Working Paper No. 65/04

Rewarding effort

by

Alexander W. Cappelen
Bertil Tungodden

SNF Project No. 2515
From circumstance to choice: Implications of the new genetics
for social justice and health policy

The project is financed by the Research Council of Norway

INSTITUTE FOR RESEARCH IN ECONOMICS AND BUSINESS
ADMINISTRATION
BERGEN 2004
ISSN 1503-2140

Rewarding effort

Alexander W. Cappelen*and Bertil Tungodden†

June 24, 2004

Abstract

According to liberal egalitarian ethics, individuals should be rewarded for factors under their control, but not for factors outside their control. A fundamental challenge to liberal egalitarian theories of justice is how to do this without violating minimal egalitarian and liberal requirements. The paper analyses the effects of two such requirements: the principle of equal reward and the principle of reward independence. The exact formulations of these principles depend on how we interpret the concept of reward. We propose two different definitions of reward, contrafactual and interpersonal reward, where both can be given a general and narrow interpretation. Given this, we show that it is impossible to establish a framework that is truly liberal egalitarian in all respects and that a generalized version of the egalitarian equivalent mechanism is the most plausible liberal egalitarian approach.

1 Introduction

Two fundamental ethical questions are what individual characteristics society should reward and how it should reward them. According to an important strand of liberal egalitarian ethics, agents should only be rewarded for factors under their control and not for factors beyond their control (see among others

*The University of Oslo and the Norwegian School of Economics and Business Administration, 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.

Dworkin (1981), Arneson (1989), Cohen (1989), Le Grand (1991), Roemer (1993, 1996, 1998), Bossert (1995), Fleurbaey (1994, 1995a,b,c,d) and Bossert and Fleurbaey (1996)). Let us refer to factors under an agent's control as *effort* and factors outside an agent's control as *talent*. Liberal egalitarian ethics can then be seen as consisting of two parts. First, the liberal principle that agents should be held accountable for the effort they choose to exercise, which implies that agents should be rewarded for their effort. We name this the *principle of responsibility*. Second, the egalitarian principle that the effect of differences in talent should be eliminated, which implies that talent should *not* be rewarded. We name this *the principle of equalization*.

In the context of income distribution, the principle of equalization implies that all individuals exercising the same effort should have the same income (see Bossert (1995) and Bossert and Fleurbaey (1996)). However, the principle of equalization does not tell us anything about how income should be distributed among individuals exercising different levels of effort. It is, for example, consistent with strict (or outcome) egalitarianism, but it also allows for substantial inequalities in income as long as they correspond to differences in effort.

What about the principle of responsibility? One may argue that it justifies that individuals should be rewarded with their marginal productivity, but this interpretation seems (at least) questionable in cases where talent affects marginal productivity (see also Tungodden (forthcoming)). Marginal productivity reward implies that people are not only held accountable for the effort they choose, but also for their talent. Furthermore, it is well-known that this interpretation of the principle of responsibility, in general, is not compatible with the principle of equalization (Bossert (1995), Bossert and Fleurbaey (1996)). Hence, a liberal egalitarian needs to consider alternative interpretations of the principle of responsibility.

The purpose of this paper is to show how we may structure this investigation by introducing some basic liberal egalitarian requirements on how to reward effort. We do this within the framework of a first best economy, which often is seen as a limitation of the analysis. However, this is not the case for the present study. Our aim is to understand the nature of the liberal egalitarian fairness argument for rewarding effort, and for this purpose we want to leave aside incentive considerations. [1] We do this by assuming that people

[1] Similarly, in a discussion of the nature of merit goods, it has turned out to be useful to look at the implications of various definitions in first best economies before moving on

have inelastic effort supply with respect to the design of the redistribution mechanism, which implies that all no-waste allocations of post-tax income will be Pareto optimal (as long as we assume that people have self-interested preferences and a positive marginal utility of income).

Within an egalitarian framework, it may seem trivial to argue that people should be rewarded equally. But this is not the case. The exact nature of this requirement, which we name *the principle of equal reward*, depends on how we define reward. We consider two definitions of reward, *contrafactual reward* and *interpersonal reward*, which both may be given a narrow and a general interpretation. Given a contrafactual definition of reward, it turns out to be impossible to satisfy both versions of the principle of equal reward. This is important to have in mind when considering the liberal egalitarian framework more generally. Even without taking into account incentive considerations, it is impossible to establish a framework that is truly egalitarian in all respects. For the interpersonal interpretations of reward, however, the principle of equal reward is compatible with the principle of equalization, and, moreover, characterizes a group of egalitarian redistribution mechanisms.

Liberal egalitarians only want to reward individuals for factors under their control. Certainly, the choices of others are not within a person's control, and thus another minimal liberal egalitarian requirement should be that the reward scheme is independent of other people's effort. We name this *the principle of reward independence.* The main results of this paper show how different versions of this requirement, together with the principle of equalization, characterize a generalized version of the class of egalitarian equivalent redistribution mechanisms introduced in Bossert and Fleurbaey (1996).

Even though our discussion is placed in the context of income distribution and effort in the labour market, we should like to stress that the present framework is relevant for a much broader set of policy issues. Let us briefly illustrate this by considering such different issues as health policy and interregional redistribution. People make different choices about how to live their lives and these choices affect the health risks they face and their expected need for treatment. The WHO reports that three out of the four top risk factors contributing to the burden of disease could be attributed to unhealthy life style such as unsafe sex, tobacco and alcohol consumption. The idea that individuals must take responsibility for their own health is also an increasingly focused topic in the popular press. A legitimate question is thus how

to second best analysis. See for example Schroyen (forthcoming).

the costs of treatment should be distributed between different individuals and to what extent the distribution of costs should be related to individual behavior. Liberal egalitarians claim that people who make informed and free choices should be held responsible for these choices. However, holding individuals accountable for their choices in the context of health care is extremely controversial. We believe that the main reason for this is that the liberal egalitarian framework is given the wrong interpretation. It is often assumed that responsibility for own health implies that individuals who become sick should pay for their own treatment. But this would imply that those who are unlucky or who are more disposed to become sick are punished for factors beyond their control, which violates the principle of equalization. Hence, it is important to have in mind that liberal egalitarians only attempt to hold individuals accountable for their choices, not for the consequences of their choices. But in the design of health policy, what does it mean to hold people responsible for their choices if they are not to pay the actual costs of treatment? By way of illustration, what does it mean to reward individual effort to reduce the expected need for treatment, e.g., by not smoking? We believe that the present analysis may shed some light on these questions as well, even though it is beyond the scope of this paper to pursue this particular application of liberal egalitarian reasoning.

The question of how the distribution of burdens and benefits should be related to an agent's effort is also at the core of interregional fiscal equalization. Local jurisdictions within the same country often have different capacities to raise revenue and face different costs of providing public goods. This calls for intergovernmental transfers. Fiscal equalization aims at reconciling two important political principles in such situations. First, the principle that differences in fiscal capacity among local jurisdiction should be eliminated, which reflects a concern for interregional inequality being a result of factors beyond the control of the local jurisdictions. Second, the principle that a jurisdiction should be held responsible for decisions that are under their control, in particular their tax effort, which reflects a concern for local autonomy. The fundamental challenge for central governments is how to design a system of intergovernmental transfer satisfying both these two principles, that is, a transfer system that gives all local jurisdictions equal opportunities and at the same time rewards their tax effort. Also in this case, we believe that both the reported results and the broader framework should be of value for the policy debate.

The paper is organized as follows. After presenting the basic framework in section 2, we consider how to define the concept of reward in section 3. In sections 4 and 5, we study the principle of equal reward and the principle of reward independence, and we show that the implications of these principles depend crucially on how reward is defined. In section 6, using the versions of the principle of reward independence compatible with the principle of equalization, we provide several independent characterizations of the generalized egalitarian equivalent mechanism. The final section provides an overview of the analysis and some concluding comments.

2 The basic framework

Consider a society with a population $N = \{1, ..., n\}$, $n \geq 4$. Let $\Omega^E = \{e^1, e^2, ...\}$ be the set of possible effort levels, where $e^{\min}$ is the effort level reflecting that a person does not work, and $\Omega^T = \{t^1, t^2, ...\}$ the set of possible talent levels. $\Omega^E \subseteq \Re$ and $\Omega^T \subseteq \Re$, where $\Re$ is the set of real numbers.[2] Let $a_i = (a_i^E = e, a_i^T = t)$, where $e \in \Omega^E$ and $t \in \Omega^T$, be a characteristics vector of person i and $a = (a_1, ..., a_n)$ be a characteristics profile of society. Define $\Omega_i \subset \Re^2$ as the set of possible characteristics vectors of person i, where for any $i \in N$ and $a_i, \tilde{a}_i \in \Omega_i$, $a_i^T = \tilde{a}_i^T$. In other words, we do not consider interprofile conditions with respect to talent, but assume that there is a unique characteristics profile of talent in society. In order to make the model relevant for our study, though, we assume that there are differences in talent, i.e., that there exist $j, k \in N$ such that $a_j^T \neq a_k^T$. Beyond this, we do not impose any restrictions on the characteristics profile of talents.

