Abstract
This paper proposes four concepts of exploitation that encapsulate common uses of the word in social interactions: unfair advantage, unequal exchange, using persons as means, and free-riding. It briefly discusses how these concepts appear in the literature (the first two are prominent in Roemer’s classical work), and then examines how these forms of exploitation are related and how they can occur.

Keywords
Exploitation; Roemer; social justice

1. Introduction
In his landmark book *A General Theory of Exploitation and Class*, John Roemer thoroughly examined how to model the Marxian theory of labor exploitation. This theory relies on the idea that workers give more labor to their employers than they receive through the goods their wages can afford. One difficulty in this theory is to identify the quantity of labor that the workers receive in payment. In Marx’s approach, the labor theory of value was supposed to make this computation easy because market prices and wages would reflect the labor cost of producing the goods (and of producing the labor power, in the case of wages). But the labor theory of value is hard to reconcile with modern general equilibrium theory and appears discredited. Moreover, Roemer showed that, formally, there is little reason to focus on labor, as any other commodity (e.g. coal) could serve as the basis of value or as the currency for the computation of the unequal exchange workers are involved in.

The general conclusion of Roemer’s book was that Marx’s approach, with its focus on labor, should be abandoned and replaced by an approach that highlights the inequalities in endowments that give an advantage to the wealthy in the distribution of commodities and leisure. He proposed an alternative abstract definition of exploitation which involves
the cooperative-game notion of a value function, i.e. a function $v(S)$ defining the total payoff that coalition $S$ (a subgroup of the population $N = \{1, \ldots, n\}$) would obtain if it had to live with a fair share of the total endowment available in society. The coalition $S$ is then said to exploit its complement $N \setminus S$ in the allocation $(x_1, \ldots, x_n)$ if it is better off than under a fair distribution, while its complement is worse off:

\[(a) \sum_{i \in S} x_i > v(S);
(b) \sum_{i \notin S} x_i < v(N \setminus S).\]

In addition, Roemer stipulated that exploitation really occurs when there is a relation between the two groups and $S$ ‘dominates’ $N \setminus S$ in this interaction (e.g. members of $S$ hire and direct the work of members of $N \setminus S$).

Elsewhere (Fleurbaey, 1996) I have argued that the Marxian focus on labor is useful when it highlights the fact that unequal societies, since the origins of history, have witnessed an unequal division of advantages characterized by the coexistence of an elite living well and enjoying pleasant activities while the unpleasant, dangerous, degrading jobs are the lot of the poor. In this perspective, as shown by Roemer, the accounting of quantities of labor incorporated in commodities can be performed independently of the labor theory of value, taking the market prices as given. The labor theory of exploitation can then be consistent and morally relevant even if the labor theory of value is flawed.1

Roemer, however, rightly questioned whether the labor theory of exploitation can serve as the basis for a comprehensive theory of distributive justice. When individual preferences over leisure and consumption differ in the population, some well-endowed individuals may be exploited in labor terms because they don’t mind working a lot. Avoiding exploitation would impose proportionality between consumption and labor. The principle that consumption should be proportional to labor has indeed been studied in the theory of fair allocation (Moulin, 1990; Roemer and Silvestre, 1993), but this theory has made it clear that there are many other ways of conceiving equality or fairness in the distribution of consumption and leisure.

So, the broader conclusion that Roemer (1982, 1985) put forth, and that is reflected in his own seminal work on equality of opportunity over the decades following the 1982 book, is that studying theories of social justice is more relevant than focusing on a very specific notion of exploitation. This broad move to a more general question, however, leaves it possible for the notion of exploitation to play a role in the definition of justice or, perhaps more directly, in the definition of injustice. This is the alley that this paper explores.

The purpose of this paper is to examine various ways in which the notion of exploitation can be understood and given a precise formulation. It appears that four different sorts of exploitation, at least, can be distinguished. First, exploitation can be understood as ‘unfair advantage’ due to inequalities in endowments, and Roemer’s proposal of an alternative definition pertains to this approach. Second, exploitation can be spotted when there is an ‘unequal exchange’, and the Marxian approach seems to be inspired by this idea. Third, exploitation can also refer to a situation in which some individuals are used as means or resources by others, and stand to benefit from their own capacities less than their exploiters. Finally, exploitation is also often mentioned when the deviant behavior
of some individuals gives them an advantage over their conscientious fellows. The typical example is the surfer who lives on income support funded by the taxes paid by the hardworking citizens. The four notions obviously refer, in different ways, to a situation in which some take an unfair advantage at the expense of others.

