

Responsibility and Reward*

Alexander W. Cappelen[†] and Bertil Tungodden[‡]

September 3, 2001

Abstract

It is not straightforward to define the ethics of responsibility in cases where the consequences of factors under our control are partly affected by factors outside our control. One way to approach this issue is to ask how an increase in one individual's effort should be allowed to affect the post-tax income of others. In this paper we show how different answers to this question can be used to characterize interesting redistributive mechanisms.

[Discussion paper 25/01](#)

1 Introduction

Inequality can be seen as the result of two types of factors, those that are under the control of individuals and those that are outside their control. Let us refer to the first type of factors as *effort* and the second type of factors as *talent*. The fundamental moral intuition underlying the concept of equal opportunity is that society should accept inequalities that arise from differences in effort, *the principle of responsibility*, but should eliminate inequalities that are due to differences in talent, *the principle of compensation* (Dworkin

*We are grateful to Geir Asheim for a valuable discussion of the framework of this paper. The usual disclaimer applies.

[†]Norwegian School of Economics and Business Administration and the Norwegian Research Centre in Organization and Management, Bergen, Norway. e-mail: alexander.cappelen@nhh.no.

[‡]Norwegian School of Economics and Business Administration and Chr. Michelsen Institute, Bergen, Norway. e-mail: bertil.tungodden@nhh.no.

(1981), Arneson (1989)).¹ A central controversy in contemporary ethics concerns where to draw the ‘cut’ between those things that are within a persons control and those things that are outside a persons control.² However, even if we agree on how to make a distinction between effort and talent, it is not straightforward to design a redistributive scheme that satisfies both the ethics of responsibility and the ethics of compensation (see Bossert (1995), Fleurbaey (1994, 1995b, 1995c, 1995d), and Bossert and Fleurbaey (1996)). The reason is that it is not always possible to separate the consequences of effort and the consequences of talent. Unless the pre-tax income function is additively separable in effort and talent, the consequences of exercising high effort will depend on a person’s talent. According to the ethics of responsibility, a person with a specific talent should be held responsible for *exercising high effort*, but *not* for *being a person with a specific talent* exercising high effort. Hence, from the ethics of responsibility, it follows that a person should bear the consequences of exercising high effort per se, but not that she should bear the consequences of exercising high effort as a more or less talented person. The question is therefore to what extent we can reward effort without rewarding talent.

There are two ways to approach this question. The standard approach is to look at the effect on the person who changes her effort and ask what the appropriate reward for effort should be. Alternatively, one may look at people who do *not* change their effort and ask what the appropriate effect on their post-tax income should be when another person changes her effort. In this paper, we will explore this latter approach, which has not been much studied in the literature, and show how it can be used to characterize interesting redistributive mechanisms that attempt to compensate for differences in talent while at the same time reward effort.³

The analysis is based on the assumption that effort is unaffected by the tax system. We adopt this approach because we want to focus on how responsibility may justify rewarding effort even in the absence of incentive considerations. There is, of course, the further question about how to incorporate incentives in such a scheme, but we do believe that a clarification of the concept of responsibility is needed before we move to more complex situations.

¹For critical reviews of parts of this literature, see Fleurbaey (1995a) and Anderson (1999).

²See Roemer (1996) for an introduction to the philosophical literature on this issue.

³See also Tungodden (forthcoming) for a related discussion.

The rest of the paper is organized around different answers to the question about how an increase in effort by a person should be allowed to affect others post-tax income. After presenting the formal framework in section 2, we discuss in section 3 the requirement that the post-tax income of individuals should be independent of the increase in effort of other individuals. As is well-known in the literature, this requirement is not compatible with the ethics of compensation, and we provide a brief discussion of the reason for this impossibility. In section 4, we consider two ways of relaxing this requirement. First, we consider the implications of the view that the increase in one person's effort should have no negative effects on others, and, second, we consider the opposite view saying that the increase in one person's effort should have no positive effects on others. In section 5, we suggest a rather different point of view, saying that an individual should not gain or lose from the increase in effort of individuals who already exercise superior effort. Finally, in section 6, we provide some further comments on how our discussion of responsibility and reward relates to the more common incentive arguments in redistribution.

2 Analysis

Consider a society with a population $N = \{1, \dots, n\}$, $n \geq 2$, where agent i 's effort is a_i^E and her talent a_i^T . We assume that $a_i^E, a_i^T \in \mathbb{R}$, where \mathbb{R} is the set of real numbers.⁴ Let $a_i = \langle a_i^E, a_i^T \rangle$ be a characteristics vector of i , $a = \langle a_1, \dots, a_n \rangle$ a characteristics profile of society (which can be partitioned into $a^E = \langle a_1^E, \dots, a_n^E \rangle$ and $a^T = \langle a_1^T, \dots, a_n^T \rangle$), $\Omega \subseteq \mathbb{R}^2$ the set of all possible characteristics vectors (where Ω_E is the set of all possible effort levels and Ω_T the set of all possible talents), and $\Omega^n \subseteq \mathbb{R}^{2n}$ the set of all possible characteristics profiles. Let $\tilde{\Omega}^n \subset \Omega^n$ be the set of admissible characteristics profiles, where for any $a, \tilde{a} \in \tilde{\Omega}^n$, $a^T = \tilde{a}^T$. In other words, we do not consider interprofile conditions with respect to talent,⁵ but assume that there is a single characteristics profile of talent in society. This profile, however, can be any profile within the set of possible profiles. We impose no other restrictions on the set of permissible characteristics vectors and profiles.