Let Ω_i^E be the set of effort levels available for person i, where we assume that $\Omega_i^E = \Omega^E, \forall i \in N$. We also assume that everyone can, at least, choose between working and not working, i.e., $a_i = (e^{\min}, a_i^T), \tilde{a}_i = (\tilde{a}_i^E > e^{\min}, a_i^T) \in \Omega_i$, $\forall i \in N$, but we do not impose any further restrictions on the set of effort levels. Hence, the framework covers both continuous and discrete cases. Define $\mathring{a}$ as the situation where everyone exercises minimum effort, i.e., $\mathring{a}_i^E = e^{\min}, \forall i \in N$. Finally, let $E(a)$ represent the distribution of effort in a, where, for any $e \in \Omega^E$, $E_e(a)$ is the cardinality of the set $\{i \in N \mid a_i^E = e\}$. $\bar{E}$ is the set of all possible effort distributions.

In most of the analysis, we will assume that $\Omega^N = \Omega_1 \times \Omega_2 \times ... \times \Omega_n$ is the set of possible characteristics profiles of society, where $\Omega^N \subseteq \Re^{2n}$. In the

[2]Hence, we do not consider the multidimensional version of this problem; see Bossert (1995) and Bossert and Fleurbaey (1996).

final part, however, we assume *restricted domain richness*, i.e., we demand that every talent group in society is represented on every chosen effort level in a particular situation. Formally speaking, this implies that the set of possible characteristics profiles of society is given by $\bar{\Omega}^N = \{a \in \Omega^N$ *|for any* $j, k \in N$, *where* $a_j^T \neq a_k^T$, *there exists* $l, m \in N$ *such that* $a_l^E = a_j^E$, $a_m^E = a_k^E$, $a_l^T = a_k^T$, *and* $a_m^T = a_j^T\}$. This should be a straightforward assumption to make when studying redistribution in large societies, and nothing of importance seems to be lost by restricting the domain in this way.

The pre-tax income function $f : \Omega \rightarrow \Re$, where $\Omega = \Omega^E \times \Omega^T$, is assumed to be strictly increasing in effort and talent, where $f(e^{\min}, t) = 0$, $\forall t \in \Omega^T$. Moreover, we assume that f is not additively separable, i.e., there exist $e^1, e^2 \in \Omega^E$ and $j, k \in N$ such that $f(e^2, a_j^T) - f(e^1, a_j^T) > f(e^2, a_k^T) - f(e^1, a_k^T)$. We only pay attention to information about effort and talent when choosing allocations, and thus our object of study can be described as a redistribution mechanism F: $\Omega^N \rightarrow \Re^n$ (or, in the final part, $F : \bar{\Omega}^N \rightarrow \Re^n$). We assume that F satisfies the no-waste condition $\sum_{i=1}^n F_i(a) = \sum_{i=1}^n f(a_i)$, $\forall a \in \Omega^N$.

The analysis will take place within the framework of some basic conditions. First, we will only consider anonymous redistribution mechanism (even though we only use this restriction explicitly in some parts of the analysis).

Anonymity (A): For any $a, \tilde{a} \in \Omega^N$ *and* $j, k \in N$, $[a_j^T = a_k^T$, $a_j^E = \tilde{a}_k^E$, $a_k^E = \tilde{a}_j^E$ and $a_i^E = \tilde{a}_i^E$, $\forall i \neq j, k] \rightarrow [F_j(a) = F_k(\tilde{a})$, $F_k(a) = F_j(\tilde{a})$ *and* $F_i(a) = F_i(\tilde{a}\ ), \forall i \neq j, k]$.

More importantly, the underlying aim of the analysis will be to see how we can reward effort within a framework satisfying the core liberal egalitarian principle of equalization, to wit that people exercising the same effort should receive the same post-tax income.

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

An important group of redistribution mechanisms, the class of *egalitarian equivalent mechanisms*, was introduced and characterized by Bossert and Fleurbaey (1996). In the present framework, we may write this class as follows.

$F_k^{EE}(a) = f(t, a_k^E) - \frac{1}{n} \sum_{i \in N} \left[f(t, a_k^E) - f(a_i) \right]$, for some $t \in \Omega^T$ and $\forall k \in N$, $\forall a \in \Omega^N$.

For any given reference talent $t \in \Omega^T$, an egalitarian equivalent mechanism assigns to every individual a post-tax income that consists of two parts. The first part is a transfer equal to the pre-tax income she would have had if her talent were equal to the reference talent. This creates an overall deficit or surplus and the second part is a uniform transfer to all individuals that balances the budget. In the following, we will show that a generalized version of F^{EE} is the most plausible liberal egalitarian approach on how effort should be rewarded.

3 What is a reward?

In order to analyze how effort is to be rewarded, we need a definition of the concept of reward. Intuitively, one may think of reward as the additional you get because you exercise a certain level of effort rather than another. This formulation, however is not sufficiently precise. In order to measure how much a person is rewarded, we need to know *who* to compare with. Two main alternatives are possible. We may either compare with the effort the person herself exercises in another situation or we may compare with the effort another person exercises in the same situation. A reward can thus be defined either through *contrafactual comparisons* or through *interpersonal comparisons.*

3.1 Contrafactual reward

According to the contrafactual definition, a reward is the increase in post-tax income of a person when she increases her effort. The most general version covers all cases where such a change takes place.

Definition 1 *General Contrafactual Reward (GCR). For any* $a, \tilde{a} \in \Omega^N$ *and* $k \in N$*, where* $a_k^E > \tilde{a}_k^E$*, contrafactual reward is defined by* $F_k(a) - F_k(\tilde{a})$*.*

Alternatively, we may narrow our definition of contrafactual reward to situations where only the person in question makes a change in effort.

Definition 2 *Narrow Contrafactual Reward (NCR): For any* $a, \tilde{a} \in \Omega^N$ *and* $k \in N$*, where* $a_k^E > \tilde{a}_k^E$ *and* $\tilde{a}_i^E = a_i^E$*,* $\forall i \neq k$*, contrafactual reward is defined by* $F_k(a) - F_k(\tilde{a})$*.*

We will shortly see how these definitions can be used to impose restrictions on how to reward effort within a liberal egalitarian framework.

3.2 Interpersonal reward

Contrafactual reward has strong intuitive appeal, but still it does not seem to capture fully our understanding of reward. Consider, for example, the strict egalitarian redistribution mechanism giving everyone the same income independent of their effort. Clearly, in one way, this redistribution mechanism cannot be said to reward effort at all. However, according to both versions of contrafactual reward, the person is given a positive reward when increasing her effort.

In contrast, the interpersonal definition views reward as the difference between what two persons get in a given situation. In the most general version, the definition covers all comparisons between persons exercising different levels of effort.

Definition 3 *General Interpersonal Reward (GIR). For any* $a \in \Omega^N$ *and* $j, k \in N$*, where* $a_j^E > a_k^E$*, interpersonal reward is defined by* $F_j(a) - F_k(a)$.

Clearly, if a person's post-tax income partly depends on her talent, then sometimes the general version of interpersonal reward will not only capture differences in post-tax income due to differences in effort.

This problem is avoided if we apply a more narrow definition of interpersonal reward.

Definition 4 *Narrow Interpersonal Reward (NIR). For any* $a \in \Omega^N$ *and* $j, k \in N$*, where* $a_j^T = a_k^T$ *and* $a_j^E > a_k^E$*, interpersonal reward is defined by* $F_j(a) - F_k(a)$.

The narrow definition of interpersonal reward only covers comparisons between people with the same talent, where differences in post-tax income solely can be explained by differences in effort.

We will now put the four definitions into work and study how they can be used to impose restrictions on how to reward effort within a liberal egalitarian framework.

4 Equal reward

The principle of equalization implies that people are rewarded equally in some situations. First, if two persons make the same change in effort when moving from one situation to another, then EIEE implies that they receive the same contrafactual reward. Second, in a given situation, EIEE implies that two persons being at the same effort level receive the same interpersonal reward (independently of whom they are compared with).

However, we propose that the principle of equal reward should be interpreted more broadly within an egalitarian framework, and we now consider how to formulate this more precisely and the implications of imposing the various interpretations of the principle of equal reward on the redistribution mechanism.

4.1 Equal contrafactual reward

First, let us consider the demand for equal general contrafactual reward, to wit, that an increase in post-tax income following a given increase in effort should be independent of the person's talent (but not necessarily independent of what others do). Consider a situation a where two individuals j and k exercise the same effort level. Compare this to two other situations $\tilde{a}, \hat{a}$, where everyone but j and k makes the same change when moving from a to $\tilde{a}$ as from a to $\hat{a}$. In this respect, the move from a to $\tilde{a}$ is identical to the move from a to $\hat{a}$. Moreover, assume that the only difference between $\tilde{a}$ and $\hat{a}$ is that j exercises the same effort level in $\tilde{a}$ as k in $\hat{a}$ and vice versa. In this case, we propose that equal general contrafactual reward implies that the change in post-tax income for j when moving from a to $\tilde{a}$ should be equal to the change in post-tax income for k when moving from a to $\hat{a}$. Formally speaking, this can be stated as follows.

