Social norms and optimal incentives in firms

Steffen Huck, Dorothea Kübler and Jörgen Weibull*

February 6, 2002

Abstract. This paper deals with the interplay between economic incentives and social norms in firms. We outline a simple model of team production and provide preliminary results on linear incentive schemes in the presence of a social norm that may cause multiple equilibria. The effect of the social norm on the optimal bonus rate is discussed, as well as the effectiveness of temporary changes in the bonus rate as a means to move a firm from a bad to a good equilibrium.

JEL classification: D21, Z13

Keywords: incentives, social norms, team production

1. Introduction

Behavior in firms is most likely not only governed by economic incentives but also by social norms. This paper deals with the interplay of these two forces. We show that optimal incentive schemes may fundamentally change if social norms are accounted for. Today, it is not unusual to include social norms in microeconomic analysis.[1] However, there have not been many attempts at studying how social norms change the optimal incentive structures in firms.[2] This note should be read as such an attempt, the results of which are preliminary.

For a firm owner, social norms concerning work are important because they can affect the productivity of the firm and hence profits. For example, norms may influence how much effort workers put into projects where only joint output is observable.

---

*Huck is at University College London, Department of Economics & ELSE, Gower Street, London WC1E 6BT, United Kingdom; Kübler at Humboldt University Berlin, Department of Economics, Spandauer Str. 1, D-10178 Berlin, Germany; and Weibull at the Stockholm School of Economics, Department of Economics, P.O. Box 6501, SE - 113 83 Stockholm, Sweden, and at the Research Institute of Industrial Economics, P.O. Box 5501, SE - 114 85 Stockholm, Sweden. We are grateful for comments from Tore Ellingsen and from participants in a seminar at the Stockholm Institute for Transition Economics. This note is an elaboration of the rough sketch in Weibull (1997).

[1]See e.g. Akerlof (1980), Moffitt (1983), Besley and Coate (1992), Bernheim (1994), Hart (2001), Lindbeck, Nyberg, and Weibull (1999) and, recently, Kübler (2001) and the literature cited therein.

[2]See Kandel and Lazear (1992) and Barron and Gjerde (1997). Also Hart (2001) focuses on norms and firms, but rather deals with the question whether the degree of trust between agents influences the optimal ownership structure.

Social norms may also keep workers from working hard under relative performance schemes or piece rate schemes that are adjusted according to past performance. Under many such contracts, the compensation to a worker not only depends on his or her own effort level, but also on the effort of other workers. Moreover, peer pressure penalizes those who deviate from the group norm, and depending on the norm, output may be higher or lower than without the norm.

We analyze work norms in a static model of team production, much along the lines of Holmström (1982). Each agent's effort level is unobserved by the principal, but total output can be observed and verified. The principal chooses a linear incentive scheme in order to maximize profits, and the agents in the team simultaneously choose their efforts thereafter. We study the effect of a social norm concerning work effort among the team members. In particular, we show how the optimal incentive scheme depends on the social norm.

We believe that the theoretical possibility of multiple equilibria, which easily arises in such models, has empirical relevance. Indeed, different output levels under identical incentive schemes have been observed in different branches of firms.[3] Whether multiple equilibria exist in theory depends on agents' social preferences. While Kandel and Lazear (1992) rule out multiplicity by assuming the peer pressure function to meet certain regularity conditions (including convexity), we side with Lindbeck *et al.* and argue that such restrictions are hard to justify *a priori*. If the regularity conditions are violated, then equilibria with low efforts and low social pressure can coexist, under the same incentive scheme, with equilibria with high efforts and high social pressure. This multiplicity is relevant for a principal who strives to find a profit-maximizing incentive scheme. For example, a firm trapped in a low-effort equilibrium may "jump" to a high-effort equilibrium even by way of a small increase in the bonus if the equilibrium correspondence has a fold just above the current bonus rate (see example below). Or, from a dynamic perspective, it may be beneficial for the principal to temporarily raise the bonus, until a high-effort equilibrium with a high work norm is reached. Afterwards, the bonus rate can be decreased to its original value while the workers' efforts only decrease gradually, and thereafter remain high, due to the new and more demanding work norm.[4]

The paper is organized as follows: The model is introduced in the next section. As a benchmark, it is first solved in the absence of social norms. Then, we characterize optimal effort levels in the presence of a work norm and derive a sufficient condition

[3]See *e.g.* Ichino and Maggi (2000). They present an empirical investigation of shirking differentials between branches of an Italian firm. Group-interaction effects are identified which allow for multiple equilibria. However, they do not analyze the interplay of these effects with economic incentives.