The income function $f : \Omega \rightarrow \mathbb{R}$ is assumed to be strictly increasing in

⁴Hence, we do not consider the multidimensional version of this problem: see Bossert and Fleurbaey (1996).

⁵See Bossert (1995).

both arguments and regular. An income function is regular if and only if for any $i, j \in N$, $a_i^T, a_j^T \in \mathbb{R}$ and any $a_i^E, \tilde{a}_i^E, \hat{a}_i^E \in \mathbb{R}$, where $a_i^E > \tilde{a}_i^E > \hat{a}_i^E$, $f(a_i^T, \tilde{a}_i^E) - f(a_i^T, \hat{a}_i^E) > f(a_j^T, \tilde{a}_j^E) - f(a_j^T, \hat{a}_j^E) \rightarrow f(a_i^T, a_i^E) - f(a_i^T, \tilde{a}_i^E) \geq f(a_j^T, a_j^E) - f(a_j^T, \tilde{a}_j^E)$.⁶ To simplify the analysis, but without loss of generality, we also assume that the marginal productivity of the more talented always is at least as great as the marginal productivity of the low talented, i.e. for any $i, j \in N$, $a_i^T, a_j^T \in \mathbb{R}$ and $a_i^E, \tilde{a}_i^E \in \mathbb{R}$, where $a_i^E > \tilde{a}_i^E$, $a_i^T > a_j^T \rightarrow f(a_i^T, a_i^E) - f(a_i^T, \tilde{a}_i^E) \geq f(a_j^T, a_j^E) - f(a_j^T, \tilde{a}_j^E)$. Finally, an efficient redistribution function $F: \tilde{\Omega}^n \rightarrow \mathbb{R}^n$ satisfies the feasibility condition $\sum_{i=1}^n F_i(a) = \sum_{i=1}^n f(a_i)$, $\forall a \in \tilde{\Omega}^n$.

By the ethics of compensation, talent is an irrelevant factor and cannot justify any inequalities. Thus, if two persons only differ in talent and not in effort, the redistributive scheme should assign the same post-tax income to both individuals. Formally, this requirement can be written as follows.

Equal Income for Equal Effort (EIEE): For any $a \in \tilde{\Omega}^n$ and $i, j \in N$, $a_i^E = a_j^E \rightarrow F_i(a) = F_j(a)$.

EIEE is an unquestionable implication of the ethics of compensation, and throughout the paper we will maintain this requirement on the redistributive mechanism.

It is more problematic to determine how income should be distributed when considering situations where individuals exercise different levels of effort. The ethics of responsibility suggests that people should be held responsible for their choice of effort, but not for being a person with a specific talent exercising a particular effort level. In general, it is not obvious how to interpret this idea, and therefore we will approach this problem by looking at what type of effects a change in one person's effort should be allowed to have on others.

3 No effect on others

Intuitively it might seem plausible to claim that the post-tax income of those individuals who do not change their effort should be unaffected by the change

⁶In most cases, this condition should be considered acceptable. By way of illustration, it is common to assume (for example in labour markets) that the marginal productivity of the more talented is higher than of the less talented for every effort level. However, the regularity condition does not rule out the opposite case. What it rules out is that there is a change in the ranking of marginal productivity at a certain effort level.

in effort of other people. This intuition is captured by the following requirement.

No effect on others (NE): For any $a, \tilde{a} \in \tilde{\Omega}^n$ and $k \in N$, $a_j^E = \tilde{a}_j^E, \forall j \neq k \rightarrow F_j(a) = F_j(\tilde{a})$.⁷

According to NE, a change in total pre-tax income due to a change in one agent's effort should only affect this person's post-tax income. However, as shown by Bossert (1995), there does not exist any efficient redistribution mechanism that satisfies this requirement and EIEE unless the pre-tax income function is additively separable, i.e. unless the marginal productivity of effort is independent of talent.

We can illustrate this result by considering a simple example. Assume that there are two people in society, one with high talent and one with low talent. They can either exercise high or low effort, and let Figure 1 give their pre-tax income as a function of effort.

		Effort	
		<i>High</i>	<i>Low</i>
Talent	<i>High</i>	100	20
	<i>Low</i>	70	0

Figure 1

We note that in Figure 1, the gain in pre-tax income following an increase in effort depends on talent, where the marginal productivity of the high talented is greater than of the low talented.

The redistributive scheme has to cover four cases in this economy.

