You Don't Always Get What You Pay For: Bonuses, Perceived Income, and Effort

Wendelin Schnedler*
Department of Economics
University of Heidelberg, IZA, and CMPO

September 29, 2011

Abstract

Consider a principal-agent relationship in which more effort by the agent raises the likelihood of success. This paper provides conditions such that no success bonus induces the agent to exert more effort and the optimal contract is independent of success. Moreover, success bonuses may even reduce effort and thus the probability of success. The reason is that bonuses increase the perceived income of the agent and can hence reduce his willingness to exert effort. This perceived income effect has to be weighed against the incentive effect of the bonus. The trade-off is determined by the marginal effect of effort on the success probability in relation to this probability itself *(success hazard-rate of effort)*. The paper also discusses practical implications of the finding.

Keywords: bonus, premium, incentives, income effect, moral hazard

JEL-Codes: M5, D8, J3

*wendelin.schnedler@awi.uni-heidelberg.de
Department of Economics, University of Heidelberg, Bergheimer Straße 58, D-69115 Heidelberg.
I thank Peter Dürsch, Matthias Kräkel, Roland Kirstein, Jörg Oechssler, participants at the Personalökonomisches Kolloquium 2008 in Bonn and at the GEABA 2008 in Augsburg, the editor and an anonymous referee for their helpful comments. All errors remain my own.

1 Introduction

Success bonuses are often introduced in order to boost performance above ordinary levels. Football players are promised a bonus if they win important matches, budgets of academics are tied to their publication record, lawyers get an additional reward if they win a case and builders receive premiums for timely completion. In all these cases, success increases with the effort of an agent (the football player, academic, lawyer, etc.) but is also affected by factors beyond the agent's influence—that is, luck. Rewarding an agent for success seems like a good idea to increase his effort and thus ultimately the probability of success. The present paper critically examines this intuition. It provides conditions such that no success bonus (however large it may be) induces an agent to exert more effort. With the additional assumption that the agent's liability is limited, the optimal contract is independent of success—even though success is informative about effort, verifiable, and costless. Moreover, when the agent is initially participating and motivated, success bonuses can even reduce effort and the probability of success.

The first result is that paying a small success bonus to an agent who likes money can lead to the same or even lower effort by this agent (Proposition 1). This finding hinges on two conditions: one about agent's preferences, the other about the success signal. Success bonuses are typically introduced because the agent would not voluntarily provide more effort; a reasonable starting point is thus that effort is a 'bad' rather than a 'good.' The condition on agents' preferences is that effort is a normal bad. In other words, the willingness to engage in effort decreases when the agent gets richer; or formally: the cross-derivative of the agent's utility in effort and money is negative. If effort is a normal bad, there is a perceived income effect. Although the agent does not actually get richer before the outcome is realized, any bonus implies that the agent behaves as if he were richer; more specifically, the certainty equivalent for the lottery that he faces increases and hence the marginal costs of effort become larger. There is thus a tendency to reduce effort. This perceived income effect needs to be traded-off against the usual incentive effect of a bonus. The crucial variable that determines this trade-off is the marginal effect of effort, e, on the probability of success, $p'(e)$, in relation to this probability itself, $p(e)$, the *success hazard-rate of effort*: $p'(e)/p(e)$. The condition about the signal is that this rate is low, so that the perceived income effect dominates the incentive effect. Jointly the two conditions imply that a small success bonus reduces effort.

The second result provides an additional condition such that no success bonus induces the agent to exert more effort (Proposition 2). The familiar assumption that marginal utility of money decreases is typically justified by saying that it becomes increasingly difficult to seek good opportunities for spending money. This may be true regardless of an agent's effort. However, an agent who exerted less effort may seize opportunities that are overlooked by an agent who is tired because he exerted more effort. In this case, the decrease in the marginal utility of money is likely to be larger for larger effort. This assumption ensures that the normal character of effort is even more pronounced, if the agent becomes richer. Given this assumption, an agent cannot be enticed to exert more effort—irrespective of the size of the success bonus—because the dominance of the perceived income effect over the incentive effect for a small success bonus translates into an even a larger dominance for a large success bonus.