[4]See Lindbeck, Nyberg, and Weibull (1999) for an elaboration of a similar argument.

for uniqueness of the equilibrium. An example with multiple equilibria is presented in section 3. Section 4 concludes.

2. The model

We consider team production with a profit maximizing owner (the principal) as residual claimant. There are $n > 1$ identical workers (agents). Each worker i exerts some effort $x_i \geq 0$. Let x_{-i} denote the average effort exerted by all other workers, $x_{-i} = \sum_{j \neq i} x_j / (n-1)$. The production technology is linear: output y equals the sum of all workers' efforts, $y = \sum_{i=1}^{n} x_i$. The principal can only observe aggregate output y, not individual efforts x_i. Each worker i observes the others' average effort x_{-i} or, equivalently, output y. In order to focus on the interplay between economic incentives and social norms in the simplest possible setting, this preliminary investigation is restricted to linear contracts.[5] More exactly, each worker earns the same wage w, and this wage is an affine function of the firm's output,

$$w = a + by/n \ .$$

where the owner chooses the fixed wage a and the bonus rate b. We require a to be nonnegative, an assumption which can be justified by wealth constraints.[6] Therefore, the profit maximizing owner will optimally choose b in the open unit interval, $0 < b < 1$.

Assuming that the firm is a price taker in its product market, and normalizing the market price to unity, the firm's profit — the residual left to the owner — is

$$\pi = (1-b)y - na.$$

2.1. Without a social norm. We first analyze the benchmark case of purely economic incentives, i.e., when social norms are absent or have no influence on behavior. A worker's utility then only depends on his or her wage earning and exerted effort. We assume each worker's utility to be linear-quadratic:

$$u_i = w - \frac{1}{2}x_i^2 = a + \frac{n-1}{n}bx_{-i} + \frac{b}{n}x_i - \frac{1}{2}x_i^2 \ .$$

From this it is immediate that workers' decisions concerning effort are strategically independent. Regardless of whether workers decide simultaneously or sequentially, each worker solves

[5] Holmström and Milgrom (1987) identify conditions under which linear incentive schemes are optimal.

[6] Notice that otherwise the principal could sell the firm to the workers. In the absence of social norms, such an arrangement would make the first-best solution achievable. However, such schemes are vulnerable to both collusion and sabotage.

$$\max_{x_i \geq 0} \frac{b}{n} x_i - \frac{1}{2} x_i^2 \ .$$

Consequently, the unique equilibrium effort levels, given any contract (a, b), are $x_i = b/n$ for all i.

Inserting these effort levels, we see that the owner's equilibrium residual (profit) is linear-quadratic in the contract: $\pi = (1 - b)b - na$. Thus, the optimal contract - the unique subgame perfect equilibrium contract - is a zero fixed wage, $a = 0$, combined with the bonus rate $b = 1/2$. The unique equilibrium effort is $x_i = b/n = 0.5/n$ for all workers i.

This common equilibrium effort level can be contrasted with the common effort level the workers would like to commit to if they could, namely the level which maximizes the sum of their utility. It is easily verified that this maximum, under any given contract (a, b) is obtained when every worker exerts effort $x_i = b$. We will call this the *team optimum* effort under contract (a, b). At the team optimum, each worker thus exerts n times more effort than in the absence of commitment possibilities.[7]

Another reference point is the total *welfare maximum*, i.e., those effort levels which maximize the sum of the owner's profit and the workers' utility. This sum is simply output minus total disutility of effort, $\sum_i (x_i - x_i^2/2)$. Hence, welfare maximization requires $x_i = 1$ for all i, leading to a welfare maximum of $n/2$.

