity in the Provision of Intertemporal Incentives", *Econometrica*, 55. 302-32