Equal General Contrafactual Reward (EGCR): For any $a, \tilde{a}, \hat{a} \in \Omega^N$ *and* $j, k \in N$, $[a_j^E = a_k^E \neq \tilde{a}_j^E = \hat{a}_k^E,\ \hat{a}_j^E = \tilde{a}_k^E\ \tilde{a}_i^E = \hat{a}_i^E,\ \forall i \neq j, k\ ] \rightarrow [F_j(\tilde{a}) - F_j(a) = F_k(\hat{a}) - F_k(a)]$.[3]

EGCR is a strong egalitarian claim and it turns out that it is not compatible with a redistribution mechanism satisfying no-waste.

Proposition 1. *There does not exist any redistribution mechanism* F *satisfying EGCR.*

Proof. (1) Suppose F satisfies EGCR. f is not additively separable and hence there exist $j, k \in N$ and $e^1, e^2 \in \Omega^E$ such that $f(e^2, a_j^T) - f(e^1, a_j^T) > f(e^2, a_k^T) - f(e^1, a_k^T)$.

[3] Notice that we do not demand that $\tilde{a}$ and $\hat{a}$ are distinct alternatives, i.e. the condition also covers the situation where j and k move together from one effort level to another.

(2) Consider any $a, \tilde{a}, \hat{a}, \bar{a} \in \Omega^N$ where $a_i^E = a_j^E = e^1 \neq \bar{a}_i^E = \bar{a}_j^E = e^2, \forall i \in N$, $\tilde{a}_i^E = \hat{a}_i^E = e^2, \forall i \neq j, k$, $\tilde{a}_j^E = e^2 > \hat{a}_j^E = e^1$, $\tilde{a}_k^E = e^1$ and $\hat{a}_k^E = e^2$. By EGCR, $F_j(\tilde{a}) - F_j(a) = F_k(\hat{a}) - F_k(a)$, $F_i(\tilde{a}) - F_i(a) = F_j(\tilde{a}) - F_j(a)$, $\forall i \neq k$ and $F_i(\hat{a}) - F_i(a) = F_k(\hat{a}) - F_k(a)$, $\forall i \neq j$.

(3) By the no-waste condition, $\sum_{i \in N}(F_i(\tilde{a}) - F_i(a)) = \sum_{i \in N}(f(\tilde{a}_i) - f(a_i))$ and $\sum_{i \in N}(F_i(\hat{a}) - F_i(a)) = \sum_{i \in N}(f(\hat{a}_i) - f(a_i))$. By taking into account (2), it is easily seen that $(n-1)\ (F_j(\tilde{a}) - F_j(a)) + (F_k(\tilde{a}) - F_k(a)) = \sum_{i \in N}(f(\tilde{a}_i) - f(a_i))$ and $(n-1)\ (F_k(\hat{a}) - F_k(a)) + (F_j(\hat{a}) - F_j(a)) = \sum_{i \in N}(f(\hat{a}_i) - f(a_i))$.

(4) By (1), it follows that $\sum_{i \in N}(f(\tilde{a}_i) - f(a_i)) > \sum_{i \in N}(f(\hat{a}_i) - f(a_i))$. Hence, taking into account (3), it follows that $(F_j(\tilde{a}) - F_j(a)) + (F_k(\tilde{a}) - F_k(a)) > (F_k(\hat{a}) - F_k(a)) + (F_j(\hat{a}) - F_j(a))$. By (2), this can be simplified to $F_k(\tilde{a}) - F_k(a) > F_j(\hat{a}) - F_j(a)$.

(5) By ECGR, $F_j(\bar{a}) - F_j(a) = F_k(\bar{a}) - F_j(a)$. But this can be written as $(F_j(\bar{a}) - F_j(\hat{a})) + (F_j(\hat{a}) - F_j(a)) = (F_k(\bar{a}) - F_k(\tilde{a})) + (F_k(\tilde{a}) - F_k(a))$. By ECGR, $F_j(\bar{a}) - F_j(\hat{a}) = F_k(\bar{a}) - F_k(\tilde{a})$, and hence $F_j(\hat{a}) - F_j(a) = F_k(\tilde{a}) - F_k(a)$. But this contradicts (4) and the result follows. ■

Hence, if people differ in talent and talent affects marginal productivity, then we cannot fulfill the demand of equal general contrafactual reward.

There are, however, redistribution mechanisms satisfying the weaker demand of equal narrow contrafactual reward, where the egalitarian claim is restricted to situations where there is a change in effort of only one person.[4]

Equal Narrow Contrafactual Reward (ENCR): For any $a, \tilde{a}, \hat{a} \in \Omega^N$ *and* $j, k \in N$, $[a_j^E = a_k^E < \tilde{a}_j^E = \hat{a}_k^E,\ \tilde{a}_i^E = a_i^E,\ \forall i \neq j$, and $\hat{a}_i^E = a_i^E,\ \forall i \neq k\ ] \rightarrow [F_j(\tilde{a}) - F_j(a) = F_k(\hat{a}) - F_k(a)]$.

It turns out, however, to be impossible to satisfy even this weaker version of the principle of equal reward within a framework satisfying the other core idea of liberal egalitarianism.

Proposition 2. *There does not exist any redistribution mechanism F satisfying ENCR and EIEE.*

Proof. (1) Suppose F satisfies both ENCR and EIEE. f is not additively separable and thus there exist $j, k \in N$ where $a_j^T \neq a_k^T$ and $e^1, e^2 \in \Omega^E$ such

[4]To see this, consider the redistribution mechanism consisting of the following two parts. First, each individual is given a transfer equal to the pre-tax income she would have had if her talent were equal to the most talented and, second, everyone is sharing equally in the surplus or deficit generated by the first part for all other persons (but not for themselves). This redistribution mechanism satisfies ECR, where narrow contrafactual reward equals the marginal productivity of the most talented.

that $(f(e^2, a_j^T) - f(e^1, a_j^T)) \neq (f(e^2, a_k^T) - f(e^1, a_k^T))$.

(2) Consider some $a, \tilde{a}, \hat{a}, \bar{a} \in \Omega^N$, where $a_i^E = e^1, \forall i \in N$, $\tilde{a}_i^E = a_i^E$, $\forall i \neq j$, $\hat{a}_i^E = a_i^E$, $\forall i \neq k$, $\bar{a}_i^E = a_i^E$, $\forall i \neq j, k$, $\tilde{a}_j^E = \hat{a}_k^E = e^2$, and $\bar{a}_j^E = \bar{a}_k^E = e^2$.

(3) By EIEE, $F_j(a) = F_k(a)$. Moreover, by ENCR, $F_j(\tilde{a}) - F_j(a) = F_k(\hat{a}) - F_k(a)$.

(4) By EIEE, $F_j(\bar{a}) = F_k(\bar{a})$. Moreover, by ENCR, $F_j(\bar{a}) - F_j(\hat{a}) = F_k(\bar{a}) - F_k(\tilde{a})$. Hence, $F_j(\hat{a}) = F_k(\tilde{a})$ and, taking into account the first part of (3), $F_j(\hat{a}) - F_j(a) = F_k(\tilde{a}) - F_k(a)$.

(5) By EIEE, $F_i(\tilde{a}) = F_k(\tilde{a})$, $\forall i \neq j$ and $F_i(\hat{a}) = F_j(\hat{a})$, $\forall i \neq k$. Hence, by no-waste, $(n-1)(F_j(\hat{a}) - F_j(a)) + (F_k(\hat{a}) - F_k(a)) = f(\hat{a}_k) - f(a_k)$ and $(n-1)(F_k(\tilde{a}) - F_k(a)) + (F_j(\tilde{a}) - F_j(a)) = f(\tilde{a}_j) - f(a_j)$, i.e., $F_j(\hat{a}) - F_j(a) = \frac{1}{n-1}[(f(\hat{a}_k) - f(a_k)) - (F_k(\hat{a}) - F_k(a))]$ and $F_k(\tilde{a}) - F_k(a) = \frac{1}{n-1}[(f(\tilde{a}_j) - f(a_j)) - (F_j(\tilde{a}) - F_j(a))]$.

(6) By (4) and (5), $[(f(\hat{a}_k) - f(a_k)) - (F_k(\hat{a}) - F_k(a))] = [(f(\tilde{a}_j) - f(a_j)) - (F_j(\tilde{a}) - F_j(a))]$. By (3), this can be simplified to $f(\tilde{a}_j) - f(a_j) = f(\hat{a}_k) - f(a_k)$. But this contradicts (1) and the result follows. ■

Proposition 1 and Proposition 2 make clear that, even without taking into account incentive considerations, it is not possible to establish a redistribution mechanism that is truly egalitarian in all respects.

