Marginal productivity reward: A new justification*

Alexander W. Cappelen†and Bertil Tungodden‡

December 16, 2004

Abstract

Marginal productivity reward has been justified either as a way of ensuring efficiency or of respecting people's self-ownership. Both these arguments have their limitations. In this paper we present a new and more general justification of marginal productivity reward that avoids the limitations of the traditional arguments.

1   Introduction

Most people support some degree of redistribution, but typically also agree that people should be rewarded in accordance with the effort they exercise. Thus a fundamental question in the design of redistributive mechanisms is to what extent individual effort should be rewarded. One prominent answer to this question is that ideally people should be rewarded with their marginal productivity. This answer has traditionally been given two types of justifications. First, marginal productivity reward has been justified by efficiency considerations. Standard economic theory tells us that if we deviate

---

from marginal productivity reward, then there may be oversupply or undersupply of effort. Marginal productivity reward, however, ensures Pareto-efficiency. Secondly, it has been justified by equity considerations. According to some theories of distributive justice, in particular libertarianism (Nozick (1974)), marginal productivity reward is the only way to respect people's self-ownership (see also Kolm (1996)).

Both these arguments have their limitations. The efficiency argument only provides a justification for marginal productivity reward in situations where there are incentive problems. In situations where the supply of effort is inelastic, there is no efficiency reason for rewarding effort with its marginal product. The equity argument is problematic because it relies on some very controversial normative assumptions. Only people accepting the basic idea of full self-ownership and the view that full self-ownership implies marginal productivity reward, would be convinced by the libertarian equity argument. This position, however, is rejected both by utilitarians (for example Mirrlees (1971), Harsanyi (1978) and Broome (1991)) and liberal egalitarians (for example Rawls (1971), Fleurbaey (1995) and Moulin and Roemer (1989)).

In this paper we present a new and more general justification for marginal productivity reward. The justification applies even in the absence of incentive considerations and it relies on a much less controversial normative assumption than the self-ownership argument. We show that marginal productivity reward follows from a very appealing requirement, namely that people always should experience an increase in their post-tax income when they increase their effort. We name this the *non-negative reward requirement*. To illustrate, consider two situations a and b, where you work harder or longer hours in b than in a and thus have a higher total production in b than in a. The requirement then states that your post-tax income in b should not be lower than your post-tax income in a.

We present the formal framework in Section 2 and the proposition in Section 3. Section 4 provides some discussion of how to interpret the result.

2 Formal framework

Consider a society with a population $N = \{1, ..., n\}$, $n \geq 2$, where person i's effort is e_i and $e = (e_1, ..., e_n)$ is the effort distribution in situation e. Let Ω be the set of all effort distributions. We assume that all individuals can choose between all effort levels $e_i \in \left[e^{\min}, e^{\max}\right\rangle \subseteq \Re$, where $\Re$ is the set of

real numbers. The pre-tax income for each individual i, $f_i : \left[e^{\min}, e^{\max} \right\rangle \to$ $\Re$ is continuous and strictly increasing in effort. Our object of study is a redistribution mechanism F: $\Omega \to \Re^n$, where $F_i(e)$ is the post-tax income of person i in situation e. F satisfies the no-waste condition $\sum_{i=1}^{n} F_i(e) = \sum_{i=1}^{n} f_i(e)$, $\forall e \in \Omega$.

3 Rewarding effort

We argue that any redistributive system should aim at satisfying a minimal reward condition saying that persons increasing their effort, and thus increasing their pre-tax income, should not experience a decrease in their post-tax income. In other words, if your effort is higher in one situation than another, then your post-tax income should at least not be lower in the situation where you exercise more effort. Formally, we can write this requirement as follows.

Non-negative reward (NNR): For any $e, \tilde{e} \in \Omega$ and $j \in N$, where $\tilde{e}_j > e_j \to F_j(\tilde{e}) \geq F_j(e)$.

Surprisingly, it turns out that the non-negative reward requirement implies marginal productivity reward, and thus is incompatible with anything else than lump sum redistribution.

Proposition 1 *A redistribution mechanism F satisfies NNR if and only if effort is rewarded with its marginal productivity.*

Proof. The if part of the proposition is trivial. Hence, we will only prove the only-if part.

(i) Suppose there exist $e, \tilde{e} \in \Omega$ and $k \in N$ such that $\tilde{e}_k > e_k$ and $F_k(\tilde{e}) - F_k(e) \neq f_k(e_k) - f(e_k)$. Assume that $F_k(\tilde{e}) - F_k(e) > f_k(\tilde{e}_k) - f_k(e_k)$.

(ii) Consider a new situation $\hat{e} \in \Omega$ such that $\hat{e}_k = \tilde{e}_k + \epsilon$
and $\hat{e}_i = e_i + \epsilon, \forall i \neq k$

(iv) By NNR, $F_k(\hat{e}) \geq F_k(\tilde{e})$ and $F_i(\hat{e}) \geq F_i(e), \forall i \neq k$.