		Low Talent	
		<i>High Effort</i>	<i>Low Effort</i>
High Talent	<i>High Effort</i>	case 1	case 2
	<i>Low Effort</i>	case 3	case 4

Figure 2

Let us first look at case 4 where the individuals exercise the same amount of effort and thus only differ in talent. EIEE demands an equal split in this situation. Moreover, according to NE, the high talented should not be affected by a move from case 4 to case 3 and the low talented should not be affected by a move from case 3 to case 1. Hence, if we take case 4 as the

⁷This condition is usually referred to as Individual Monotonicity in Effort in the literature (see for example Bossert and Fleurbaey (1996)). However, in order to focus on the similarities between this requirement and other requirements introduced in this paper, we find the renaming of the condition useful.

point of departure and assume an efficient redistributive scheme, NE has the following implications.

		Low Talent	
		<i>High Effort</i>	<i>Low Effort</i>
High Talent	<i>High Effort</i>	90,80	
	<i>Low Effort</i>	10,80	10,10

Figure 3

As is easily seen, the redistributive scheme in Figure 3 violates EIEE in case 1. And it turns out that this conflict is present in any situation where the pre-tax income function is not additively separable (Bossert (1995)), unless we give up the demand of an efficient redistributive scheme.

If we accept inefficiency, then we can satisfy both EIEE and NE by letting the reward of effort be equal to the marginal productivity of the least talented, as illustrated in Figure 4.

		Low Talent	
		<i>High Effort</i>	<i>Low Effort</i>
High Talent	<i>High Effort</i>	80,80	80,10
	<i>Low Effort</i>	10,80	10,10

Figure 4

However, we should like to maintain both the requirement of efficiency and EIEE, and hence our only option is to relax NE. This means that people's increase in effort will not always be rewarded by their marginal productivity, and as a consequence there will be a deficit or a surplus to be distributed equal to the difference between the change in pre-tax income and the change in post-tax income of the persons increasing their effort. In other words, some individuals must experience an increase or a decrease in post-tax income as a result of an increase in effort by other individuals. But who should be affected? And how? In the rest of the paper, we look at the implications of different answers to these questions. Let us stress, however, that our aim is not to defend a right answer, but rather to clarify the implications of different views and thereby contribute to the process of establishing a reflective equilibrium in the Rawlsian sense on this issue.

“When a person is presented with an intuitively appealing account of his sense of justice (one, say, which embodies various reasonable and natural presumptions), he may well revise his judgments to conform to its principles even though the theory

does not fit his existing judgments exactly. He is especially likely to do this if he can find an explanation for the deviations which undermines his confidence in his original judgments and if the conception presented yields a judgment which he finds he can now accept. From the standpoint of moral philosophy, the best account of a person's sense of justice is not the one which fits his judgments prior to his examining any conception of justice, but rather the one which matches his judgments in reflective equilibrium. As we have seen, this state is one reached after a person has weighed various proposed conceptions and he has either revised his judgments to accord with one of them or held fast to his initial convictions (and the corresponding conception)" (Rawls, 1971, p. 48).

With this in mind, we move on to a discussion of different ways of restricting the type of effects an increase in effort by someone should have on others.

4 Restricting the type of effects on others

If we want to satisfy the ethics of compensation within an efficient redistributive scheme, there will be cases where there is a sacrifice or loss to be distributed among those who do not change their effort. How should that be done? An intuitively appealing idea is to say that the surplus or deficit should be split equally, and we will take this idea as the starting point for the present discussion. In section 5, however, we will argue that it is not obvious that we should accept this view, because there might be morally relevant differences among those who are not affected that should make us support an unequal split.

Formally speaking, we can state our basic condition as follows.

Equal Effect on Others (EE): For any $a, \tilde{a} \in \tilde{\Omega}^n$ and $j, k \in \{i \in N \mid a_i^E = \tilde{a}_i^E\}$, $F_j(a) - F_j(\tilde{a}) = F_k(a) - F_k(\tilde{a})$.

EE and EIEE characterize a large class of redistribution schemes, which includes strict egalitarianism (saying that we should always split the resources equally). Certainly, strict egalitarianism does not capture our concern for rewarding effort, and thus immediately we should like to narrow this class somewhat by imposing a restriction on the redistributive scheme that guarantees that the reward of effort is not below the marginal productivity of the

least talented and not above the marginal productivity of the most talented. In our view, this restriction follows directly from any reasonable interpretation of the ethics of responsibility. Let a_{\min}^T refer to the minimal talent and a_{\max}^T to the maximal talent in the population. The minimal restriction we impose on the ethics of responsibility can now be captured by the following condition.

Restricted Reward (RR): For any $a \in \tilde{\Omega}^n$ and $j, k \in N$, $a_j^E > a_k^E \rightarrow f(a_{\min}^T, a_j^E) - f(a_{\min}^T, a_k^E) \leq F_j(a) - F_k(a) \leq f(a_{\max}^T, a_j^E) - f(a_{\max}^T, a_k^E)$.

Even though we narrow the class of permissible redistributive schemes somewhat by imposing RR, there is still a vast number of perspectives satisfying all our present demands. Among them the so-called egalitarian equivalent mechanisms F^{EE} introduced by Bossert and Fleurbaey (1996), which will turn out to be of particular interest in the rest of this discussion.

$F_k^{EE}(a) := f(a_k^E, t^{REF}) - \frac{1}{n} \sum_{i \in N} [f(a_i^E, t^{REF}) - f(a_i)]$, $\forall k \in N$, $\forall a \in \tilde{\Omega}^n$.