2.2. With a social norm. We now add the following social norm or "work ethic" to each worker's preferences: to exert the team optimum effort level, $x_i = b$. In other words, under any given contract (a, b), each worker feels that he or she should ideally exert the same (high) effort level b; the effort that maximizes the sum of all workers' utility. We assume that a worker's embarrassment or disutility of exerting less effort than this is an increasing function of the other workers' average effort. More specifically, we assume that each worker's utility is additively separable in wage earnings, disutility of effort and disutility of norm-deviation, as follows:[8]

$$u_i = w - \frac{1}{2} x_i^2 - \frac{1}{2} v(x_{-i})(b - x_i)^2.$$

Here $v : \mathbb{R}_+ \to \mathbb{R}_+$ is continuous and strictly increasing; the more effort others exert, the more embarrasing it is to shirk.

[7]Interestingly, when the workers can commit to a common effort level, the optimal contract is the same as without commitment power, since then we would have $x_i = b$ for all i, and thus $\pi = n(1 - b)b - na$, implying $a = 0$ and $b = 1/2$.

[8]Kandel and Lazear (1992) model workers' utility as additive in wage earnings, disutility of effort, and social pressure.

The earlier strategic independence of effort choices is lost. What is optimal for one worker depends on what other workers do. Thus, the timing of effort choice is now relevant. In the following, we assume that workers decide simultaneously. More exactly, we solve for subgame perfect equilibria in the game where the owner (principal) first chooses a contract whereupon all workers observe this and simultaneously choose their efforts.[9]

Given any contract (a, b), worker i thus solves (irrespective of a)

$$\max_{x_i \geq 0} \frac{b}{n} x_i - \frac{1}{2} x_i^2 - \frac{1}{2} v(x_{-i})(b - x_i)^2 \ .$$

It is easily verified that this gives

$$x_i = b \frac{v(x_{-i}) + 1/n}{v(x_{-i}) + 1}$$

for all i. Hence, each worker's effort level is linear in the bonus rate. We focus on symmetric Nash equilibria. Given any bonus rate b and firm size n, the set of such Nash equilibria in the subgame is characterized by the fixed-point equation $x = bf(x)$, where $f : \mathbb{R}_+ \to \mathbb{R}_+$ is the continuous and increasing function defined by

$$f(x) = \frac{v(x) + 1/n}{v(x) + 1} \ . \tag{1}$$

Each fixed point x is the common equilibrium effort level in a symmetric Nash equilibrium, and *vice versa*.

It is easily verified that there exists at least one fixed point x, and that no equilibrium effort is lower than in the model without the social norm. Moreover, if v is twice differentiable and concave, then so is f, and hence multiplicity of equilibria is excluded in this case. More generally:

Remark 1. *There exists at least one fixed point x. If x is a fixed point, then $b/n \leq x < b$. If v is twice differentiable with $[1 + v(x)]\, v''(x) \leq 2\,[v'(x)]^2$ for all $x \in (0, b)$, then there exists exactly one fixed point.*

Proof: The first two claims follow from f being continuous and positive with $1/n < f(x) < 1$ for all $x \in \mathbb{R}_+$. The third claim follows from

$$f''(x) = (1 - 1/n)\, b \left([1 + v(x)]\, v''(x) - 2\,[v'(x)]^2 \right) / \left[1 + v(x)\right]^2 .$$

[9] In particular, worker i does not then know the others' average effort x_{-i}. Nevertheless, the latter enter i's utility function. One may imagine that workers learn about each others efforts after the simultaneous choices, and utilities are evaluated then. In equilibrium, no worker would like to unilaterally change his or her effort.

We note in passing that if there were no embarrassment of deviating from the social norm, *i.e.*, if $v(x) \equiv 0$, then the unique equilibrium effort would be $x = b/n$, just as in the benchmark case. More generally, if the embarrassment were independent of others' efforts, $v(x) \equiv \theta \geq 0$, then the unique equilibrium effort decreases in team size n but increases continuously with the fixed embarrassment θ from b/n when $\theta = 0$ towards b as θ increases towards plus infinity.

We conclude by noting that, irrespective of whether subgame equilibria are unique or not, the optimal contract has zero fixed wage, $a = 0$, and a bonus rate b which maximizes profits, $n(1-b)x$, subject to the (equilibrium) constraint $x = bf(x)$. The function f being everywhere positive, this optimization program can be solved in two steps, by first solving

$$\max_{x \in (0,1)} \left[1 - \frac{x}{f(x)}\right] x \ ,$$

and thereafter computing the associated optimal bonus $b = x/f(x)$. In the absence of the social norm, this program boils down to $\max_x (1 - nx)\, x$, resulting in $x = 0.5/n$ and $b = 0.5$, as shown in the preceding subsection. The effect of the social norm is thus to replace the exogenous coefficient n by the endogenous coefficient $1/f(x)$ in the above maximization program As will be seen in the following example, the effect of the social norm on the optimal bonus and resulting efforts and profits can be significant.