Proposition 1 shows that this is partly due to the fact that we do not want to waste resources. Proposition 2, however, shows that there is also a fundamental conflict between two basic egalitarian ideals. If we aim at giving people the same contrafactual reward in situations where this is possible, then we cannot at the same always assign equal income to people exercising the same effort.

4.2 Equal interpersonal reward

Given an interpersonal definition of reward, we also have two possible interpretations of the principle of equal reward. The demand of equal general interpersonal reward states that if two situations are equal in all other respects than that the effort levels of two individuals, j and k, are permuted with the effort levels of two other people, l and m, then the post-tax income difference between l and m in the new situation should be the same as the post-tax income difference between j and k in the initial situation. Formally, this can be written as follows.

Equal General Interpersonal Reward (EGIR): For any $a, \tilde{a} \in \Omega^N$ *and* $j, k, l, m \in N$, $[a_j^E = \tilde{a}_l^E$, $\tilde{a}_j^E = a_l^E$, $a_k^E = \tilde{a}_m^E$, $\tilde{a}_k^E = a_m^E$, $a_j^E \neq a_k^E$ *and* $a_i^E = \tilde{a}_i^E$,

$\forall i \neq j, k, l, m\ ] \rightarrow F_j(a) - F_k(a) = F_l(\tilde{a}) - F_m(\tilde{a})$.

EGIR should have strong egalitarian appeal. If we want to reward differences in effort equally, then it is hard to see how to justify a difference in general interpersonal reward in situations covered by EGIR.

Immediately, we can notice that EGIR is equivalent to claiming that general interpersonal reward should be the same in all situations where the overall effort structure is the same (and not only in all situations were everything else is equal).

Lemma 1. *The redistribution mechanism F satisfies EGIR if and only if for any $a, \tilde{a} \in \Omega^N$ and $j, k, l, m \in N$,* where $a_j^E = \tilde{a}_l^E$, $\tilde{a}_j^E = a_l^E$, $a_k^E = \tilde{a}_m^E$, $\tilde{a}_k^E = a_m^E$, $a_j^E \neq a_k^E$ *and* $E(a) = E(\tilde{a})$, $F_j(a) - F_k(a) = F_l(\tilde{a}) - F_m(\tilde{a})$.

Proof. The if part.

(1) It follows from observing that $E(a) = E(\tilde{a})$ in all situations covered by EGIR.

The only-if part

(2) Consider any $a, \tilde{a} \in \Omega^N$ and $j, k, l, m \in N$, where $a_j^E = \tilde{a}_l^E$, $\tilde{a}_j^E = a_l^E$, $a_k^E = \tilde{a}_m^E$, $\tilde{a}_k^E = a_m^E$, $a_j^E \neq a_k^E$, and $E(a) = E(\tilde{a})$. It follows that the cardinality $\hat{N}_C(a, \tilde{a})$ of the set $\hat{N}(a, \tilde{a}) = \{i \neq j, k, l, m \mid a_i^E \neq \tilde{a}_i^E\}$ is $0 \leq \hat{N}_C(a, \tilde{a}) \leq n-4$. If $\hat{N}_C(a, \tilde{a}) = 0$, then it follows from EGIR that $F_j(a) - F_k(a) = F_l(\tilde{a}) - F_m(\tilde{a})$. By the fact that $E(a) = E(\tilde{a})$, $\hat{N}_C(a, \tilde{a}) \neq 1$. We will now prove that for any $\hat{N}_C(a, \tilde{a}) > 1$, there exists $\hat{a} \in \Omega^N$ such that $\hat{N}_C(a, \tilde{a}) - \hat{N}_C(\hat{a}, \tilde{a}) \geq 2$, $a_i^E = \hat{a}_i^E, \forall i \notin \hat{N}(a, \tilde{a})$, $E(a) = E(\hat{a})$ and $F_j(a) - F_k(a) = F_j(\hat{a}) - F_k(\hat{a})$.

(3) If $\hat{N}_C(a, \tilde{a}) > 1$, there exist some $r, s, u, v \neq j, k, l, m$ (where r, s are not necessarily distinct from u, v) such that $a_i^E \neq \tilde{a}_i^E, i = r, s, u, v$ and $a_u^E = \tilde{a}_r^E$ and $a_v^E = \tilde{a}_s^E$. Consider $a^1 \in \Omega^N$, where $a_u^{1E} = a_j^E$, $a_j^{1E} = a_u^E$, $a_v^{1E} = a_k^E$, $a_k^{1E} = a_v^E$, and $a_i^{1E} = a_i^E$, $\forall i \neq j, k, u, v$. By EGIR, $F_j(a) - F_k(a) = F_u(a^1) - F_v(a^1)$.

(4) Consider $a^2 \in \Omega^N$, where $a_u^{2E} = a_r^{1E}$, $a_r^{2E} = a_u^{1E}$, $a_v^{2E} = a_s^{1E}$, $a_s^{2E} = a_v^{1E}$, and $a_i^{2E} = a_i^{1E}$, $\forall i \neq r, s, u, v$. By EGIR, we have that $F_u(a^1) - F_v(a^1) = F_r(a^2) - F_s(a^2)$.

(5) Consider $a^3 \in \Omega^N$, where $a_j^{3E} = a_r^{2E}$, $a_r^{3E} = a_j^{2E}$, $a_k^{3E} = a_s^{2E}$, $a_s^{3E} = a_k^{2E}$, and $a_i^{3E} = a_i^{2E}$, $\forall i \neq j, k, r, s$. By EGIR, $F_r(a^2) - F_s(a^2) = F_j(a^3) - F_k(a^3)$.

(6) By (3) - (5), it follows that $F_j(a) - F_k(a) = F_j(a^3) - F_k(a^3)$, $a_i^E = a_i^{3E}$, $\forall i \neq r, s, u, v$, $a_r^{3E} = \tilde{a}_r^E$ and $a_s^{3E} = \tilde{a}_s^E$. Hence, $\hat{N}_C(a, \tilde{a}) - \hat{N}_C(a^3, \tilde{a}) \geq 2$ and $E(a) = E(a^3)$, and we have established the result promised in (2).

(7) From (6), it follows straightforwardly by induction that $F_j(a) - F_k(a) = F_j(\bar{a}) - F_k(\bar{a})$, where $\hat{N}_C(\bar{a}, \tilde{a}) = 0$ and $a_i^E = \bar{a}_i^E$, $\forall i \notin \hat{N}(a, \tilde{a})$. By EGIR,

$F_j(\bar{a}) - F_k(\bar{a}) = F_l(\tilde{a}) - F_m(\tilde{a})$ and the result follows. ■

Given EIEE, EGIR narrows the class of admissible redistribution mechanisms. To see this, let us first define the function $\tilde{r} : \Omega^E \times \bar{E} \to \Re$.[5] Consider now the following class of egalitarian redistribution mechanisms, which includes the egalitarian equivalent mechanism F^{EE} as a special case.

$F_k^{GE}(a) = \tilde{r}(a_k^E, E(a)) - \frac{1}{n}\sum_{i\in N}\left[\tilde{r}(a_k^E, E(a)) - f(a_i)\right], \forall k \in N, \forall a \in \Omega^N.$

It turns out that we have to adopt F_k^{GE} if we endorse EGIR and EIEE.[6]

Proposition 3. *The redistribution mechanism F satisfies EGIR and EIEE if and only if* $F = F^{GE}$.

Proof. The if part.

(1) Let us show that F^{GE} satisfies EGIR. For any $j, k, l, m \in N$ and $a, \tilde{a} \in \Omega^N$, we have that $F_j^{GE}(a)$ - $F_k^{GE}(a) = \tilde{r}(a_j^E, E(a))$ - $\tilde{r}(a_k^E, E(a))$ and $F_l^{GE}(\tilde{a})$ - $F_m^{GE}(\tilde{a}) = \tilde{r}(\tilde{a}_l^E, E(\tilde{a}))$ - $\tilde{r}(\tilde{a}_m^E, E(\tilde{a}))$. The result follows by observing that $E(a) = E(\tilde{a})$, $a_j^E = \tilde{a}_l^E$ and $a_k^E = \tilde{a}_m^E$ in all cases covered by EGIR. It is easily seen that F^{GE} satisfies EIEE.

The only-if part.

(2) From Lemma 1, it follows that for any two effort levels, a difference in interpersonal reward can only be due to a difference in the overall effort structure in society. From EIEE, it follows that there should never be any difference in post-tax income between two persons exercising the same effort. Hence, it is possible to define some $\tilde{r} : \Omega^E \times \bar{E} \to \Re$ such that for any $a \in \Omega^N$ and any $j \in N$:

$F_j(a) - F_1(a) = \tilde{r}(a_j^E, E(a)) - \tilde{r}(a_1^E, E(a)),$

.

.

.