These various forms have been mentioned in the literature. While the Marxian literature has emphasized the unequal exchange conception, this conception also lies at the heart of Wertheimer’s (1996) theory of exploitation as departure from hypothetical competitive market prices in a trade. Exploitation due to unfair advantage in the prevailing circumstances (e.g. the distribution of endowments) is Roemer’s alternative proposal and also features prominently in Wood (1995) and Goodin (1987). The idea that the exploited are used as means is pervasive in all the literature, although it is seldom treated as a separate idea, except perhaps in Sample’s (2003) conception of exploitation as degradation and in Wolff (1999). The exploitation of good citizens by feckless clients of welfare support has become a popular theme in the luck egalitarian literature after Rawls (1982) and Dworkin (1981).

In this paper we put aside the issue of whether exploitation requires some direct relation or interaction between the exploiters and their victims, a relation of ‘domination’ as stipulated by Roemer. It is clear that one can take an unfair advantage without directly interacting with the disadvantaged. This issue is very well examined in Skillman (1995) in the context of the unequal exchange theory of exploitation. Whether exploitation in any other sense requires the combination of unfair advantage with a direct interaction between the parties is debatable. It seems odd to say that A exploits B when there is no relation whatsoever between them, but there appears to be considerable latitude about the required form of interaction. This question is by and large put aside here.

Obviously, it should not be ignored in a broader theory of justice, and one may even claim that power relations (especially workplace democracy) are more strategic for social progress in our era than exploitation issues. But this paper is about exploitation, not about social justice widely conceived. However, it turns out that one notion of exploitation studied in this paper lends itself naturally to illustrations in terms of power relations in the workplace. It is not surprising that some notions of exploitation are more closely connected with direct social relations while other notions have more to do with arm’s-length or indirect resource transfers. Moreover, for every notion of exploitation, we will seek to formulate a bilateral definition that makes it possible to say that $i$ exploits $j$, rather than simply say that $i$ is exploited or is an exploiter (a form of balance-sheet computation that simply assesses $i$’s situation with respect to the rest of society). Bilateral definitions can more easily be enriched with an analysis of the associated power relations.

This paper is organized as follows. The next section proposes stylized formalizations of the four notions of exploitation. Then, Section 3 incorporates them in a more general model in which they can all appear. This section examines the relations between the four notions in this context. The last section concludes.

2. Four sorts of exploitation

In this section, four simple models are used to formalize the four notions of exploitation highlighted in this paper. In all models there is a finite population $N = \{1, \ldots, n\}$. 
2.1. Unfair endowments

Suppose that individual \( i \in N \) has an endowment \( c_i \in \mathbb{R}_+ \) (like ‘capital’ or ‘capacity’ or ‘circumstance’), the profile of endowments being denoted \( c_N = (c_1, \ldots, c_n) \). There is a mechanism that distributes advantages to individuals as a function of the profile of endowments: \( x_i = \varphi_i (c_N) \). It is assumed that \( \varphi_i \) is increasing in \( c_i \), continuous in \( c_N \), and that the total pie \( \sum_i x_i \) is increasing in each component of \( (c_1, \ldots, c_n) \).

Let us focus on two individuals, \( i \) and \( j \), fixing the endowments of the other individuals, which are denoted \( c_{-ij} \). If one takes equality of the endowments as the benchmark for fairness, one can think that there is exploitation when \( c_i > c_j \) and one of the following situations prevails:

1. \( i \) benefits and \( j \) suffers compared to an equal sharing of their total endowment:

\[
x_i > \varphi_i \left( \frac{c_i + c_j}{2}, \frac{c_i + c_j}{2}, c_{-ij} \right) \quad \text{and} \quad x_j < \varphi_j \left( \frac{c_i + c_j}{2}, \frac{c_i + c_j}{2}, c_{-ij} \right)
\]

2. \( i \) benefits and \( j \) suffers from the fact that \( i \) has more than \( c_j \):

\[
x_i > \varphi_i \left( c_j, c_j, c_{-ij} \right) \quad \text{and} \quad x_j < \varphi_j \left( c_j, c_j, c_{-ij} \right)
\]

3. \( i \) benefits and \( j \) suffers from the fact that \( j \) has less than \( c_i \):

\[
x_i > \varphi_i \left( c_i, c_i, c_{-ij} \right) \quad \text{and} \quad x_j < \varphi_j \left( c_i, c_i, c_{-ij} \right)
\]

A more general approach to fair shares can be adopted. Let us assume that for every \( C \in \mathbb{R}_+ \), there is an exogenous fairness rule that defines the optimal sharing of \( C \) between \( i \) and \( j \), i.e., there is an optimal \( (\hat{c_i}(C), \hat{c_j}(C)) \) such that \( \hat{c_i}(C) + \hat{c_j}(C) = C \). Equal sharing is one example, but unequal shares may be justified, for instance, if one of the individuals has greater needs or a greater ability to benefit from the endowment. We further assume that \( \hat{c_i}(C), \hat{c_j}(C) \) are both continuous and increasing in \( C \).