3. Example

According to Remark 1, a necessary condition for the existence of multiple Nash equilibria is that the disutility function v at least locally has sufficient curvature upwards in the sense of (at least locally) violating the inequality. In particular, v has to be at least locally convex. One class of such functions are given by the equation

$$v(x) = \alpha e^{-(x-1)^2/\sigma^2} \tag{2}$$

for some $\alpha, \sigma > 0$. These functions are rescalings of the density function of the normal distribution with mean value 1 and variance σ^2. Figure 1 below shows the graph of the associated function f for $\alpha = 5$, $\sigma^2 = 1/6$, $n = 7$, and $b = 0.4$, 0.87, and 1 respectively. Higher values of the bonus rate b correspond to higher curves.

At the maximal bonus rate $b = 1$, the equilibrium effort level is unique and quite high. By contrast, at the bonus rate $b = 0.4$ there is a unique equilibrium with rather

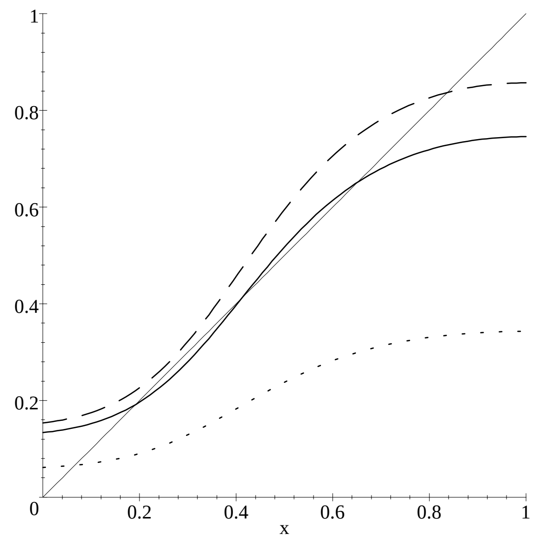

Figure 1: The fixed-point equation $x = bf(x)$, for $b = 0.4$, $b = 0.87$, and $b = 1$, respectively.

low effort levels. Finally, if the bonus rate b is 0.87 we observe multiple equilibria — with low, medium, and high effort. Of these three, only the two extreme ones would be stable in a dynamic where workers gradually adjust their work efforts in the light of the current average effort. So, with v defined as above, we find that at low bonus rates there exists one low-effort equilibrium. At a certain intermediate bonus rate two equilibria exist, one stable with an effort level that is a continuous extension of the effort at lower bonus rates, and another unstable equilibrium with significantly higher effort. Above this critical bonus rate there is an interval of bonus rates with three equilibria: one stable equilibrium with high effort, another stable equilibrium with lower effort, and one unstable equilibrium in-between. At a higher critical bonus rate, two of these equilibria "merge," and a stable equilibrium with high effort and an unstable equilibrium with low effort remain. Finally, for bonus rates above this critical value, only one stable high-effort equilibrium remains. Hence, it is sufficient to raise the economic incentive above the second critical value in order to get rid of the low effort equilibria.

Figure 2 plots the equilibrium correspondence that gives the set of equilibrium effort levels x for each bonus rate b (the other parameters are the same as in Figure

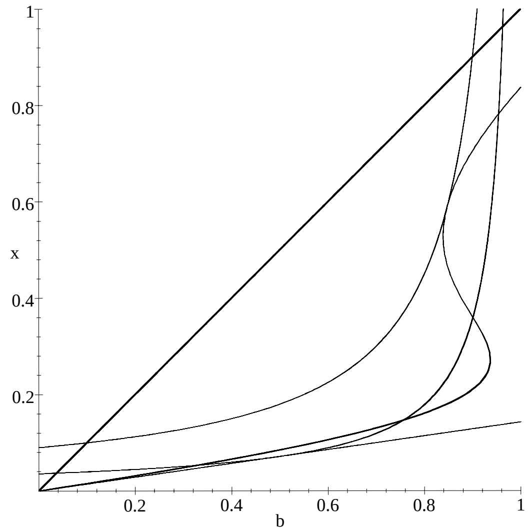

Figure 2: Subgame equilibrium combinations of bonus rate b and effort x.