$F_j(a) - F_n(a) = \tilde{r}(a_j^E, E(a)) - \tilde{r}(a_n^E, E(a)).$

(3) It follows from (2) that $nF_j(a) - \sum_{i\in N} F_i(a) = n\tilde{r}(a_j^E, E(a)) - \sum_{i\in N} \tilde{r}(a_i^E, E(a))$. By the no-waste condition, $\sum_{i\in N} F_i(a) = \sum_{i\in N} f_i(a)$ and hence we have that $F_j(a) = \tilde{r}(a_j^E, E(a)) - \frac{1}{n}\sum_{i\in N}\left[\tilde{r}(a_i^E, E(a) - f(a_i)\right]$. The result follows. ■

F^{EE} implies that interpersonal reward is independent of the overall effort structure. But there are other possibilities within the framework of F^{GE}. To

[5] Let $\tilde{r}(e, E(a)) = 0$ if e is not chosen in a.

[6] The framework of F^{GE} excludes a number of egalitarian redistribution mechanisms satisfying EIEE. By way of illustration, it does not include the proportional egalitarian equivalent mechanism and the subgroup solidarity mechanism discussed in Cappelen and Tungodden (2003).

illustrate, let $\theta_C(a_k^E)$ be the cardinality of the set $\theta(a_k^E) = \left\{i \in N \mid a_i^E > a_k^E\right\}$. Consider now the following rank order dependent class of redistribution mechanisms.

$F_k^{ROE}(a) = \tilde{r}(a_k^E, \theta_C(a_k^E)) - \frac{1}{n}\sum_{i\in N}\left[\tilde{r}(a_k^E, \theta_C(a_k^E)) - f(a_i)\right], \forall k \in N$, $\forall a \in \Omega^N$.

F^{ROE} relies both on absolute and relative performance when rewarding effort. It is easily seen that it satisfies EIEE and EGIR. However, in the next section, we will argue that it violates a fundamental liberal egalitarian intuition, and for that reason is not a plausible liberal egalitarian redistribution mechanism.

Finally, let us consider the implications of imposing the weaker version of equal interpersonal reward on the redistribution mechanism, where we only consider interpersonal reward between equally talented individuals.

Equal Narrow Interpersonal Reward (ENIR): For any $a, \tilde{a} \in \Omega^N$ *and* $j, k, l, m \in N$, *where* $a_j^T = a_k^T = a_l^T = a_m^T$, $[a_j^E = \tilde{a}_l^E$, $\tilde{a}_j^E = a_l^E$, $a_k^E = \tilde{a}_m^E$, $\tilde{a}_k^E = a_m^E$, $a_j^E \neq a_k^E$ *and* $a_i^E = \tilde{a}_i^E$, $\forall i \neq j, k, l, m\,] \rightarrow F_j(a) - F_k(a) = F_l(\tilde{a}) - F_m(\tilde{a})$.

It turns out that this is a very weak requirement that is satisfied by all anonymous redistribution mechanisms.

Proposition 4. *If the redistribution mechanism F satisfies* A, *then it satisfies ENIR.*

Proof. (1) Consider any $a, \tilde{a} \in \Omega^N$ and $j, k, l, m \in N$, where $a_j^T = a_k^T = a_l^T = a_m^T$ and $a_j^E = \tilde{a}_l^E$, $\tilde{a}_j^E = a_l^E$, $a_k^E = \tilde{a}_m^E$, $\tilde{a}_k^E = a_m^E$, $a_j^E \neq a_k^E$ *and* $a_i^E = \tilde{a}_i^E$, $\forall i \neq j, k, l, m$.

(2) Consider $\hat{a}$, where $\hat{a}_j^E = a_l^E$, $\hat{a}_l^E = a_j^E$ and $a_i^E = \hat{a}_i^E, \forall i \neq j, l$. By A, $F_l(\hat{a}) = F_j(a)$, $F_j(\hat{a}) = F_l(a)$ and $F_i(\hat{a}) = F_i(a), \forall i \neq j, l$.

(3) By A, $F_k(\tilde{a}) = F_m(\hat{a})$, $F_m(\tilde{a}) = F_k(\hat{a})$ and $F_i(\bar{a}) = F_i(\hat{a}), \forall i \neq k, m$.

(4) By (2) - (3), it follows that $F_j(a) = F_l(\tilde{a})$ and $F_k(a) = F_m(\tilde{a})$, i.e. $F_j(a) - F_k(a) = F_l(\tilde{a}) - F_m(\tilde{a})$. ■

4.3 Overview on equal reward

In summary, the analysis in this section shows that the contrafactual interpretation of equal reward is too demanding. It is in general impossible to satisfy general contrafactual reward and, moreover, the weaker version is not consistent with a framework satisfying EIEE. Equal interpersonal reward, on the other hand, turns out to be compatible with the principle of equalization. The weaker version is satisfied by any anonymous redistribution mechanism, whereas the stronger version (together with EIEE) characterizes a broad class of redistribution mechanisms.

5 Reward Independence

We will argue that liberal egalitarians should not be concerned with relative performance when rewarding effort (as is case for the rank order dependent class of redistribution mechanisms F^{ROE}). Liberal egalitarians aim at rewarding what is within the control of individuals, and hence should value what we refer to as reward independence. In the case of interpersonal reward, reward independence says that the reward assigned to a person increasing her effort should be independent of what others do. It is the choice of this person a liberal egalitarian wants to reward and not the possible fact that, say, this person ends up being the one exercising most effort in society. The relative position of an individual is beyond her control and thus should not affect the structure of the reward scheme. Similarly, in the case of interpersonal reward, liberal egalitarians should be concerned with the difference in choice between two individuals and not their relative positions (which are beyond their control).

Notice that the principle of reward independence has no egalitarian content. All versions are compatible with pure libertarianism, where no redistribution takes place. Hence, it is an independent requirement of a very different kind than the principle of equalization and the principle of equal reward. But it follows directly from the liberal egalitarian idea of only keeping people responsible for factors within their control. We now consider how to formulate these conditions more precisely and the implications of imposing them on the redistribution mechanism.

5.1 Contrafactual reward independence

The requirement that general contrafactual reward should be independent of the effort of other individuals can be stated formally as follows.

General Contrafactual Reward Independence (GCRI): For any $a, \tilde{a}, \hat{a}, \bar{a} \in \Omega^N$ *and* $k \in N$, $[a_k^E = \tilde{a}_k^E \neq \hat{a}_k^E = \bar{a}_k^E,] \rightarrow [F_k(a) - F_k(\hat{a}) = F_k(\tilde{a}) - F_k(\bar{a})]$.[7]

[7]Given the no-waste condition, *GCRI* is equivalent to the individual monotonicity condition introduced by Bossert (1995) and Bossert and Fleurbaey (1996).

This is clearly a very demanding condition, and not surprisingly it has strong implications for the kind of redistribution mechanism we may adopt.

Proposition 5. *A redistribution mechanism F satisfies GCRI if and only if for any* $a, \tilde{a} \in \Omega^N$ *and person* $k \in N$, $F_k(a) - F_k(\tilde{a}) = f(a_k) - f(\tilde{a}_k)$.

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

(1) Suppose there exist some $a, \tilde{a} \in \Omega^N$ and $k \in N$, where $F_k(a) - F_k(\tilde{a}) \neq f(a_k) - f(\tilde{a}_k)$.

(2) If $a_k^E = \tilde{a}_k^E$, then the supposition in (1) implies that $F_k(a) \neq F_k(\tilde{a})$. By domain richness, there exists $\hat{a} \in \Omega^N$, where $\hat{a}_k^E \neq a_k^E$. By GCRI, $F_k(a) - F_k(\hat{a}) = F_k(\tilde{a}) - F_k(\hat{a})$. But this implies that $F_k(a) = F_k(\tilde{a})$, and hence the supposition in (1) is not possible if $a_k^E = \tilde{a}_k^E$.

(3) Assume that $a_k^E \neq \tilde{a}_k^E$. By domain richness, there exists some $\bar{a} \in \Omega^N$, where $\bar{a}_k^E = a_k^E$ and $\bar{a}_i^E = \tilde{a}_i^E$, $\forall i \neq k$. By (2), $F_i(\bar{a}) = F_i(\tilde{a}), \forall i \neq k$. Hence, by the no-waste condition, $F_k(\bar{a}) - F_k(\tilde{a}) = f(\bar{a}_k) - f(\tilde{a}_k) = f(a_k) - f(\tilde{a}_k)$. But by GCRI, $F_k(\bar{a}) - F_k(\tilde{a}) = F_k(a) - F_k(\tilde{a}) \neq f(a_k) - f(\tilde{a}_k)$, and hence the supposition in (1) is not possible. The result follows. ■

The underlying intuition is straightforward. If some person, say k, gets more (less) than her marginal productivity when increasing effort, then there will be created a deficit (surplus) that, given the no-waste condition, must be distributed among the others. However, it is easily seen that this implies that some other people's general contrafactual reward will depend on what k does, which violates GCRI.

Proposition 5 shows that there is a close link between GCRI and the libertarian redistribution mechanism.

$F_k^L(a) = f(a_k), \forall k \in N, \forall a \in \Omega^N$.