With this notion one can then define three notions of exploitation. The first one is the immediate generalization of definition 1 above.

**Simple exploitation:** \( i \) benefits and \( j \) suffers compared to a fair sharing of their total endowment:

\[
c_i > \hat{c}_i \left( c_i + c_j \right) \quad \text{and} \quad c_j < \hat{c}_j \left( c_i + c_j \right)
\]

\[
x_i > \varphi_i \left( \hat{c}_i \left( c_i + c_j \right), \hat{c}_j \left( c_i + c_j \right), c_{-ij} \right) \quad \text{and} \quad x_j < \varphi_j \left( \hat{c}_i \left( c_i + c_j \right), \hat{c}_j \left( c_i + c_j \right), c_{-ij} \right)
\]

In this definition, the condition on endowments may appear superfluous. However, it rules out the case in which \( \varphi_i \) is more sensitive to \( c_j \) than \( c_i \) so that \( i \) benefits when the deviation from the fair shares is in favor of \( j \), while symmetrically \( j \) suffers from such a deviation because \( \varphi_j \) is more sensitive to \( c_i \) than \( c_j \). Perhaps one can also talk about exploitation in this case, but it creates a strange pattern in which \( i \) would like to give part of his endowment to \( j \), but \( j \) would refuse. Here we restrict attention to the more standard case in which every agent would benefit from grabbing the others’ endowment.

This approach, therefore, defines exploitation as the combination of two things: unfair shares and opposite consequences for the individuals. It is the latter that differentiates
exploitation from a mere notion of unfairness. There is unfairness whenever the shares differ from what they should be, but there is exploitation when, in addition, one individual benefits from having more than warranted and the other suffers from having less.

It is important not to interpret \( \hat{c}_i (c_i + c_j) , \hat{c}_j (c_i + c_j) \) as the fair shares of \( i,j \) in the global distribution. It might be that they are, together, grossly advantaged or disadvantaged. What is at stake here is only a fair distribution of their current joint endowment, without any assessment of the fairness of \( c_i + c_j \) with respect to the rest of society.

Simple exploitation is the natural generalization of the first of the three cases listed in the beginning of this section. The other two cases refer to two extreme counter-factual egalitarian distributions, one in which both individuals have either the lowest or the greatest endowment. These two cases can be accommodated in our more general setting if one modifies the definition of simple exploitation by considering alternative values for \( C \).

Two notions can be proposed, a weak and a strong one. The weak notion is similar to saying that there is exploitation if either of the three cases listed in the beginning holds, and says that there is exploitation if for a fair sharing of some \( C \), one has the typical configuration of one individual benefitting from having more than a fair share and the other suffering from having less.

**Weak exploitation:** \( i \) benefits and \( j \) suffers compared to a fair sharing of some total endowment: there is \( C \) such that

\[
\begin{align*}
  & c_i > \hat{c}_i (C) \quad \text{and} \quad c_j < \hat{c}_j (C) \\
  & x_i > \varphi_i \left( \hat{c}_i (C), \hat{c}_j (C), c_{-ij} \right) \quad \text{and} \quad x_j < \varphi_j \left( \hat{c}_i (C), \hat{c}_j (C), c_{-ij} \right)
\end{align*}
\]

The more restrictive variant is like saying that there is exploitation if all three cases listed above hold at the same time. This more restrictive definition checks that the imbalance of advantage and disadvantages holds for a whole set of relevant situations.