F^{EE} assigns to every agent a post-tax income equal to the pre-tax income she would earn if her talent were equal to the reference talent t^{REF} , plus a uniform transfer. The uniform transfer secures that the redistributive mechanism is efficient. It is rather easy to see that any version of F^{EE} satisfies EIEE, EE, and RR (as long as we assume that the reference talent is above the lowest and below the highest talent in society). Bossert and Fleurbaey (1996) provide characterisations of this class of redistributive schemes, but *not* of any single mechanism within this class. However, they underline that “the choice of a particular reference [talent] is an important issue” (p. 344), and in the following we will provide characterisations that turn out to be useful in this choice. In particular, we will present complete characterisations of two egalitarian equivalent mechanisms, where the least and most talented, respectively, is the reference talent.

In order to do this, however we will have to say more about the type of effects that an increase in effort should have on others. Two opposing views easily come to mind. On the one hand, we may argue that there should be no negative effects on others; on the other hand, we may argue that there should be no positive effects on others. Let us consider the implications of each of these views?

4.1 No negative effect on others

If we approach the question about how an increase in effort should affect others from the perspective of those who do not change their effort, it may seem reasonable to argue that they should not receive a lower post-tax income because others increase their effort. The increase in effort is the responsibility of the person increasing her effort, and it might seem unfair that others should experience a loss in such a situation.

This requirement can be captured by the following principle.

No negative effect on others (NNE): For any $j \in N$ and $a, \tilde{a} \in \tilde{\Omega}^n$, where $\tilde{a}_j^E > a_j^E$ and $a_i = \tilde{a}_i \forall i \neq j$, $F_i(\tilde{a}) \geq F_i(a), \forall i \neq j$.

By adding NNE to our framework, we obtain a complete characterisation of a particular egalitarian equivalent mechanism.

Theorem 1 *A redistribution mechanism F satisfies EIEE, EE, RR and NNE if and only if $F = F^{EE}$ and the reference talent equals the lowest talent in society.*

Proof. The if-part of the theorem is trivial, and hence we will only prove the only-if part.

(i) Suppose that there exists an F satisfying EIEE, NNE, EE and RR such that for some $a \in \tilde{\Omega}^n$ and $j, k \in N$, $F_j(a) - F_k(a) \neq f(a_{\min}^T, a_j^E) - f(a_{\min}^T, a_k^E)$. Obviously, by EIEE, this is not possible if $a_j^E = a_k^E$. Hence, without loss of generality, we assume that $a_j^E > a_k^E$.

(ii) By RR and (i), it follows that $F_j(a) - F_k(a) > f(a_{\min}^T, a_j^E) - f(a_{\min}^T, a_k^E)$.

(iii) Consider some $\tilde{a} \in \tilde{\Omega}^n$, where $\tilde{a}_i^E = a_k^E, \forall i \in N$. By EIEE, $F_i(\tilde{a}) = F_k(\tilde{a}), \forall i \in N$.

(iv) Let m refer to the least talented in society. Consider some $\hat{a} \in \tilde{\Omega}^n$, where $\hat{a}_i^E = \tilde{a}_i^E, \forall i \neq m$ and $\hat{a}_m^E = a_j^E$. By NNE, $F_i(\hat{a}) \geq F_i(\tilde{a}), \forall i \neq m$.

(v) Moreover, by construction, $\sum_{i \in N} f(\hat{a}_i) - \sum_{i \in N} f(\tilde{a}_i) = f(a_{\min}^T, a_j^E) - f(a_{\min}^T, a_k^E)$. Hence, taking into account the result established in (iv), $F_m(\hat{a}) - F_m(\tilde{a}) \leq f(a_{\min}^T, a_j^E) - f(a_{\min}^T, a_k^E)$, and consequently $F_m(\hat{a}) - F_i(\hat{a}) \leq f(a_{\min}^T, a_j^E) - f(a_{\min}^T, a_k^E), \forall i \neq m$.

(vi) Consider some $\bar{a} \in \tilde{\Omega}^n$, where $\bar{a}_i = \hat{a}_i, \forall i \neq j, k$ and $\bar{a}_j = a_j$ and $\bar{a}_k = a_k$. By EE, $F_m(\bar{a}) - F_i(\bar{a}) = F_m(\hat{a}) - F_i(\hat{a}), \forall i \neq j, k$. Moreover, by EIEE, $F_j(\bar{a}) = F_m(\bar{a})$ and $F_k(\bar{a}) = F_i(\bar{a}), \forall i \neq j, m$. Hence, $F_j(\bar{a}) - F_k(\bar{a}) = F_m(\hat{a}) - F_i(\hat{a}) \leq f(a_{\min}^T, a_j^E) - f(a_{\min}^T, a_k^E)$. By EE, $F_j(a) - F_k(a) = F_j(\bar{a}) - F_k(\bar{a})$, which violates (ii). Hence, the supposition in (i) is not possible.