It turns out that we can characterize F^L by combining GCRI with the following very plausible condition.

No Redistribution for Equal Effort and Equal Pre-tax Income (NREEP): For any $a \in \Omega^N, [\, a_1^E = a_2^E = ... = a_n^E$ *and* $f(a_1) = f(a_2) = ... = f(a_n)\,] \rightarrow [\, F_1(a) = F_2(a) = ... = F_n(a)\,]$.

Corollary 1. *A redistribution mechanism F satisfies GCRI and NREEP if and only if* $F = F^L$.

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

(1) Consider any $a \in \Omega^N$ and $k \in N$. We will now show that given the assumptions of the corollary, $F_k(a) = f(a_k)$. By domain richness, there exists $\mathring{a} \in \Omega^N$, where $\mathring{a}_1^E = \mathring{a}_2^E = ... = \mathring{a}_n^E = e^{\min}$ and $f(\mathring{a}_1) = f(\mathring{a}_2) = ... = f(\mathring{a}_n) = 0$. By NREEP and the no-waste condition, $F_1(\mathring{a}) = F_2(\mathring{a}) = ... = F_n(\mathring{a}) = 0$. Hence, $F_k(\mathring{a}) = f(\mathring{a}_k)$.

(2) By Proposition 5, $F_k(a) - F_k(\mathring{a}) = f(a_k) - f(\mathring{a}_k)$. By (1), this can be simplified to $F_k(a) = f(a_k)$. ■

GCRI is too demanding if we want to make redistribution conditional on the effort people exercise. But within such a framework, we will also argue that it is not suitable. If our aim is to redistribute on the basis of effort, then general contrafactual reward captures too much. In many cases, it will reflect both a reward part and a redistribution part, and there is, of course, no reason to demand that the redistribution part should be independent of the effort of others.

In order to isolate the reward part, we consider the following weaker version of contrafactual reward independence.

Narrow Contrafactual Reward Independence (NCRI): For any $a, \tilde{a}, \hat{a}, \bar{a} \in \Omega^N$ *and* $k \in N$, $[a_k^E = \tilde{a}_k^E,\ \hat{a}_k^E = \bar{a}_k^E,\ a_k^E \neq \hat{a}_k^E,\ a_i^E = \hat{a}_i^E$ and $\tilde{a}_i^E = \bar{a}_i^E$, $\forall i \neq k \in N] \rightarrow [F_k(a) - F_k(\hat{a}) = F_k(\tilde{a}) - F_k(\bar{a})]$.

NCRI covers situations where only a single person changes her effort. In such cases, any change in post-tax income of this person must reflect a contrafactual reward of effort, and we suggest that a liberal egalitarian should demand that this contrafactual reward is independent of the effort chosen by others. This requirement is consistent with a wide range of redistribution mechanisms, including both the libertarian mechanism F^L and the egalitarian equivalent mechanism F^{EE}. As we will return to in section 6, however, within the framework of EIEE, it characterizes uniquely a generalized version of the egalitarian equivalent mechanism.

5.2 Interpersonal reward independence

The demand for general interpersonal reward independence can be written as follows.

General Interpersonal Reward Independence (GIRI): For any $a, \tilde{a} \in \Omega^N$ *and* $j,\ k \in N$, where $a_j^E \neq a_k^E$, $\left[a_i^E = \tilde{a}_i^E, i = j, k\right] \rightarrow [F_j(a) - F_k(a) = F_j(\tilde{a}) - F_k(\tilde{a})]$.

GIRI states that the general interpersonal reward assigned to someone exercising more effort than another person should be independent of their relative effort performance in society. This captures precisely the principle of reward independence if a person's post-tax income is independent of her talent. If not, then we need to consider the following weaker version.

Narrow Interpersonal Reward Independence (IRI): For any $a, \tilde{a} \in \Omega^N$ *and* $j, k \in N$, where $a_j^T = a_k^T$ and $a_j^E \neq a_k^E$, $[a_i^E = \tilde{a}_i^E, i = j, k] \rightarrow [F_j(a) - F_k(a) =$

$F_j(\tilde{a}) - F_k(\tilde{a})]$.

NIRI only considers interpersonal reward between persons of the same talent, where the difference in post-tax income only can be explained by differences in effort. It is satisfied by a wide range of redistribution mechanisms, including both the libertarian mechanism F^L and the egalitarian equivalent mechanism F^{EE}. But as we will show in section 6, if we accept a certain restriction on the domain of the redistribution mechanism and combines it with EIEE, we have a unique characterization of a generalized version of the egalitarian equivalent mechanism. If we do not accept restricting the domain, we need GIRI in order to characterize the same class of redistribution mechanisms.

5.3 Overview on reward independence

In summary, most versions of the requirement of reward independence are compatible with the principle of equalization. Only by applying a general definition of contrafactual reward, do we get a conflict with EIEE. In this particular case, there is a close link between the libertarian redistribution mechanism and the requirement of reward independence.

6 The Generalized Egalitarian Equivalent Mechanism

We will now show that each of the versions of reward independence consistent with EIEE supports an independent characterization of a generalized version of the egalitarian equivalent redistribution mechanism.[8]

$F_k^{GEE}(a) = r(a_k^E) - \frac{1}{n}\sum_{i \in N}\left[r(a_k^E) - f(a_i)\right], \forall k \in N, \forall a \in \Omega^N,$

where $r : \Omega^E \to \Re$. The generalized egalitarian equivalent redistribution mechanism consists of two parts. First it rewards all individuals according to a given reward scheme, $r(a_k^E)$; secondly, it distributes the surplus or deficit generated by the first part equally among them.

It turns out that there is a close link between the principle of equalization, the principle of reward independence and F^{GEE}. This is most easily seen in

[8]Notice that this is also a generalization of the formulation of the egalitarian equivalent mechanism suggested by Bossert and Fleurbaey (1996), because we do not demand *any* link between $r(a_k^E)$ and the pre-tax income function f. This implies that the characterization offered in Bossert and Fleurbaey (1996) does not cover the generalized version.

the case where we impose general interpersonal reward independence on the redistribution mechanism.

Proposition 6. *A redistribution mechanism F satisfies EIEE and GIRI if and only if $F = F^{GEE}$.*

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

(1) Consider any $a, \tilde{a} \in \Omega^N$ and $j, k, l, m \in N$, where $a_j^E = \tilde{a}_l^E \neq a_k^E = \tilde{a}_m^E$. We will now prove that $F_j(a) - F_k(a) = F_l(\tilde{a}) - F_m(\tilde{a})$.

(2) Consider $\hat{a} \in \Omega^N$, where $\hat{a}_k^E = a_k^E$, $\hat{a}_j^E = a_j^E$ and $\hat{a}_i^E = \tilde{a}_i^E$, $\forall i \neq j, k$. By GIRI $F_l(\hat{a}) - F_m(\hat{a}) = F_l(\tilde{a}) - F_m(\tilde{a})$, and by EIEE $F_j(\hat{a}) = F_l(\hat{a})$ and $F_k(\hat{a}) = F_m(\hat{a})$. Hence $F_j(\hat{a}) - F_k(\hat{a}) = F_l(\hat{a}) - F_m(\hat{a})$. By GIRI, $F_j(a) - F_k(a) = F_j(\hat{a}) - F_k(\hat{a})$, and we have established the result announced in (1).

(3) From (2), it follows that for any two effort levels, there exists a given difference between the post-tax income of persons exercising these two effort levels. From EIEE, it follows that there should never be any difference in post-tax income between two persons exercising the same effort. Hence it is possible to define a function $r : \Omega^E \to \Re$ such that for any $a \in \Omega^N$ and any $j \in N$:

$F_j(a) - F_1(a) = r(a_j^E) - r(a_1^E)$

.

.

.

$F_j(a) - F_n(a) = r(a_j^E) - r(a_n^E)$

(4) It follows from (3) that $nF_j(a) - \sum_{i \in N} F_i(a) = nr(a_j^E) - \sum_{i \in N} r(a_i^E)$. By the no-waste condition, $\sum_{i \in N} F_i(a) = \sum_{i \in N} f_i(a)$, and hence we have that $F_j(a) = r(a_j^E) - \frac{1}{n} \sum_{i \in N} \left[r(a_i^E) - f_i(a) \right]$. The result follows. ■

In order to see the link between narrow contrafactual reward independence and EIEE, let us first establish the following lemma.

Lemma 2. *If a redistribution mechanism F satisfies EIEE and NCRI, then it satisfies GIRI.*

Proof. (1) We will first establish by induction that for any $a, \tilde{a} \in \Omega^N$ and $j, k \in N$, where $a_i^E = \tilde{a}_i^E, \forall i \neq j$, $a_j^E \neq \tilde{a}_j^E$, $F_i(a) - F_i(\tilde{a}) = F_k(a) - F_k(\tilde{a})$, $\forall i \neq j$.