The third finding is a direct corollary to the second one: given the previous conditions and assuming that the agent's liability is limited, the optimal contract is not based on the success information (Corollary 1). The logic is straightforward. With limited liability, the only way to provide incentives to the agent is to use success bonuses. Since no bonus can induce the agent to exert more effort, bonuses only lead to costs, which can be avoided by ignoring the success information.

One of the seminal findings from the moral-hazard literature is Holmström's sufficient statistic result: any freely available independent information about agent's effort will be used in the optimal contract (Holmström, 1979, 1982). Amershi and Hughes (1989) point out that the result only holds for a restricted class of distributions, while Holmström and Milgrom (1991) show that it is limited to single-task settings. The present paper complements these observations: even in a standard single-task principal-agent model, agent's preferences may prevent all freely available information from being used in the optimal contract.

The fact that success bonuses have an immediate effect on income and can thus detrimentally affect effort and success has to my knowledge not yet been discussed in the literature.[1] However, there are two contributions that deal with the somewhat related effect of wealth on incentives. Mookherjee (1997) shows that agents who are extremely poor or rich are harder to motivate than agents of intermediate wealth; the very poor

[1] The reason is probably that most most principal-agent models assume for mathematical convenience that the agent's preferences are additively separable in effort and money or can be represented by a negative exponential utility (for an exception see Grossman and Hart, 1983). These assumptions, however, eliminate the effect.

because they have little wealth to use as a collateral and the very rich because they derive relatively little marginal utility from money. Thiele and Wambach (1999) ask the related question whether the principal's surplus is greater from poor than from rich agents and provide a corresponding condition. The crucial difference of both articles to the present paper is that wealth in these contributions is exogenous. In contrast, the present paper examines how promising a success bonus endogenously determines agent's expected 'wealth' and thus his behavior.

The finding that bonuses can reduce effort is consistent with experimental as well as empirical evidence (see the survey articles by Bowles, 2008; Frey and Jegen, 2001) and complements various models that explain this evidence by psychological effects of incentives (see Bénabou and Tirole, 2003; Sliwka, 2007; Friebel and Schnedler, 2007; Schnedler and Vadovic, 2007; Herold, forthcoming). At the heart of all these models is an informational asymmetry: the principal knows more than the agent and this information is revealed by the provision of incentives. In contrast to the often involved signaling models, the explanation here is simple: bonuses reduce the agent's marginal willingness to engage in effort because they make him richer.

The next section presents a simple model to illustrate the mechanisms at work. This model is then analyzed in Section 3. Section 4 summarizes the findings.

2 Model

Consider an agent (he) who works for a principal (she) on a project that can be successful ($s = 1$) or not ($s = 0$). Success can be verified by courts and its probability p increases in the effort e of the agent, where e is a non-negative real number and p a twice continuously differentiable and concave function of e with $p'(0) > 0$.

Agent's utility. The agent's utility depends on money m and some non-negative effort e: $u(m, e)$, where u is a twice continuously differentiable concave function in m and e. While the utility is strictly concave in e to ensure a unique solution, it may be weakly or strictly concave in m, so that risk-neutrality as well as risk-aversion is covered. The utility is supposed to strictly increase in money, so that success bonuses can have an effect on effort. In the absence of the success bonus, there is some initial effort level $e^0 := \arg\max_e u(0, e)$, which may be zero as in most moral-hazard models or positive as in Holmström and Milgrom (1991). The latter captures situations in which the agent is initially motivated and allows us later to study the consequences of success bonuses in this case.

Interaction between the marginal utility of money and effort. Principal-agent models often assume away any interaction between effort and money in the agent's utility, i.e. $u(m, e)$ is supposed to be additively separable, so that $\frac{\partial^2 u(m,e)}{\partial m \partial e} = 0$. Here, interactions are allowed. In particular, it is possible that more effort reduces the marginal utility from money or (equivalently) that more money increases the marginal dis-utility from effort:[2] $\frac{\partial^2 u(m,e)}{\partial m \partial e} < 0$. In this case, the agent strictly prefers less effort when he is richer and in line with standard terminology effort is *normal*; the term *weakly normal* is used if the preference is not strict.