1).Hence, the plotted curve is given by the equation $x = bf(x)$, which here becomes

$$x = b\frac{1 + 35e^{-6(1-x)^2}}{7 + 35e^{-6(1-x)^2}}. \tag{3}$$

This is the S-shaped curve in the diagram. The steep straight line represents the team optimum effort level in the absence of the social norm, $b = x$, and the less steep straight line the equilibrium effort in the absence of the social norm, then $b = nx$. The two hyperbolas are iso-profit curves.[10] Tangency with such a hyperbola is thus necessary for optimum.

Which bonus rate b maximizes the firm's profit? Plugging equation (3) into the expression $\pi = n(1-b)x$ gives

$$\pi = 7\left[1 - \frac{7 + 35e^{-6(1-x)^2}}{1 + 35e^{-6(1-x)^2}}x\right]x$$

The graph of this function is plotted in Figure 3, showing that the equilibrium effort which maximizes the firm's profit is approximately $x = 0.6$. By plugging this value

[10]Each iso-profit curve is of the form $b = 1 - c/x$, for $c = \pi/n$.

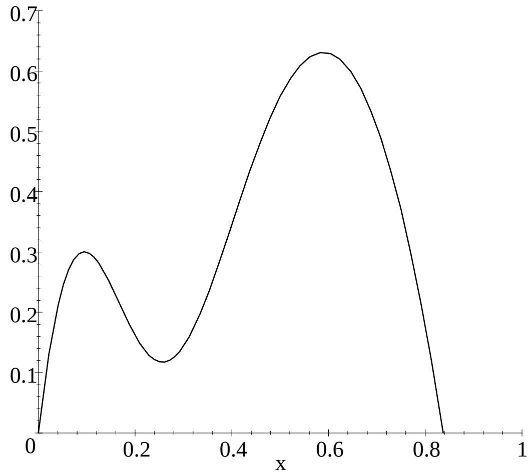

Figure 3: Profit as function of equilibrium effort.

into equation (3) we find that the optimal bonus rate is $b \approx 0.85$, yielding profit $\pi = n(1-b)x \approx \pi \approx 0.63$. By contrast, in the absence of the social norm, we found that the optimal bonus rate was $b = 0.5$, with effort $x = b/n \approx 0.07$ and profit $\pi \approx 0.25$ in this example.

At the optimal bonus rate, the subgame played by the workers, after the contract has been set, has multiple equilibria, as shown in Figure 2.[11] Thus, at the optimal bonus, the firm cannot induce the optimal effort for sure. Workers might coordinate on one of the two other equilibria, for instance the low-effort equilibrium, which has $x \approx 0.18$. In this subgame equilibrium the profit is only about 0.19, and also the utility to each worker is lower, 0.82 utiles instead of 1.86. However, the multiplicity of subgame equilibria, with the accompanying possibility of mis-coordination, can be avoided by the owner by instead choosing a slightly higher bonus rate, say $b = 0.94$. The unique subgame equilibrium effort at that bonus rate is $x \approx 0.77$, yielding a profit of about 0.32.

We finally note that the double-peakedness of the profit function in Figure 3 does not hinge on the multiplicity of subgame equilibria. By continuously moving from the model without the social norm to the current model with a social norm, by way of increasing the factor α in equation (2) from 0 to 5, this profit function will first be single-peaked at $x = 0.5/n$, then double-peaked, with the new peak small but

[11]Multiple equilibria arise for $0.84 \lesssim b \lesssim 0.94$.

rising until one obtains the current graph. Hence, the social norm *per se* induces a discontinuity in this equilibrium correspondence.