**Strong exploitation:** \( i \) benefits and \( j \) suffers compared to a fair sharing in a specific range: there is \( C \) such that \( c_i > \hat{c}_i (C) \) and \( c_j < \hat{c}_j (C) \), and for all such \( C \),

\[
\begin{align*}
  & x_i > \varphi_i \left( \hat{c}_i (C), \hat{c}_j (C), c_{-ij} \right) \quad \text{and} \quad x_j < \varphi_j \left( \hat{c}_i (C), \hat{c}_j (C), c_{-ij} \right)
\end{align*}
\]

There is a convenient graphical way of illustrating and analyzing these definitions (Figure 1). The path of \( \hat{c}_i (C) , \hat{c}_j (C) \) is an increasing curve in \( (c_i, c_j) \) space. The functions \( \varphi_i, \varphi_j \) define indifference maps in the same space (for a fixed value of \( c_{-ij} \)).

The following proposition studies the logical relationship between these three notions.

**Proposition 2.1.** **Strong⇒ Simple ⇒ Weak.** The three notions are equivalent if \( \varphi_i \) is decreasing in \( c_j \) and \( \varphi_j \) is decreasing in \( c_i \).

**Proof.** That Simple ⇒ Weak follows from the definitions.

**Strong ⇒ Simple.** By assumption, there is \( C \) such that \( c_i > \hat{c}_i (C) \) and \( c_j < \hat{c}_j (C) \). Suppose that \( \hat{c}_i (c_i + c_j) \geq c_i \). As \( \hat{c}_i (c_i + c_j) + \hat{c}_j (c_i + c_j) = c_i + c_j \), this implies
\[ \hat{c}_j (c_i + c_j) \leq c_j. \]

But as \( \hat{c}_i (C), \hat{c}_j (C) \) are increasing, it is then impossible to find \( C \) such that \( \hat{c}_i (C) < c_i \) and \( \hat{c}_j (C) > c_j \). Therefore \( \hat{c}_i (c_i + c_j) < c_i \), which implies \( \hat{c}_j (c_i + c_j) > c_j \). Strong exploitation then implies that for \( C = c_i + c_j, x_i > \varphi_i (\hat{c}_i (C), \hat{c}_j (C), c_{-ij}) \) and \( x_j < \varphi_j (\hat{c}_i (C), \hat{c}_j (C), c_{-ij}) \), which corresponds to simple exploitation.

The three notions are equivalent if \( \varphi_i \) is decreasing in \( c_j \) and \( \varphi_j \) is decreasing in \( c_i \).

Weak exploitation implies that there is \( C \) such that \( c_i > \hat{c}_i (C) \) and \( c_j < \hat{c}_j (C) \). Let \( C' \) be such that \( \hat{c}_i (C') < c_i \) and \( \hat{c}_j (C') > c_j \). Since \( \varphi_i \) is increasing in \( c_i \) and decreasing in \( c_j \), \( \varphi_i (\hat{c}_i (C'), \hat{c}_j (C'), c_{-ij}) < \varphi_i (c_i, c_j, c_{-ij}) \). Similarly, \( \varphi_j (\hat{c}_i (C'), \hat{c}_j (C'), c_{-ij}) > \varphi_j (c_i, c_j, c_{-ij}) \), which proves that strong exploitation holds.

The case in which \( \varphi_i \) is decreasing in \( c_j \) and \( \varphi_j \) is decreasing in \( c_i \) easily occurs when payoffs are proportional to \( c_i \) but with a coefficient of proportionality (e.g. a wage rate) that decreases with the sum of \( c \): \( \varphi_i (c_N) = p \left( \sum_k c_k \right) c_i \), with \( p \) a decreasing function. The sum of payoffs remains increasing in \( c_i \) provided that \( p \) is not too decreasing (the elasticity of \( p \) must be greater than \(-1\)).

One can have any of the three forms of exploitation even in absence of interactions between \( i \) and \( j \), i.e. even when \( \varphi_i \) does not depend on \( c_j \) and \( \varphi_j \) does not depend on \( c_i \).

The above definitions can be extended to groups. In particular, if one wants to describe the fact that coalition \( S \) exploits its complement \( N \setminus S \) in the simple sense, one can write

\[
\sum_{i \in S} x_i > \sum_{i \in S} \varphi_i \left( \hat{c}_N \left( \sum_{j \in N} c_j \right) \right) \quad \text{and} \quad \sum_{i \notin S} x_i < \sum_{i \notin S} \varphi_i \left( \hat{c}_N \left( \sum_{j \in N} c_j \right) \right)
\]

These two inequalities are the same as those posited by Roemer if one defines the value function in his approach as

\[
v(S) = \sum_{i \in S} \varphi_i \left( \hat{c}_N \left( \sum_{j \in N} c_j \right) \right)
\]
One drawback of this coalitional approach, though, is that it is compatible with some members of S being disadvantaged even if the group as a whole is advantaged. An alternative definition that checks that every member of one group is advantaged and every member of the other group is disadvantaged is possible.