(2) Consider first any $\hat{a}, \bar{a} \in \Omega^N$, where $\hat{a}_j^E = a_j^E$, $\bar{a}_j^E = \tilde{a}_j^E$, and $\hat{a}_i^E = \bar{a}_i^E = a_k^E$, $\forall i \neq j$. By EIEE, it follows that $F_i(\hat{a}) - F_i(\bar{a}) = F_k(\hat{a}) - F_k(\bar{a})$ $\forall i \neq j$.

(3) Suppose that for some $2 \leq \theta \leq n-1$ and every $a^1, a^2 \in \Omega^N$, where $a_j^{1E} = a_j^E$, $a_j^{2E} = \tilde{a}_j^E$, $a_k^{1E} = a_k^E$, $a_i^{1E} = a_i^{2E}$, $\forall i \neq j$, and the cardinality of the set $M(a^1, a^2) = \left\{ i \neq j \in N \mid a_i^{1E} = a_k^{1E} \right\}$ is $M_C(a^1, a^2) = \theta$, we have that $F_i(a^1) - F_i(a^2) = F_k(a^1) - F_k(a^2) = F_k(\hat{a}) - F_k(\bar{a})$, $\forall i \neq j$. Consider now any $a^3, a^4 \in \Omega^N$ where $a_j^{3E} = a_j^E$, $a_j^{4E} = \tilde{a}_j^E$, $a_k^{3E} = a_k^E$, $a_i^{3E} = a_i^{4E}$, $\forall i \neq j$, and the cardinality of the set $M(a^3, a^4) = \left\{ i \neq j \in N \mid a_i^{3E} = a_k^{3E} \right\}$ is $M_C(a^3, a^4) = \theta - 1$. We will now show, in (4)-(6), that given the supposition in the first part of this paragraph, $F_i(a^3) - F_i(a^4) = F_k(a^3) - F_k(a^4) = F_k(\hat{a}) - F_k(\bar{a})$, $\forall i \neq j$.

(4) Consider any $l \neq j, k \notin M((a^3, a^4)$ and $a^5, a^6 \in \Omega^N$, where $a_i^5 = a_i^3$ and $a_i^6 = a_i^4$, $\forall i \neq l$ and $a_l^{5E} = a_l^{6E} = a_k^E$. It follows that the cardinality of the set $M(a^5, a^6)$ is $M_C(a^5, a^6) = \theta$ and hence from the supposition in (3) that $F_l(a^5) - F_l(a^6) = F_k(\hat{a}) - F_k(\bar{a})$. By NCRI, $F_l(a^3) - F_l(a^5) = F_l(a^4) - F_l(a^6)$, i.e., $F_l(a^3) - F_l(a^4) = F_l(a^5) - F_l(a^6) = F_k(\hat{a}) - F_k(\bar{a})$.

(5) By (4), $F_i(a^3) - F_i(a^4) = F_k(\hat{a}) - F_k(\bar{a}), \forall i \neq j \notin M(a^3, a^4)$. By EIEE, $F_i(a^3) - F_i(a^4) = F_k(a^3) - F_k(a^4)$, $\forall i \in M(a^3, a^4)$. By NCRI, $F_j(a^3) - F_j(a^4) = F_j(\hat{a}) - F_j(\bar{a})$.

(6) By the no-waste condition, $\sum_{i \in N}(F_i(a^3) - F_i(a^4)) = \sum_{i \in N}(f(a^3) - f(a^4))$ and $\sum_{i \in N}(F_i(\hat{a}) - F_i(\bar{a})) = \sum_{i \in N}(f(\hat{a}) - f(\bar{a}))$. By the definition of $\hat{a}, \bar{a}$ in (2) and a^3, a^4 in (3), $\sum_{i \in N}(f(a^3) - f(a^4)) = \sum_{i \in N}(f(\hat{a}) - f(\bar{a}))$, and thus $\sum_{i \in N}(F_i(a^3) - F_i(a^4)) = \sum_{i \in N}(F_i(\hat{a}) - F_i(\bar{a}))$. By taking into account (5), this implies that $F_i(a^3) - F_i(a^4) = F_k(\hat{a}) - F_k(\bar{a})$, $\forall i \in M(a^3, a^4)$. In sum, in (4) - (6), we have established that $F_i(a^3) - F_i(a^4) = F_k(a^3) - F_k(a^4) = F_k(\hat{a}) - F_k(\bar{a})$, $\forall i \neq j$, as announced in (3).

(7) In (2), $M_C(\hat{a}, \bar{a}) = n - 1$. Hence, taking together (2) - (6), it follows, by induction, that for any $a^r, a^s \in \Omega^N$, where $a_j^{rE} = a_j^E$, $a_j^{sE} = \tilde{a}_j^E$, $a_k^{rE} = a_k^E$, $a_i^{rE} = a_i^{sE}$, $\forall i \neq j$, we have that $F_i(a^r) - F_i(a^s) = F_k(a^r) - F_k(a^s)$, $\forall i \neq j$. The result announced in (1) follows by observing that we may define $a^r = a$ and $a^s = \tilde{a}$.

(8) Given (7), we can establish that the redistribution mechanism satisfies GIRI, to wit that for any $a, \tilde{a} \in \Omega^N$ and j, $k \in N$, where $a_i^E = \tilde{a}_i^E, i = j, k$ and $a_j^E \neq a_k^E$, $F_j(\tilde{a}\ ) - F_j(a) = F_k(\tilde{a}) - F_k(a)$. Consider the sequence of alternatives $a^t \in \Omega^N$, where $t = 0...n$, $a^0 = a$, $a_t^t = \tilde{a}_t$, and $a_i^t = a_i^{t-1}$, $\forall i \neq t$. If $a^t \neq a^{t-1}$, then it follows from (7) that $F_j(a^t) - F_j(a^{t-1}) = F_k(a^t) - F_k(a^{t-1})$ for every $t = 1, ..., n$. The result follows by observing that $a^n = \tilde{a}$ and $F_i(\tilde{a}) - F_i(a) = \sum_{t=1,...,n}(F_i(a^t) - F_i(a^{t-1}))$, $\forall i \in N$, which implies that $F_j(\tilde{a}\ ) - F_j(a) = F_k(\tilde{a}) - F_k(a)$. ■

We are now ready to report the following proposition.

Proposition 7. *A redistribution mechanism F satisfies EIEE and NCRI if and only if $F = F^{GEE}$.*

Proof. The result follows from combining Lemma 2 and Proposition 6. ■

Finally, we may even weaken our demand for interpersonal reward independence and still establish a characterization of F^{GEE}. In this case, however, we have to assume restricted domain richness and apply the anonymity condition A.

Proposition 8: *Given restricted domain richness, a redistribution mechanism F satisfies EIEE, NIRI and A if and only if $F = F^{GEE}$.*

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

(1) Consider any $a, \tilde{a} \in \bar{\Omega}^N$ and $j,\, k \in \left\{i \in N \mid\; a_i^E = \tilde{a}_i^E\right\}$, where $a_j^E \neq a_k^E$. We will now show that F satisfies GIRI, i.e., that $F_j(a) - F_k(a) = F_j(\tilde{a}) - F_k(\tilde{a})$.

(2) By restricted domain richness, there exist $l, m \neq j, k \in N$ such that $a_l^E = \tilde{a}_m^E = a_j^E$ and $a_l^T = a_m^T = a_k^T$, and $\hat{a} \in \bar{\Omega}^N$, where $\hat{a}_i^E = \tilde{a}_i^E$, $\forall i \neq l, m$ and $\hat{a}_m^E = \tilde{a}_l^E$ and $\hat{a}_l^E = \tilde{a}_m^E$. By A, $F_i(\hat{a}) = F_i(\tilde{a}), \forall i \neq l, m,$ $F_m(\hat{a}) = F_l(\tilde{a})$ and $F_l(\hat{a}) = F_m(\tilde{a})$.

(3) By (2), $F_j(a) - F_j(\tilde{a}) = F_j(a) - F_j(\hat{a})$. By EIEE, $F_j(a) = F_l(a)$ and $F_j(\hat{a}) = F_l(\hat{a})$, and hence $F_j(a) - F_j(\hat{a}) = F_l(a) - F_l(\hat{a})$. By NIRI, $F_l(a) - F_l(\hat{a}) = F_k(a) - F_k(\hat{a})$. By (2), $F_k(a) - F_k(\hat{a}) = F_k(a) - F_k(\tilde{a})$. Hence, taken together, $F_j(a) - F_j(\tilde{a}) = F_k(a) - F_k(\tilde{a})$, i.e., $F_j(a) - F_k(a) = F_j(\tilde{a}) - F_k(\tilde{a})$ and we have established the result announced in (1).

(4) Even though we have now established that F satisfies GIRI, Proposition 6 does not apply immediately. Step (2) in the proof of Proposition 6 is not valid given restricted domain richness. Hence, we will have to show independently that for any $a, \tilde{a} \in \bar{\Omega}^N$ and $j, k, l, m \in N$, where $a_j^E = \tilde{a}_l^E \neq a_k^E = \tilde{a}_m^E$, $F_j(a) - F_k(a) = F_l(\tilde{a}) - F_m(\tilde{a})$.