(vii) Consequently, we know that for any $a \in \tilde{\Omega}^n$ and $j, k \in N$, $F_j(a) - F_k(a) = f(a_{\min}^T, a_j^E) - f(a_{\min}^T, a_k^E)$. Moreover, by efficiency, $F_k(a) = \sum_{i \in N} f(a_i) - \sum_{i \neq k} F_i(a)$ and similarly for j . Without loss of generality, let us only consider the post-tax income of k . By straightforward manipulation, we find that $nF_k(a) = \sum_{i \in N} f(a_i) - \sum_{i \in N} (F_i(a) - F_k(a))$, and thus, taking into account the result established in (vi), we have that $nF_k(a) = \sum_{i \in N} f(a_i) - \sum_{i \in N} (f(a_{\min}^T, a_i^E) - f(a_{\min}^T, a_k^E))$. Hence, $F_k(a) = f(a_{\min}^T, a_k^E) - \frac{1}{n} \sum_{i \in N} [f(a_{\min}^T, a_i^E) - f(a_i)]$, which completes the proof. ■

Let us briefly illustrate how this redistributive mechanism works in the example considered in section 3.

		Low Talent	
		High	Low
High Talent	High	85,85	85,15
	Low	10,80	10,10

Figure 5

As we can see from Figure 5, the difference between the marginal productivity of the less talented and the marginal productivity of the more talented is shared equally when the more talented increases her effort. More generally, there will be an equal split of such a surplus among all the members of society.

An immediate objection to this property of the redistribution function is that it is unreasonable that the low talented should gain simply by the fact that the more talented increases her effort. It can be argued that there is no justification for a compensation of this kind, and we now turn to a discussion of the implications of endorsing such a perspective.

4.2 No positive effect on others

If we take the perspective of the individual who increases her effort, it may be argued that there is no reason that others should benefit from such a move. This intuition is captured by the following requirement.

No positive effect on others (NPE): For any $j \in N$ and $a, \tilde{a} \in \tilde{\Omega}^n$, where $\tilde{a}_j^E > a_j^E$ and $a_i = \tilde{a}_i \forall i \neq j$, $F_i(\tilde{a}) \leq F_i(a), \forall i \in N$.

By substituting NPE for NNE in our framework, we obtain a complete reversal of the conclusion in the previous subsection.

Theorem 2 *An efficient redistribution mechanism F satisfies EIEE, EE, RR and NPE if and only if $F = F^{EE}$ and the reference talent equals the*

highest talent in society.

Proof. The if-part of the theorem is trivial, and hence we will only prove the only-if part.

(i) Suppose that there exists an F satisfying EIEE, NPE, EE and RR such that for some $a \in \tilde{\Omega}^n$ and $j, k \in N$, $F_j(a) - F_k(a) \neq f(a_{\max}^T, a_j^E) - f(a_{\max}^T, a_k^E)$. Obviously, by EIEE, this is not possible if $a_j^E = a_k^E$. Hence, without loss of generality, we assume that $a_j^E > a_k^E$.

(ii) By RR and (i), it follows that $F_j(a) - F_k(a) < f(a_{\max}^T, a_j^E) - f(a_{\max}^T, a_k^E)$.

(iii) Consider some $\tilde{a} \in \tilde{\Omega}^n$, where $\tilde{a}_i^E = a_k^E, \forall i \in N$. By EIEE, $F_i(\tilde{a}) = F_k(\tilde{a}), \forall i \in N$.

(iv) Let m refer to the most talented in society. Consider some $\hat{a} \in \tilde{\Omega}^n$, where $\hat{a}_i^E = \tilde{a}_i^E, \forall i \neq m$ and $\hat{a}_m^E = a_j^E$. By NPE, $F_i(\hat{a}_i) \leq F_i(\tilde{a}_i), \forall i \neq m$.

(v) Moreover, by construction, $\sum_{i \in N} f(\hat{a}_i) - \sum_{i \in N} f(\tilde{a}_i) = f(a_{\max}^T, a_j^E) - f(a_{\max}^T, a_k^E)$. Hence, taking into account the result established in (iv) and the efficiency of F , $F_m(\hat{a}) - F_m(\tilde{a}) \geq f(a_{\max}^T, a_j^E) - f(a_{\max}^T, a_k^E)$, and consequently $F_m(\hat{a}) - F_i(\hat{a}) \geq f(a_{\max}^T, a_j^E) - f(a_{\max}^T, a_k^E), \forall i \neq m$.

(vi) Consider some $\bar{a} \in \tilde{\Omega}^n$, where $\bar{a}_i = \hat{a}_i, \forall i \neq j, k$ and $\bar{a}_j = a_j$ and $\bar{a}_k = a_k$. By EE, $F_m(\bar{a}) - F_i(\bar{a}) = F_m(\hat{a}) - F_i(\hat{a}), \forall i \neq j, k$. Moreover, by EIEE, $F_j(\bar{a}) = F_m(\bar{a})$ and $F_k(\bar{a}) = F_i(\bar{a}), \forall i \neq j, m$. Hence, $F_j(\bar{a}) - F_k(\bar{a}) = F_m(\hat{a}) - F_i(\hat{a}) \geq f(a_{\max}^T, a_j^E) - f(a_{\max}^T, a_k^E)$. By EE, $F_j(a) - F_k(a) = F_j(\bar{a}) - F_k(\bar{a})$, which violates (ii). Hence, the supposition in (i) is not possible.