**Simple exploitation** The members of subgroup $M$ exploit the members of subgroup $M'$ if:

(a) for every $i \in M$, $c_i > \hat{c}_i \left( \sum_{k \in M \cup M'} c_k \right)$ and $x_i > \varphi_i \left( \hat{c}_{M \cup M'} \left( \sum_{k \in M \cup M'} c_k \right), c_{M-M'} \right)$;

(b) for every $i \in M'$, $c_i < \hat{c}_i \left( \sum_{k \in M \cup M'} c_k \right)$ and $x_i < \varphi_i \left( \hat{c}_{M \cup M'} \left( \sum_{k \in M \cup M'} c_k \right), c_{M-M'} \right)$.

Similar extensions of the notions of weak and strong exploitation are immediately conceivable.

For further reference, the notions of exploitation studied in this subsection will be called A-exploitation (A for ‘advantage’).

### 2.2. Unequal exchange

The simplest way to model the phenomenon of unequal exchange is to assume that $j$ sells some object $e_j$ to $i$, for which $i$ pays $p_i$. There is unequal exchange when $p_i \neq V(e_j)$, for some ‘fair’ value function $V$.

The difficulty in this case is to define what the fair value is. In Marxian theory, it is the quantity of labor contained in $e_j$, which must be embodied in the goods that the wage $p_i$ will buy on the commodity market. Exploitation is then due to the fact that the market wage does not pay the quantity of labor contained in $e_j$, but only the lower quantity of labor needed to produce $e_j$ (labor power). Similarly, in Yoshihara and Veneziani (2009), $i$ is exploited when the share of $i$’s labor in the total labor of the economy is less than his income share. This is tantamount to requiring labor to be paid at a wage equal to the average labor productivity in the economy.

An alternative approach to the fair value is to refer to a counterfactual distribution of endowment parameters that have an impact on the payment. Suppose that, for a given value of $e_j$, the level of $p_i$ is influenced by $c_N$. The fair value could be defined as $p_i \left( e_j, \hat{c}_N \left( \sum_{k \in N} c_k \right) \right)$, where $\hat{c}_N \left( \sum_{k \in N} c_k \right)$ is the fair distribution of the total endowment $\sum_{k \in N} c_k$.

To take a simple example, suppose that $p_i$ is determined by a bargaining process between $i$ and $j$, depends only on $c_i, c_j$, and is decreasing in $c_i$ and increasing in $c_j$. Then

$$V(e_j) = p_i \left( e_j, \hat{c}_i \left( c_i + c_j \right), \hat{c}_j \left( c_i + c_j \right) \right)$$

and unequal exchange occurs when either $c_i > \hat{c}_i \left( c_i + c_j \right)$ and $c_j < \hat{c}_j \left( c_i + c_j \right)$, or $c_i < \hat{c}_i \left( c_i + c_j \right)$ and $c_j > \hat{c}_j \left( c_i + c_j \right)$.

One then sees a possible connection between exploitation due to unfair endowment and unequal exchange. Assume that $e_j$ is fixed and that

$$\varphi_i \left( c_N \right) = F(e_j) - p_i \left( e_j, c_i, c_j \right)$$

$$\varphi_j \left( c_N \right) = p_i \left( e_j, c_i, c_j \right) - U(e_j).$$

(One may think of $F(e_j)$ as the product of $j$’s effort $e_j$, and of $U(e_j)$ as the utility loss for $j$ of making this effort.) In this case, the three notions of A-exploitation from the previous
subsection are equivalent, by Proposition 2.1, and A-exploitation occurs if and only if unequal exchange takes place.

If \( e_j \) is not fixed and is influenced by \( c_i, c_j \), the relation between the two approaches (A-exploitation and unequal exchange) is more complex, but the equivalence still obtains under certain assumptions.