Principal's profit. The principal likes success and dislikes paying money to the agent.

[2]The equivalence follows from Young's theorem.

The details of the profit function, however, are not essential. For simplicity, suppose the profit is $\pi(s,m) := s - m$.

Contractual environment. All that can be observed is the success or failure of the project: $s = 1$ and $s = 0$. Any contract has to specify transfers (t_0, t_1), where t_0 is the payment in case of failure and t_1 that in case of success. In line with everyday language, we define a *success bonus* as transfers that satisfy: $t_1 > t_0 \geq 0$.

3 Analysis

The effort choice that maximizes the joint surplus of principal and agent in the absence of incentives is:

$$e^* := \text{argmax}_e u(0,e) + \mathrm{E}_S\left[\pi(S,0)\right] = u(0,e) + p(e),$$

where S is the random variable that describes success and can take on the realizations $s = 1$ and $s = 0$. The maximizer e^* of the joint surplus is unique because the objective function is concave. Recall that the actual choice in the absence of transfers is:

$$e^0 = \text{argmax}_e u(0,e).$$

Comparing the two objective functions, it becomes clear that the agent fails to internalize the externality of his effort choice on the principal and that the effort in the absence of incentives e^0 is too small from a social point of view: $e^0 < e^*$. Since effort itself cannot be contractually stipulated, the principal has to rely on the project's success in order to provide incentives. This leads to the central question: Can the principal induce a higher effort level than e^0 by using a success bonus?

First, note that weakly normal effort implies that a bonus in the case of failure, $t_0 > 0$, cannot be optimal (see Lemma 1 in the appendix). The intuition for this result is that such a transfer is irrelevant for participation, costly to the principal, and would reduce the agent's motivation.

Under the assumption that effort is weakly normal, a bonus implies that the agent's maximization problem is well-behaved. Exploiting the agent's first-order condition with the implicit-function theorem, one can study how the introduction of a success bonus affects effort. The proof for this and the following results can be found in the appendix.

Proposition 1. *Let effort be weakly normal. Suppose the agent initially exerts no effort, $e^0 = 0$, then a small success bonus does not increase effort if and only if the income effect dominates the incentive effect:*

$$p(e) \cdot (-1) \cdot \left.\frac{\partial^2 u(t,e)}{\partial m \partial e}\right|_{t=0,e=e^0} > p'(e) \cdot \left.\frac{\partial u(t,e)}{\partial m}\right|_{t=0,e=e^0}. \tag{1}$$

Suppose the agent initially exerts effort, $e^0 > 0$, then a small success bonus reduces effort if and only if inequality (1) *holds.*

The important insight from this proposition is that a small bonus payment may well decrease effort. Since success is an increasing function of effort, paying for success can thus lower the probability of success. Notice that this can only happen if the marginal utility of transfers is lower when the agent exerts more effort ($\frac{\partial^2 u(m,e)}{\partial m \partial e} < 0$). So, a reduction of effort cannot occur if the utility is separable in effort and transfers; effort needs to be normal.

If effort is normal, introducing a success bonus has two qualitative effects, which are reflected by inequality (1). First, it increases the marginal disutility of effort in case of success; the agent becomes richer in the sense that the certainty equivalent of the lottery that he faces is larger (left-hand side). Second, it increases the gains from effort; the agent thus has more incentives to exert effort (right-hand side). The trade-off of these two effects depends on the ratio $p'(e)/p(e)$. In other words, the effect of the next unit of effort on the probability of success in relation to this probability, the *hazard-rate of success (in effort)*, is crucial. This is interesting for two reasons. First, the parameter has a simple interpretation: the marginal effect of effort on success needs to be large in relation to the probability of being lucky. Second, the same parameter already features in other contexts. It is used in econometrical hazard-rate models (hence its name) but also induces an ordering in terms of employment probabilities and rents in adverse selection models (see Schnedler, 2002).