(v) By no-waste $\sum_i [F_i(\hat{e}) - F_i(e)] = \sum_i [f_i(\hat{e}_i) - f_i(e_i)]$. By rearranging, this can be written as $F_k(\hat{e}) - F_k(e) = \sum_i [f_i(\hat{e}_i) - f_i(e_i)] + \sum_{i \neq k} [F_i(e) - F_i(\hat{e})]$.

(vi) By continuity of f, for a sufficiently small ϵ, $\sum_i [f_i(\hat{e}_i) - f_i(e_i)] = [f_k(\tilde{e}_k) - f_k(e_k)] + \hat{\epsilon}$, where $\hat{\epsilon} < [F_k(\tilde{e}) - F_k(e)]$ - $[f_k(\tilde{e}_k) - f_k(e_k)]$. Hence, by (v), we have that $F_k(\hat{e}) - F_k(e) < [F_k(\tilde{e}) - F_k(e)]$ - $\sum_{i \neq k} [F_i(e) - F_i(\hat{e})]$. By (iv), $\sum_{i \neq k} [F_i(e) - F_i(\hat{e})] \leq 0$. Hence, $F_k(\hat{e}) - F_k(e) < F_k(\tilde{e}) - F_k(e)$, i.e., $F_k(\hat{e}) < F_k(\tilde{e})$. But this violates NNR, as stated in (iv).

(8) The proof is symmetric in the case where $F_k(\tilde{e}) - F_k(e) < f_k(\tilde{e}_k) - f_k(e_k)$, and the result follows. ∎

4   Discussion

The underlying intuition of the proposition is that any non lump-sum redistribution, that is, any system of redistribution where your net transfer depends on your effort, creates interdependences between the individuals in the economy, and the existence of such interdependences makes it impossible to satisfy the non-negative reward requirement.

To illustrate, consider a very simple economy with two individuals, person 1 and person 2. Individual i has the pre-tax income function $f(w_i, L_i) = w_i L_i$, where w_i is person i's marginal productivity and L_i is person i's labor effort. To make this an interesting case, assume that the two individuals differ in marginal productivity, i.e., $w_1 \neq w_2$ (even though the proof does not rely on this assumption). Moreover, assume that the government policy is limited to a linear tax scheme, where the tax income is shared equally among the two individuals in society. The post-tax incomes are then given by $F_1 = w_1 L_1 (1-t) + \frac{w_1 L_1 + w_2 L_2}{2} t$ and $F_2 = w_2 L_2 (1-t) + \frac{w_1 L_1 + w_2 L_2}{2} t$. It is easy to see that in this case, there is no redistributive tax scheme that satisfies the requirement of non-negative reward. For any positive t, there are situations in which person 1 has a lower post-tax income in a situation with high labor effort than in a situation with low labor effort, and similarly for person 2. The reason is that this redistributive mechanism creates an interdependence between the two individuals. The increase in post-tax income due to an increase in a person's own labor effort will sometimes be offset by a decrease in post-tax income due to a reduction in the other persons labor effort. The only way of avoiding this possibility is to have a the tax rate equal to zero and thereby eliminating the interdependence. Our proposition shows that this is not only a feature of a linear tax scheme with uniform transfers, but applies to *any* redistributive mechanism that does not rely on lump sum redistribution. In this respect, notice that we do not require the post-tax income of all individuals to be positive in all situations. We only require that the redistributive system balances, i.e., that the sum of post-tax incomes cannot exceed the sum of pre-tax incomes.

An alternative interpretation of the idea of non-negative reward, is the view that only in cases where we have a *unilateral* increase in effort should we

demand no decrease in a person's post-tax income. This very weak requirement does not imply marginal productivity reward and is consistent with any reasonable redistribution mechanism. But at the same time, we doubt that it captures all of our moral intuitions on how to reward effort. We find the idea that an increase in effort should imply no decrease in post-tax income, *independent* of what others do, extremely attractive, and thus we do believe that it is of much importance to observe that lump sum redistribution is the only redistributive policy that has this feature. There are some other problems with lump sum distribution, however, both from an an informational and distributive point of view, and thus it may be that it is impossible to satisfy all reasonable requirements in the design of a redistributive system.

References

[1] Broome, J. (1991) *Weighing Goods.* Basil Blackwell.

[2] Fleurbaey, M. (1995) "Three solutions for the compensation problem". *Journal of Economic Theory* 6: 96-106.

[3] Harsanyi, J. C. (1987) "Bayesian decision theory and utilitarian ethics". *American Economic Review*, Papers and Proceedings, 68: 223-228.

[4] Kolm, S-C. (2001) *Modern Theories of Justice.* MIT Press.

[5] Mirrlees, J. A. (1971) "An exploration in the theory of optimal taxation". *Review of Economic Studies* 38: 175-208.

[6] Moulin, H. and J. Roemer (1989) "Public ownership of the external world and private ownership of the self." *Journal of Political Economy* 97: 347-367.

[7] Nozick, R. (1974) *Anarchy, State, and Utopia.* Basic Books.

[8] Rawls, J. (1971) *A Theory of Justice.* Harvard University Press.