**Proposition 2.2.** Assume that

\[
\phi_i(c_i, c_j) = F(e_j(c_i, c_j)) - p_i(e_j(c_i, c_j), c_i, c_j)
\]

\[
\phi_j(c_i, c_j) = p_i(e_j(c_i, c_j), c_i, c_j) - U(e_j(c_i, c_j))
\]

with \( e_j(c_i, c_j) \) increasing in \( c_i \) and decreasing in \( c_j \), \( F \) and \( U \) increasing, and \( p_i \) decreasing in \( c_i \) and increasing in \( c_j \). Unequal exchange and A-exploitation coincide if and only if \( \phi_i(c_i, c_j) \) is non-decreasing in \( c_i \) and non-increasing in \( c_j \) (and never constant in both) and \( \phi_j(c_i, c_j) \) is non-increasing in \( c_i \) and non-decreasing in \( c_j \) (and never constant in both). Each of the following conditions is sufficient for this to hold:

(i) \( f(c_i, c_j) = p_i(e_j(c_i, c_j), c_i, c_j) \) is decreasing in \( c_i \) and increasing in \( c_j \).

(ii) \( F(e) - p_i(e, c_i, c_j) \) increases in \( e \) and \( p_i(e, c_i, c_j) - U(e) \) decreases in \( e \).

(iii) All functions are differentiable and for \( k = i, j, X = F, U \),

\[
\frac{\partial c_k}{\partial e_j} \bigg|_{X-p_i=\text{cst}} \frac{\partial e_j}{\partial c_k} > -1
\]

**Proof.** The first part is proved by observing that under the assumption that \( \phi_i(c_i, c_j) \) is non-decreasing in \( c_i \) and non-increasing in \( c_j \) (and never constant in both) and \( \phi_j(c_i, c_j) \) is non-increasing in \( c_i \) and non-decreasing in \( c_j \) (and never constant in both), A-exploitation occurs if and only if \( \phi_i(c_i, c_j) \) is increasing in \( c_i \) and increasing in \( c_j \), unequal exchange occurs if and only if \( \phi_j(c_i, c_j) \) is decreasing in \( c_i \) and increasing in \( c_j \), unequal exchange occurs if and only if \( \phi_i(c_i, c_j) \) is decreasing in \( c_i \) and increasing in \( c_j \), unequal exchange occurs if and only if \( \phi_j(c_i, c_j) \) is decreasing in \( c_i \) and increasing in \( c_j \).

That (i) is a sufficient condition is obvious.

Condition (ii) is also sufficient because when \( c_i \) increases or \( c_j \) decreases, \( F(e) - p_i(e, c_i, c_j) \) increases with \( e = e_j(c_i, c_j) \), and further increases because \( p_i(e, c_i, c_j) \) decreases in \( c_i \) and increases in \( c_j \). A similar reasoning applies to \( p_i(e, c_i, c_j) - U(e) \).

Condition (iii) is derived from computing \( \partial \phi_i(c_i, c_j)/\partial c_i > 0, \partial \phi_i(c_i, c_j)/\partial c_j < 0, \partial \phi_j(c_i, c_j)/\partial c_i < 0, \partial \phi_j(c_i, c_j)/\partial c_j > 0 \). For instance, \( \partial \phi_i(c_i, c_j)/\partial c_i > 0 \) is equivalent to:

\[
\left[ \frac{dF}{de_j} - \frac{\partial p_i}{\partial e_j} \right] \frac{\partial e_j}{\partial c_i} + \frac{\partial p_i}{\partial c_i} > 0
\]

\[
\left[ \frac{\partial (F-p_i)/\partial e_j}{\partial (F-p_i)/\partial c_i} \right] \frac{\partial e_j}{\partial c_i} > -1
\]
while \( \partial \varphi_i (c_i, c_j) / \partial c_j < 0 \) is equivalent to

\[
\left[ \frac{dF}{de_j} - \frac{\partial p_i}{\partial e_j} \right] \frac{\partial e_j}{\partial c_j} - \frac{\partial p_i}{\partial c_j} < 0 \]
\[
\left[ \frac{\partial (F - p_i)}{\partial e_j} / \partial c_j \right] \frac{\partial e_j}{\partial c_j} > -1
\]

Condition (i) is easy to understand, as it means that the total payment, not just the unit price, favors the individual with greater \( c \). This can occur, for instance, when \( c \) has a strong influence on bargaining power. Condition (ii) captures the case in which \( e_j \) is paid less than its marginal productivity and costs \( j \) more than its marginal payment. In this case, an advantage for \( c_i \) over \( c_j \) not only distorts the unit price \( p_i/e_j \) in favor of \( i \), but as it also raises \( e_j \), it is doubly good for \( i \) and bad for \( j \). Condition (iii) means that the marginal rate of substitution between \( c_j \) (or \( c_i \)) and \( e_j \) in the payoff function never offsets the direct influence of \( c_i \) (or \( c_j \)) on \( e_j \). This condition is sufficient, and is almost necessary for \( \varphi_i (c_i, c_j) \) being non-decreasing in \( c_i \) and non-increasing in \( c_j \) (and never constant in both) and \( \varphi_j (c_i, c_j) \) being non-increasing in \( c_i \) and non-decreasing in \( c_j \) (and never constant in both), because under differentiability it is equivalent to \( \varphi_i (c_i, c_j) \) being increasing in \( c_i \) and decreasing in \( c_j \) and \( \varphi_j (c_i, c_j) \) being decreasing in \( c_i \) and increasing in \( c_j \).