Notice that although bonuses affect income, the effect is different from the income effect in a standard labor supply model. While the standard effect is based on a deterministic link between hours worked and wage, the effect here is based on a stochastic link and becomes less relevant if the agent gains more control over the outcome (in the sense that the hazard-rate of success increases). It disappears entirely if the relationship between effort and success is deterministic because the agent can then no longer rely on luck to obtain the bonus.

The preceding proposition only deals with small bonuses. So the dictum of Gneezy and Rustichini (2000), 'Pay Enough or Don't Pay at All', may well hold in our model. As the following proposition shows, this is no longer possible if the utility from money becomes more concave for larger effort.

Proposition 2. *Suppose that effort is normal, $\frac{\partial^2 u(m,e)}{\partial m \partial e} < 0$, the marginal utility of transfers decreases more if more effort is exerted, $\frac{\partial}{\partial e}\frac{\partial^2 u(m,e)}{\partial m^2} < 0$, and inequality* (1) *holds. Then, no success bonus can implement an effort larger than e^0.*

The intuition behind this result is that the marginal effect of incentives weakens with larger bonus payments while the effect on marginal costs gets stronger. The result highlights that under some circumstances, it is impossible to increase effort using success bonuses. It also directly leads to the observation that under limited liability, it may not be optimal to use the success information at all. Without success bonuses the effort level is e^0. A necessary condition for the optimal contract to be based on the success information is that the contract implements an effort level above e^0. Otherwise, the expected payment to the agent only increases the principal's costs. Then, the only way to increase effort is through a success bonus (see Lemma 1). However, under the conditions from Proposition 2, there is no success bonus that generates a higher effort level than e^0. Hence, transfers in the optimal contract must be constant in the success information. The argument is independent of whether the participation constraint is binding; it thus holds in particular when the agent's expected utility is equal to his outside option. These insights can be summarized as follows.

Corollary 1. *Suppose the conditions from Proposition 2 hold and the agent's liability is limited. Then, the optimal contract does not condition on the freely available information about success and failure.*

The corollary emphasizes that the sufficient statistic result fails under seemingly innocent assumptions about agent's preferences.

4 Conclusion

Success bonuses can lead to a lower probability of success. As with any bonus payment, they increase agents' perceived income. Under the arguably reasonable assumption that more income reduces the agent's willingness to engage in effort, there is a perceived income effect that needs to be traded-off against the incentive effect of a success bonus. This perceived income effect is different from the income effect in a standard labor supply problem: the extent to which bonuses reduce effort depends crucially on the link between effort and success; it gets smaller when the agent has more control over success and disappears if the link is deterministic.

This paper shows that incentives may fail under innocuous assumptions even if they do not fall foul of the "Folly of Rewarding for A while hoping for B" (Kerr, 1975). The paper thus complements well-known reasons for the failure of incentive schemes such as multi-tasking (Holmström and Milgrom, 1991) and helps explain why explicit incentives are not more widely spread. The failure crucially depends on the assumption that the agent's liability is limited. Without this restriction, it is possible to set-off the bonus by a fine, keep the agent's income constant, and eliminate the perceived income effect. Presumably, this is the reason why the problem is not mentioned even if the utility function allows interaction between income and effort as in Grossman and Hart (1983).

A tentative conclusion from our results could be that success bonuses should be accompanied by a fine to keep perceived income constant. While practitioners tend to advocate success bonuses, they often advise against fines in case of failure. From a theoretical perspective and bearing in mind the results of this paper, this seems an unnecessary and potentially harmful restriction. More behaviorally oriented economists, however, argue that taking away from an agent is problematic because it may result in negative reciprocity (see e.g. Bewley, 1995; Fehr and Rockenbach, 2003; Falk and Kosfeld, 2006).

A direct practical implication of this paper is that before trying to boost effort by paying for success, it is sensible to ensure that the incentive effect is not dominated by the perceived income effect. For example, one should verify that the agent's effort strongly affects the probability of success and that his chances to succeed are not too large to begin with.

References

Amershi, Amin H. and John S. Hughes, "Multiple Signals, Statistical Sufficiency, and Pareto Orderings of Best Agency Contracts," *RAND Journal of Economics*, Spring 1989, *20* (1), 102–112.