(vii) Consequently, we know that for any $a \in \tilde{\Omega}^n$ and $j, k \in N$, $F_j(a) - F_k(a) = f(a_{\max}^T, a_j^E) - f(a_{\max}^T, a_k^E)$. Moreover, by efficiency, $F_k(a) = \sum_{i \in N} f(a_i) - \sum_{i \neq k} F_i(a)$ and similarly for j . Without loss of generality, let us only consider the post-tax income of k . By straightforward manipulation, we find that $nF_k(a) = \sum_{i \in N} f(a_i) - \sum_{i \in N} (F_i(a) - F_k(a))$, and thus taking into account the result established in (vi), $nF_k(a) = \sum_{i \in N} f(a_i) - \sum_{i \in N} (f(a_{\max}^T, a_i^E) - f(a_{\max}^T, a_k^E))$. Hence, $F_k(a) = f(a_k^E, a_{\max}^T) - \frac{1}{n} \sum_{i \in N} [f(a_i^E, a_{\max}^T) - f(a_i)]$, which completes the proof. ■

Let us briefly illustrate how F^{EE} works when the most talented is the reference talent in our example.

		Low Talent	
		<i>High</i>	<i>Low</i>
High Talent	<i>High</i>	85,85	90,10
	<i>Low</i>	5,85	10,10

Figure 6

As we can see from Figure 6, the consequence of imposing NPE is that the more talented experiences a loss when the less talented increases her effort. The reward assigned to the less talented equals the marginal productivity of the more talented, and this creates a deficit which in this example is shared equally by the two persons. More generally, there will be an equal split of such a deficit among all the members of society.

4.3 Trade-offs between positive and negative effects

We have so far considered what can be viewed as the extreme positions in this debate, where either we only assign importance to avoiding positive or to avoiding negative effects on others. An intermediate position would be to find both effects problematic, and thus to seek trade-offs between them. One way of doing that would be to adopt equivalent egalitarian mechanisms that are in between the polar cases discussed in the two previous subsections. Bossert and Fleurbaey (1996) considers one such option in particular, to wit when the reference talent equals the average talent in the population. In our standard example, this mechanism works as illustrated in Figure 7.

		Low Talent	
		<i>High</i>	<i>Low</i>
High Talent	<i>High</i>	85,85	87.5,12.5
	<i>Low</i>	7.5,82.5	10,10

Figure 7

As we can see, when the average talent is the reference talent, there are both negative and positive effects on others. When the more talented increases her effort, the less talented gains, whereas, when the less talented increases her effort, the more talented loses. However, comparing this mechanism with the two polar cases, one should notice that allowing for both positive and negative effects reduces the size of the effect on others. In Figure 7, the size of the loss (gain) imposed on the other person is 2.5, whereas the loss (gain) in Figure 5 (Figure 6) is 5.

This may give us a hint about how to characterise the average mechanism, to wit by imposing a condition saying that we should seek to minimize the overall effect on others. However, we will not pursue this line of reasoning here, but rather turn to a discussion of an alternative perspective on this problem. So far, we have worked within the framework of EE, saying that any sacrifice or loss to be distributed among those who do not change their effort should be split equally. This might seem uncontroversial, but we should

now like to argue that it is not.

5 Restricting who should be affected

Consider a situation where the only factor under a person's control is the number of hours she works. Let us furthermore assume that the hourly wage given to a high talented person is less than her marginal productivity. In this situation an increase in the hours she works will create a surplus that has to be distributed among the individuals in the population. Is it fair that this surplus is distributed equally among all individuals independent of how many hours they work? Or is it more reasonable that the share of burdens and benefits should somehow be related to how many hours a person work so as to give those who work long hours a larger share? In other words, when considering how to split the surplus among those who do not increase their effort, should we consider their effort level a morally relevant factor that justifies an unequal split?

One way to relate the distribution of burdens and benefits to a person's effort level is to hold that the post-tax income of a person should not be affected if people who already exercise a higher effort level increase their effort even further. This view could be justified by viewing each effort level as a cooperative venture and argue that the benefits or cost associated with any such cooperative venture should only be distributed among those who participate in that venture. Formally, we capture the perspective by the following condition.

No effect from superior effort (NESE): For any $j \in N$ and $a, \tilde{a} \in \tilde{\Omega}^n$, where $\tilde{a}_j^E = a_j^E$ and $\tilde{a}_k^E > a_k^E \geq a_j^E, \forall k \in \{i \in N \mid \tilde{a}_i^E \neq a_i^E\} \longrightarrow F_j(\tilde{a}) = F_j(a)$.

One part of this requirement, that you should not gain from others superior effort, can also be seen as capturing the idea that 'you can not complain if you do not try', i.e. only those who make a comparable effort can make a claim on the gains from others effort.

While NNE and NPE were compatible with EE on a general basis, this is not the case for NESE.

Theorem 3 *There does not exist any efficient redistribution mechanism F satisfying EIEE, EE and NESE, unless f is additively separable.*

Proof. It is easily seen that the conditions are compatible if f is additively separable, and hence we will only prove that they are not compatible on a general basis.