This simple model shows how a discrepancy between A-exploitation and unequal exchange can occur. In the model, \( c_i, c_j \) appear only as arguments of \( p_i \) and \( e_j \). If, instead, an advantage in \( c \) gave extra benefit independently of its influence on the payment \( p_i \) or the quantity exchanged, then one could have A-exploitation in absence of unequal exchange, or even when unequal exchange goes in the opposite direction. This can occur, for instance, with the slightly richer model in which \( c \) has a direct influence on productivity and disutility:

\[
\varphi_i (c_N) = F(e_j, c_i) - p_i (e_j, c_i, c_j) \\
\varphi_j (c_N) = p_i (e_j, c_i, c_j) - U(e_j, c_j)
\]

Let us conclude this subsection with a remark on the Marxian approach. In the Marxian models like Roemer’s, exploitation is understood as unequal exchange but is treated as an impersonal feature of the situation. An individual can be an exploiter or an exploited agent (or neither), as a function of the labor quantities he contributes and receives, but the identity of his trade partners is not used. Therefore, it would be theoretically possible for an agent to be, for instance, a net exploiter even though he suffers from unequal exchange with some trade partners (for instance, he could pay some workers too little and others too much). In the approach proposed here, individual relations are disentangled and it is possible for an agent to exploit some partners and be exploited by other partners.

However, the approach proposed here still makes it possible to compute whether the net balance of all of \( i \)'s trades are in his favor or his disadvantage, compared to a fair benchmark. This will be illustrated in the next section.

### 2.3. Being used as a means

It is common to define exploitation in terms of the exploiter using the exploited as a means rather than an end, in reference to Kant. But this notion has, to the best of my
knowledge, never been formalized. This is a challenging notion because, in standard economic models of trade and strategic interaction, one can think that every agent sees the other agents as means to the pursuit of his own objectives.

In order to give a more restrictive scope to the notion, one can perhaps refer to the archetypal form of use as a means that is found in slavery. There, what happens is that the capacity of the slave is under the control of the master. This has two consequences. First, the master does not pay the slave a fair price for his services. This is simply the idea of unequal exchange. Second, the master reaps all the benefits of any enhancement to the (observable) capacity of the slave. Just as the owner of a mine benefits when the reserves prove to be greater than expected, the owner of a slave benefits when the slave is more able than expected.

Therefore, defining exploitation in terms of using the exploited as a means can be related both to the concept of unequal exchange and to this new notion of greater benefit from the capacity of the exploited. For the sake of clarity we will associate the notion of exploitation as a means to the latter. For further reference, let M-exploitation (M as in ‘means’) denote this form of exploitation.

Consider the model of the previous subsection, in which \( e_j \) is sold to \( i \) by \( j \):

\[
\phi_i(c_N) = F(e_j) - p_i(e_j, c_i, c_j)
\]

\[
\phi_j(c_N) = p_i(e_j, c_i, c_j) - U(e_j)
\]

One can say that \( j \) is used as a means when an increase in \( e_j \) benefits \( i \) more than it benefits \( j \).\(^4\) Assuming differentiability, this occurs when for all \( e \) in a relevant range,\(^5\)

\[
F'(e) - \frac{\partial p_i}{\partial e_j}(e, c_i, c_j) > \frac{\partial p_i}{\partial e_j}(e, c_i, c_j) - U'(e)
\]

Condition (ii) in Proposition 2.2, in the case of differentiability, amounts to

\[
F'(e) - \frac{\partial p_i}{\partial e_j}(e, c_i, c_j) > 0 > \frac{\partial p_i}{\partial e_j}(e, c_i, c_j) - U'(e)
\]

Therefore this condition, which is sufficient to render A-exploitation and unequal exchange equivalent, is also sufficient to reveal the occurrence of M-exploitation.

But one can see that M-exploitation and the other forms of exploitation are quite independent. One can have M-exploitation in the absence of A-exploitation or unequal exchange, and conversely. This is due to the fact that M-exploitation refers to variations whereas the other forms are about levels.