Bénabou, Roland and Jean Tirole, "Intrinsic and Extrinsic Motivation," *Review of Economic Studies*, 2003, *70*, 489–520.

Bewley, Truman F., "A Depressed Labor Market as Explained by Participants," *American Economic Review*, May 1995, *85* (2), 250–254. Papers and Proceedings.

Bowles, Samuel, "Policies Designed for Self-Interested Citizens May Undermine "the Moral Sentiments": Evidence from Economic Experiments," *Science*, June 2008, *320* (5883), 1605 — 1609.

Falk, Armin and Michael Kosfeld, "The Hidden Costs of Control," *American Economic Review*, 2006, *96* (5), 1611–1630.

Fehr, Ernst and Bettina Rockenbach, "Detrimental Effect of Sanctions on Human Altruism," *Nature*, March 2003, *422*, 137–140.

Frey, Bruno S. and Reto Jegen, "Motivation Crowding Theory: A Survey of Empirical Evidence," *Journal of Economic Surveys*, 2001, *15*, 589–611.

Friebel, Guido and Wendelin Schnedler, "Team Governance: Empowerment or Hierarchical Control," Discussion Paper 6575, C.E.P.R. November 2007.

Gneezy, Uri and Aldo Rustichini, "Pay Enough or Don't Pay at All," *The Quarterly Journal of Economics*, August 2000, *115* (3), 791–810.

Grossman, Sanford J. and Oliver D. Hart, "An Analysis of the Principal-Agent Problem," *Econometrica*, January 1983, *51* (1), 7–46.

Herold, Florian, "Contractual Incompleteness as a Signal of Trust," *Games and Economic Behavior*, forthcoming.

Holmström, Bengt, "Moral Hazard and Observability," *Bell Journal of Economics*, 1979, *10* (1), 74–91.

_ , "Moral Hazard in Teams," *Bell Journal of Economics*, 1982, *13* (2), 324–340.

_ **and Paul Milgrom**, "Multitask Principal-Agent-Analysis: Incentive Contracts, Asset Ownership, and Job Design," *Journal of Law, Economics, and Organization*, 1991, *7*, 24–52. special issue.

Kerr, Steven, "On the Folly of Rewarding A While Hoping for B," *Academy of Management Journal*, 1975, *18* (4), 769–783.

Mookherjee, Dilip, "Wealth Effects, Incentives, and Productivity," *Review of Development Economics*, February 1997, *1* (1), 116–33.

Schnedler, Wendelin, "The Virtue of Being Underestimated: A Note on Discriminatory Contracts in Hidden Information Models," *Economics Letters*, 2002, *75* (2), 171–178.

_ **and Radovan Vadovic**, "Legitimacy of Control," 2007. Discussion paper 3013, Institute for the Study of Labor (IZA), 2007.

Sliwka, Dirk, "On the Hidden Costs of Incentive Schemes," *American Economic Review*, 2007, *97* (3), 999–1012.

Thiele, Henrik and Achim Wambach, "Wealth Effects in the Principal Agent Model," *Journal of Economic Theory*, December 1999, *89* (2), 247–260.

Proofs

Lemma 1. *If the marginal utility of money weakly decreases in effort,* $\frac{\partial^2 u}{\partial m \partial e} \leq 0$*, a bonus for failure* ($t_0 > t_1$) *cannot be optimal.*

Proof. The agent's expected utility can be written as $u(t_0, e) + (1 - p(e)) \cdot (u(t_0, e) - u(t_1, e))$. Taking the derivative with respect to effort yields:

$$\frac{\partial u(t_0, e)}{\partial e} - p'(e) \cdot (u(t_0, e) - u(t_1, e)) + p(e) \cdot \left(\frac{\partial u(t_0, e)}{\partial e} - \frac{\partial u(t_1, e)}{\partial e} \right). \tag{2}$$

Given that $t_0 > t_1$ and $\frac{\partial^2 u}{\partial m \partial e} \leq 0$, this derivative is strictly negative. So, the agent has an incentive to reduce rather than increase effort. Since this reduction leads to costs for the principal, $t_0 > t_1$ is not optimal. □

Lemma 2.