(i) Suppose F satisfies EIEE, EE and NESE. Consider some $a \in \tilde{\Omega}^n$, where for some $j, k \in N$, $a_j^T > a_k^T$ and $a_i^E = a_j^E = a_k^E, \forall i \in N$. By EIEE, $F_j(a) = F_k(a)$.

(ii) Consider $\tilde{a} \in \tilde{\Omega}^n$, where $\tilde{a}_i = a_i, \forall i \neq j$ and $\tilde{a}_j^E > a_j^E$. By NESE, $F_i(\tilde{a}) = F_i(a), \forall i \neq j$. Hence, from the efficiency of F , it follows that $F_j(\tilde{a}) - F_j(a) = f(\tilde{a}_j) - f(a_j)$.

(iii) Consider $\hat{a} \in \tilde{\Omega}^n$, where $\hat{a}_i = \tilde{a}_i, \forall i \neq k$ and $\hat{a}_k^E = \tilde{a}_j^E$. By NESE, $F_i(\hat{a}) = F_i(\tilde{a}), \forall i \neq j, k$. Moreover, by EE, $F_j(\hat{a}) - F_i(\hat{a}) = F_j(\tilde{a}) - F_i(\tilde{a}), \forall i \neq j, k$. Hence, $F_j(\hat{a}) = F_j(\tilde{a})$ and from the efficiency of F it follows that $F_k(\hat{a}) - F_k(\tilde{a}) = f(\hat{a}_k) - f(\tilde{a}_k)$.

(iv) By assumption, $a_j^T > a_k^T$. If f is not additively separable, we know that there exist effort levels such that the marginal productivity of the more talented is strictly above the marginal productivity of the less talented. Let us assume that this is the case for a_j^E and \tilde{a}_j^E , which implies that $f(\tilde{a}_j) - f(a_j) > f(\hat{a}_k) - f(\tilde{a}_k)$. Consequently, taking into account the results established in (ii) and (iii), $F_j(\hat{a}) > F_k(\hat{a})$. But this violates EIEE, and hence the supposition in (i) is not possible. ■

Let us now look at the implications of substituting NESE for EE within our framework. For this purpose, consider the following redistribution mechanism, where for any $j \in N$, $a_{\min+j}^E$ refers to the j th least talented in society and $n(a_l^E)$ is the cardinality of $N(a_l^E) = \{i \in N \mid a_i^E \geq l\}$.

$$F_k^{NE}(a) := \frac{1}{n} \sum_{i \in N} f(a_i^T, a_{\min}^E) + \sum_{l=a_{\min+1}^E}^{a_k^E} \frac{1}{n(a_l^E)} \sum_{i \in N(a_l^E)} [f(a_i^T, a_l^E) - f(a_i^T, a_{l-1}^E)],$$

$\forall k \in N, \forall a \in \tilde{\Omega}^n$.

F^{NE} works in the following way. First, all individuals are given an equal share of the the total pre-tax income that would have been produced if all individuals exerted the minimal effort level actually exercised in society. And persons actually exercising the minimal effort level does not get anything more than this. The rest of the population also receives a share equal to $\frac{1}{n(a_{\min+1}^E)}$ of the increase in total pre-tax income (compared to the case where everyone exercises the minimal effort level) due to the fact that these people exercise at least the effort level $a_{\min+1}^E$. Persons actually exercising the second lowest effort level do not get anything more than the total of these two shares,

and in the same manner we move to the third lowest effort level and so on.

We can illustrate F^{NE} by looking at the post-tax income distribution of our standard example.

	Low Talent	
	<i>High</i>	<i>Low</i>
High Talent	<i>High</i>	85,85 90,10
	<i>Low</i>	10,80 10,10

Figure 8

The difference between this approach and the egalitarian equivalent mechanisms is best illustrated by comparing the move from case 4 to case 2 with the move from case 3 to case 1 in our example. In both situations, the only thing that happens is that the more talented increases her effort. But in one case, when going from 4 to 2, the present approach endorses that the less talented is not affected, because the less talented does not exercise a comparable level of effort, whereas in the other case the approach demands a compensation to the less talented. More generally, this approach justifies an unequal split of any surplus or deficit to be distributed among those who do not increase their effort, because it views the actual effort level exercised by these people a morally relevant factor.

It turns out that F^{NE} is the only redistribution mechanism compatible with NESE and EIEE.

Theorem 4 *An efficient redistribution mechanism F satisfies EIEE and NESE if and only if $F = F^{NE}$.*

Proof. It is easily seen that F^{NE} satisfies *EIEE* and *NESE*. Hence, we will only prove that if F satisfies *EIEE* and *NESE*, then $F = F^{NE}$. To simplify the proof and without loss of generality, we assume that $a_1^E \leq a_2^E \leq \dots a_{n-1}^E \leq a_n^E$ for any $a \in \tilde{\Omega}^n$.