4. Discussion

This paper analyzes optimal incentive schemes in the presence of a social work norm. We found in our example that the optimal bonus rate was higher than in the absence of the social norm. The reason is that an increase in the bonus rate not only increases the economic incentive to each worker ($\Delta x_i = \Delta b/n$), but also indirectly increases the "social incentive": if others work harder (because of their increased economic incentives), then I also work harder in order to mitigate my increased social embarrassment (due to the raised work norm). We also found in the example that a social norm makes all parties materially better off, firm owners as well as workers. Moreover, the example demonstrates that, at least for certain social embarrassment functions v, it might be desirable for the owner to set the bonus rate b just above the critical value above which a unique equilibrium remains. The bonus should not be lower if a firm wants to avoid the co-existence of an alternative low-effort equilibrium and it should not be higher, since this reduces the profit, see Figure 3.

The multiplicity of equilibria also suggest a dynamic perspective. For example, suppose the firm pays a bonus rate at which three equilibria co-exist, but workers only put in little effort such that the low-effort equilibrium is realized. To move away from this inefficient equilibrium, the firm may increase the bonus up to a level where equilibrium is unique. Assuming that workers adapt gradually to changes in the bonus b, along the current branch of the equilibrium correspondence, the firm can afterwards move back to its original bonus but now at the efficient, high-effort, equilibrium.[12]

In a future project we plan to study team production game protocols in a series of laboratory experiments.[13] While we conjecture that this will confirm the relevance of social norms, we also expect that institutional details will influence the evolution and strength of such norms.[14] Another avenue for future work is to use tools from evolutionary game theory to (a) analyze the relative stability of alternative subgame equibria in case of multiplicity, (b) endogenize the social norm for work effort, which here was exogenously set at the team optimum level.

[12] See sections II and VII in Lindbeck *et al.* (1999) for a discussion of similar equilibrium dynamics.

[13] In experimental labor markets social norms based on reciprocity have been shown to be extremely important. See, in particular, the work by Fehr and collaborators, e.g., Fehr, Kirchsteiger, and Riedl (1993) or Fehr, Gächter, and Kirchsteiger (1997). However, these experiments deal with reciprocal relations (or norms) between employer and employee while our study addresses social relations between different employees.

[14] See, for example, recent work by Bohnet, Frey, and Huck (2001) who study the evolution of trust in contractual relationships, both, theoretically and experimentally.

References

[1] Akerlof G. (1980). A Theory of Social Custom, of Which Unemployment May Be One Consequence, *Quarterly Journal of Economics* 84, 749-775.

[2] Barron, J.M. and K.P. Gjerde (1997). Peer Pressure in an Agency Relationship. *Journal of Labour Economics* 15: 234-254.

[3] Bernheim D. (1994). A Theory of Conformity. *Journal of Political Economy* 102(5): 841-877.

[4] Besley T. and S. Coate (1992). Understanding Welfare Stigma: Taxpayer Resentment and Statistical Discrimination. *Journal of Public Economics* 48: 165-183.

[5] Bohnet, I., B.S. Frey, and S. Huck (2001). More Order with Less Law: On Contract Enforcement, Trust, and Crowding. *American Political Science Review* 95: 131-144.

[6] Fehr, E., Gächter, S., and G. Kirchsteiger (1997). Reciprocity as a Contract Enforcement Device. *Econometrica* 65: 833-860.

[7] Fehr, E., Kirchsteiger, G., and A. Riedl (1993). Does Fairness Prevent Market Clearing? *Quarterly Journal of Economics* 108: 437-460.

[8] Hart, O. (2001). Norms and the Theory of the Firm. Harvard University. SSRN Electronic Paper Collection.

[9] Holmström, B. (1982). Moral Hazard in Teams. *Bell Journal of Economics* 13: 324-40.

[10] Ichino, A. and G. Maggi (2000). Work Environment and Individual Background: Explaining Regional Shirking Differentials in a Large Italian Firm. *Quarterly Journal of Economics* 115: 1057-1090.

[11] Kandel, E. and E.P. Lazear (1992). Peer Pressure and Partnerships. *Journal of Political Economy* 100: 801-817.

[12] Kübler, D. (2001). On the Regulation of Social Norms. *Journal of Law, Economics, and Organization* 17(2) (forthcoming).

[13] Lindbeck, A., Nyberg, S., and J. Weibull (1999). Social Norms and Economic Incentives in the Welfare State. *Quarterly Journal of Economics* CXIV: 1-35.

[14] Moffitt R. (1983). An Economic Model of Welfare Stigma. *American Economic Review* 73(5): 1023-35.

[15] Weibull J. (1997). Norms in firms: an example. Mimeo.