$$\frac{\partial^2 u}{\partial m \partial e} \leq 0 \textit{ and } t_1 \geq t_0 \quad \Rightarrow \quad \frac{\partial^2 u}{\partial e^2} E_S \left[U^A(t(S), e) \right] < 0.$$

Proof. Taking the derivative of (4) yields:

$$\begin{aligned} &(1 - p(e)) \frac{\partial^2 u}{\partial e^2} + p''(e)(u(t_1, e) - u(t_0, e)) \\ &+2 \cdot p'(e) \left(\frac{\partial u(t_1, e)}{\partial e} - \frac{\partial u(t_0, e)}{\partial e} \right) + p(e) \left(\frac{\partial^2 u}{\partial e^2} \right). \end{aligned}$$

Using that p is concave and u strictly concave in e and that $t_1 \geq t_0$, the first line becomes negative. Since $\frac{\partial^2 u}{\partial m \partial e} \leq 0$ and by evoking once more the strict concavity of e, the second line is non-positive. □

Proof Proposition 1. Given (t_0, t_1), the agent's expected utility can be written as:

$$u(t_0, e) + p(e) \cdot (u(t_1, e) - u(t_0, e)). \tag{3}$$

The derivative is:

$$\frac{\partial u(t_0, e)}{\partial e} + p'(e) \cdot (u(t_1, e) - u(t_0, e)) + p(e) \cdot \left(\frac{\partial u(t_1, e)}{\partial e} - \frac{\partial u(t_0, e)}{\partial e} \right). \tag{4}$$

If effort is weakly normal and a success bonus is used, the agent's incentive problem is concave (see Lemma 2) and the first-order condition based on (4) is necessary and sufficient for a maximum.[3] Applying the implicit-function theorem yields:

$$\frac{de}{dt_1} = -(\phi(t_1, t_0, e))^{-1} \cdot \left(p'(e) \cdot \left. \frac{\partial u(t, e)}{\partial m} \right|_{t=0, e=e^0} + p(e) \cdot \left. \frac{\partial^2 u(t, e)}{\partial m \partial e} \right|_{t=0, e=e^0} \right),$$

[3]Interestingly, the agent's objective function is not necessarily concave if the marginal utility of transfers is increasing in effort.

where $\phi(t_1, t_0, e)$ is the first derivative of the agent's payoff with respect to e. Note that $\phi(t_1, t_0, e)$ is negative near $t_1 = t_0 = 0$ and $e = e^0$. The sign of the derivative is hence determined by the second factor which is negative whenever inequality (1) holds. By continuity, the result holds for any small t_1. $\square$

Proof of Proposition 2. Since $\frac{\partial^2 u(t,e)}{\partial m \partial e} < 0$, it holds by Lemma 2 that the agent's utility is concave and the incentive constraint can be represented by the first-order condition. The induced effort is then implicitly defined by the incentive constraint as a continuous and differentiable function $\tilde{e}$ of the bonus t_1. Applying the implicit-function theorem, we obtain that this function falls at $t = 0$ and $e^0 = 0$ because of (1). Next examine any $t' > 0$ such that $\tilde{e}(t') = e^0$. Observe that $\frac{\partial u}{\partial m}\big|_{t=t', e=e^0} \leq \frac{\partial u}{\partial m}\big|_{t=0, e=e^0}$ due to the concavity of utility in effort. So, the incentive effect is weakly smaller at t: $p'(e^0) \cdot \frac{\partial u(t,e)}{\partial m}\big|_{t=0, e=e^0} \geq p'(e^0) \cdot \frac{\partial u(t,e)}{\partial m}\big|_{t=t', e=e^0}$. Moreover, since $\frac{\partial}{\partial e}\frac{\partial^2 u(t,e)}{\partial m^2} < 0$, it follows that $(-1)\frac{\partial^2 u(t,e)}{\partial m \partial e}\big|_{t=0, e=e^0} > (-1)\frac{\partial^2 u(t,e)}{\partial m \partial e}\big|_{t=t', e=e^0}$. So, $\tilde{e}$ decreases in t for any t' such that $\tilde{e}(t') = e^0$ and hence never exceeds e^0. $\square$