(i) Consider any $a \in \tilde{\Omega}^n$ and $k \in N$ and the following sequence:

$$\begin{aligned}
 a^1 &= (a_1^E, \dots, a_1^E), \\
 a^2 &= (a_1^E, a_2^E, \dots, a_2^E), \\
 &\vdots \\
 a^{k-1} &= (a_1^E, a_2^E, \dots, a_{k-1}^E, \dots, a_{k-1}^E), \\
 a^k &= (a_1^E, a_2^E, \dots, a_{k-1}^E, a_k^E, \dots, a_k^E).
 \end{aligned}$$

By NESE, for any t in the sequence, it follows that $F_i(\overset{t+1}{a}) = F_i(\overset{t}{a}), \forall i \leq t$.

(ii) By EIEE, it follows that $F_k(\overset{t}{a}) = F_i(\overset{t}{a}), \forall i \geq t \leq k$.

(iii) Hence, by the results established in (i) and (ii) and the efficiency of F , we have that $F_k(\overset{t+1}{a}) = F_k(\overset{t}{a}) + \frac{1}{n(a_{t+1}^E)} \sum_{i \in N(a_{t+1}^E)} [f(a_i^T, a_{t+1}^E) - f(a_i^T, a_t^E)]$.

Consequently, $F_k(\overset{k}{a}) = F_k(\overset{1}{a}) + \sum_{l=2}^{a_k^E} \frac{1}{n(a_l^E)} \sum_{i \in N(a_l^E)} [f(a_i^T, a_l^E) - f(a_i^T, a_{l-1}^E)]$.

(iv) By NESE, $F_k(\overset{k}{a}) = F_k(a)$, and the result follows from noticing that $F_k(\overset{1}{a}) = \frac{1}{n} \sum_{i=1}^n f(a_i^T, a_{\min}^E)$. ■

6 Responsibility and incentives

We have discussed different redistributive mechanisms under the assumption that effort is unaffected by the tax system. The justification for this approach was that we wanted to focus on the relationship between responsibility and reward. The ethics of responsibility presents a justification for rewarding effort even in the absence of incentive considerations. However, if the reward from effort differs from the increase in one's pre-tax income, some individuals must experience an increase or a decrease in post-tax income as a result of an increase in effort by others. We have shown that restrictions on who should be affected and how they should be allowed to be affected, have important implications for our choice of redistributive mechanism.

This discussion is of interest because economic theory has been preoccupied with welfaristic normative theories where considerations of responsibility do not play any independent role. Welfaristic theories view reward as an incentive used to induce certain behavior, and not as something of intrinsic importance in normative reasoning. Let us provide a simple illustration of this difference. Assume that the utility functions of the individuals are additively separable in income and leisure and, moreover, that the supply of labor is perfectly inelastic. In this situation, there will be no reason for a welfarist to reward those individuals who work longer hours. The only reason to deviate from strict egalitarianism would be if some individual had a higher marginal utility from income. However, common sense, and important non-welfarist normative traditions, hold that effort should be rewarded even if the reward does not affect behavior. People deserve a certain reward

for their effort irrespectively of the incentive effect. Clearly, incentive effects are also of much importance, and an interesting question for further research is therefore how incentive and desert considerations can be combined in a redistributive framework.⁸

References

- [1] Anderson, E. (1999) "What is the point of equality". *Ethics* 109: 287-337.
- [2] Arneson (1989) "Equality and equal opportunity for welfare". *Philosophical Studies* 56: 159-194.
- [3] Bossert, W. (1995) "Redistribution mechanisms based on individual characteristics". *Mathematical Social Sciences* 29: 1-17.
- [4] Bossert, W. and M. Fleurbaey (1996) "Redistribution and compensation". *Social Choice and Welfare* 13: 343-355.
- [5] Dworkin, R. (1981) "What is equality? Part 2: equality of resources". *Philosophy and Public Affairs* 10: 283-345.
- [6] Fleurbaey, M. (1994) "On fair compensation". *Theory and Decision* 36: 277-307.
- [7] Fleurbaey, M. (1995a) "Equal opportunity or equal social outcome". *Economics and Philosophy* 11: 25-55.
- [8] Fleurbaey, M. (1995b) "The requisites of equal opportunity". In *Social Choice, Welfare, and Ethics*. Barnett WA, Moulin H, Salles M, and Schofield N (eds). Cambridge University Press: 37-54.
- [9] Fleurbaey, M. (1995c) "Three solutions for the compensation problem". *Journal of Economic Theory* 65: 505-521.
- [10] Fleurbaey, M. (1995d) "Equality and responsibility". *European Economic Review* 39: 683-689.
- [11] Rawls, J. (1971) *A Theory of Justice*. Harvard University Press.

⁸Roemer (1993, 1996) presents one interesting perspective in this respect.

- [12] Roemer, J. (1993) "A pragmatic theory of responsibility for the egalitarian planner". *Philosophy and Public Affairs* 22: 146-166.
- [13] Roemer, J. (1996) *Theories of Distributive Justice*. Harvard University Press.
- [14] Sprumont, Y. (1997) "Balanced egalitarian redistribution of income". *Mathematical Social Sciences* 33: 185-201.
- [15] Tungodden, B. (forthcoming) "Responsibility and redistribution: the case of first best taxation". *Social Choice and Welfare*.