(5) By restricted domain richness, we know that there exist $r, s, u, v \in \bar{\Omega}^N$, where $a_r^T = a_s^T = a_u^T = a_v^T$ and $a_r^E = a_j^E$, $a_s^E = a_k^E$, $\tilde{a}_u^E = \tilde{a}_l^E$, and $\tilde{a}_v^E = \tilde{a}_m^E$.[9] Consider $\hat{a} \in \bar{\Omega}^N$, where $\hat{a}_r^E = \tilde{a}_u^E$, $\hat{a}_u^E = \tilde{a}_r^E$, and $\hat{a}_i^E = \tilde{a}_i^E$, $\forall i \neq r, u$. By A, $F_r(\hat{a}) = F_u(\tilde{a})$, $F_u(\hat{a}) = F_r(\tilde{a})$, and $F_i(\hat{a}) = F_i(\tilde{a})$, $\forall i \neq r, u$.

(6) Consider $\bar{a} \in \bar{\Omega}^N$, where $\bar{a}_s^E = \hat{a}_v^E$, $\bar{a}_v^E = \hat{a}_s^E$, and $\bar{a}_i^E = \hat{a}_i^E$, $\forall i \neq s, v$.

[9] Notice that r, j, u and s, m, v do not have to be distinct persons. Hence, as in the other proofs, we only need $n \geq 4$.

By A, $F_s(\bar{a}) = F_v(\hat{a})$, $F_v(\bar{a}) = F_s(\hat{a})$, and $F_i(\bar{a}) = F_i(\hat{a})$, $\forall i \neq s, v$. Hence, we have that $F_r(\bar{a}) - F_s(\bar{a}) = F_r(\hat{a}) - F_v(\hat{a}) = F_u(\tilde{a}) - F_v(\tilde{a})$. By EIEE, $F_u(\tilde{a}) - F_v(\tilde{a}) = F_l(\tilde{a}) - F_m(\tilde{a})$, i.e., $F_l(\tilde{a}) - F_m(\tilde{a}) = F_r(\bar{a}) - F_s(\bar{a})$.

(6) By GIRI, it follows that $F_r(a) - F_s(a) = F_r(\bar{a}) - F_s(\bar{a})$. By EIEE, $F_r(a) - F_s(a) = F_j(a) - F_k(a)$, i.e., $F_j(a) - F_k(a) = F_r(\bar{a}) - F_s(\bar{a})$. Hence, taking into account (6), $F_j(a) - F_k(a) = F_l(\tilde{a}) - F_m(\tilde{a})$, and we have established the result announced in (4).

(7) The proposition follows from applying (6) in steps (3) and (4) of the proof of Proposition 6, which also are valid given restricted domain richness. ■

To see that the propositions provide independent characterizations of F^{GEE}, let us consider the relationship between NCR, GIRI, and NIRI. Clearly, GIRI implies NIRI. But we have to restrict the domain in Proposition 8, and hence it follows that Proposition 8 is distinct from both Proposition 6 and Proposition 7.

What about the relationship between NCRI and the two interpersonal reward independence conditions? Consider the redistribution mechanism consisting of the following two parts. First, everyone is given a transfer equal to the pre-tax income they would have had if they were the most talented and, second, the overall deficit from the first part is divided equally among the most talented in society. This mechanism satisfies NCRI (where contrafactual reward always is equal to the marginal productivity of the most talented), but it does not satisfy GIRI (because people do not share equally in the deficit created when people not in possession of the highest level of talent change their effort). If we impose the further assumption that we only distribute the deficit among those who exercise the highest level of effort within the most talented group, then it will not even satisfy NIRI.

On the other hand, consider the redistribution mechanism where the first part consists of giving a person half of her pre-tax income if there is someone else exercising more effort but all if she is the one exercising most effort in society, and the second part consists of dividing the overall surplus from the first part equally among everyone in the economy. This redistribution mechanism clearly violates NCRI (because contrafactual reward depends on whether you are the one exercising most effort or not), but it satisfies GIRI and NIRI (because everyone shares equally in any surplus or deficit created by a change in someone else's effort). Certainly, none of these redistribution mechanisms satisfy EIEE, and it is only within the framework of this basic egalitarian condition that they all single out the generalized egalitarian equivalent redistribution mechanism.

7 Conclusion

Liberal egalitarians want to reward people for factors under their control. But how can this be done? We have studied the implications of two general liberal egalitarian ideas, to wit that people should be rewarded equally and that the reward structure should be independent of other people's effort. The exact nature of these principles, however, depends on how we define the concept of reward. We have proposed the distinction between contrafactual and interpersonal reward, and within each of the categories we have outlined a general and a narrow definition. In sum, as shown in Table 1, this has provided us with eight different reward conditions.

Reward conditions	Equal Reward	Reward Independence
General contrafactual reward	EGCR	GCRI
Narrow contrafactual reward	NGCR	NCRI
General interpersonal reward	EGIR	GIRI
Narrow interpersonal reward	NGIR	NIRI

Table 1. Reward conditions.

In the main part of this paper, we have studied the implications of imposing these reward conditions on the redistribution mechanism. Table 2 gives an overview of the results.

It turns out that it is in general not possible to satisfy the principle of equal contrafactual reward. There is no redistribution mechanism satisfying the general version, while the narrow version is in conflict with the principle of equalization. This shows that, even without taking into account incentive considerations, it is not possible to establish a redistribution mechanism that is truly egalitarian in all respects.

The equal interpersonal reward requirement is less demanding. Equal narrow interpersonal reward is satisfied by any anonymous redistribution mechanism, whereas equal general interpersonal reward and the principle of equalization characterize a broad group of egalitarian redistribution mechanism.

Condition	Implications
EGCR	Not compatible with any redistribution mechanism satisfying no waste.
ENCR	Not compatible with EIEE.
EGIR	Characterises a broad class of egalitarian redistribution mechanisms (together with EIEE).
NGIR	Satisfied by any anonymous redistribution mechanism
GCRI	Characterizes the libertarian redistribution mechanism (together with NREEP).
NCRI	Characterises the generalized egalitarian equivalent mechanism (together with EIEE).
GIRI	Characterises the generalized egalitarian equivalent mechanism (together with EIEE).
NIRI	Characterises the generalized egalitarian equivalent mechanism (together with EIEE, within a restricted domain).

Table 2. Implications of the reward conditions.

The principle of reward independence follows from the liberal egalitarian idea of only holding people responsible for factors under their control. Interestingly, it turns out that all versions of this principle compatible with the principle of equalization, support a characterization of the same class of redistribution mechanisms, to wit, the generalized egalitarian equivalent redistribution mechanism. This class of redistribution mechanisms also satisfies all versions of the principle of equal reward compatible with the principle of equalization, and thus we conclude that it constitutes the most promising approach to liberal egalitarian reasoning.

References

[1] Arneson, R. (1989) "Equality and equal opportunity for welfare". *Philosophical Studies 56:* 159-194.

[2] Bossert, W. (1995) "Redistribution mechanisms based on individual characteristics". *Mathematical Social Sciences* 29: 1-17.

[3] Bossert, W. and M. Fleurbaey (1996) "Redistribution and compensation". *Social Choice and Welfare 13:* 343-355.

[4] Cappelen, A. and B. Tungodden (2003) "Reward and Responsibility: How should we be affected when others change their effort". *Politics, Philosophy and Economics, 2003, 2* (2): 191-211.

[5] Cohen, G. A. (1989) "On the currency of egalitarian justice". *Ethics* 99: 906-944.

[6] Dworkin, R. (1981) "What is equality? Part 2: equality of resources". *Philosophy and Public Affairs 10:* 283-345.

[7] Fleurbaey, M. (1994) "On fair compensation". *Theory and Decision* 36: 277-307.

[8] Fleurbaey, M. (1995a) "Equal opportunity or equal social outcome". *Economics and Philosophy* 11: 25-55.

[9] 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.

[10] Fleurbaey, M. (1995c) "Three solutions for the compensation problem". *Journal of Economic Theory 65:* 505-521.

[11] Fleurbaey, M. (1995d) "Equality and responsibility". *European Economic Review* 39: 683-689.

[12] Fleurbaey, M. and F. Maniquet (1999) "Cooperative production with unequal skills: The solidarity approach to compensation". *Social Choice and Welfare 16: 569-583.*

[13] Le Grand, J. (1991) *Equity and Choice.* Harper Collins.

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

[15] Roemer, J. (1993) "A pragmatic theory of responsibility for the egalitarian planner". *Philosophy and Public Affairs* 22: 146-166.

[16] Roemer, J. (1996) *Theories of Distributive Justice.* Harvard University Press.

[17] Roemer, J. E. (1998) *Equality of Opportunity.* Harvard University Press.

[18] Schroeyen, F. (forthcoming) "An alternative way to model merit good arguments". *Journal of Public Economics.*

[19] Tungodden, B. (forthcoming) "Responsibility and redistribution: the case of first best". *Social Choice and Welfare.*