Now, go back to the model of the first subsection: \( \phi_i(c_N), \phi_j(c_N) \). One can also detect M-exploitation when \( i \) benefits more from an increase in \( c_j \) than \( j \) does, i.e. assuming differentiability:

\[
\frac{\partial \phi_i}{\partial c_j}(c_N) > \frac{\partial \phi_j}{\partial c_j}(c_N)
\]

Once again, the link with A-exploitation is tenuous. Recall that we assumed \( \phi_j \) to be increasing in \( c_j \). Therefore M-exploitation requires \( \phi_i \) to be even more increasing in \( c_j \).
If \( c_i > \hat{c}_i (c_i + c_j) \) and \( c_j < \hat{c}_j (c_i + c_j) \), then in the presence of M-exploitation it is not guaranteed that
\[
\varphi_i (c_N) > \varphi_i (\hat{c}_i (c_i + c_j), \hat{c}_j (c_i + c_j), c_{-ij})
\]
because the fact that \( c_j < \hat{c}_j (c_i + c_j) \) tends to lower \( \varphi_i (c_N) \). Or, in other words, simple A-exploitation requires M-exploitation not to be too strong.

A peculiar feature of M-exploitation is that it can occur both ways, as nothing precludes that each of the two agents can exploit the other by benefiting more from the other’s capacity than from his own.\(^6\) However, one has the following result:

**Proposition 2.3.** Assume differentiability. \( \varphi_i (c_i, c_j) + \varphi_j (c_i, c_j) = F (c_i + c_j) \), \( \varphi_i (c_i, c_j) \) is increasing in \( c_i \), and \( \varphi_j (c_i, c_j) \) is increasing in \( c_j \). If \( i \) A- and M-exploits \( j \), then \( j \) does not M-exploit \( i \).

**Proof.** If \( i \) M-exploits \( j \), then \( \partial \varphi_i / \partial c_j > \partial \varphi_j / \partial c_j (> 0) \). As \( \varphi_i (c_i, c_j) + \varphi_j (c_i, c_j) = F (c_i + c_j) \), one has
\[
\frac{\partial \varphi_i}{\partial c_j} + \frac{\partial \varphi_j}{\partial c_j} = F' (c_i + c_j)
\]
so that \( \partial \varphi_i / \partial c_j > F'/2 > \partial \varphi_j / \partial c_j \). Similarly, if \( j \) M-exploits \( i \), one has \( \partial \varphi_j / \partial c_i > F'/2 > \partial \varphi_i / \partial c_i \).

Consider a small transfer from \( c_j \) to \( c_i \), i.e. \( dc_i = -dc_j > 0 \). The resulting change in \( \varphi_i, \varphi_j \) is:
\[
d\varphi_i = \left( \frac{\partial \varphi_i}{\partial c_i} - \frac{\partial \varphi_i}{\partial c_j} \right) dc_i < 0 < \left( \frac{\partial \varphi_j}{\partial c_i} - \frac{\partial \varphi_j}{\partial c_j} \right) dc_i = d\varphi_j
\]
Therefore it is impossible to have \( i \) A-exploiting \( j \). \( \blacksquare \)

This result is rather intuitive. Exploitation in level (A-exploitation) and in variation (M-exploitation) is incompatible with the symmetric exploitation in variation.

### 2.4. Exploiting by deviating

The fourth notion of exploitation that remains to be introduced does not refer to endowments or prices, in the context of sharing private benefits. Instead, it has more to do with contributions to a public good. Those who shirk and fail to make the expected contribution exploit the others.

This can be modelled along the vein of the simple models presented in the previous subsections. Suppose that individual advantage now depends on the profile of effort (or contribution) in society: \( x_i = \varphi_i (e_N) \). Now it is assumed that \( \varphi_i \) is decreasing in \( e_i \) at least after some point.

In similar fashion as before, but with a totally different interpretation, one can introduce a norm of fair contributions that \( i \) and \( j \) should make, given the contributions of the others. So, let us fix \( e_{-ij} \) and focus on \( e_i, e_j \). For every total contribution \( E \) that \( i \) and \( j \) can make, there is a fair division \( \hat{c}_i (E), \hat{c}_j (E) \) such that \( \hat{c}_i (E) + \hat{c}_j (E) = E.\(^7\)
One can then define simple exploitation as occurring when the following happens:

**Simple exploitation:** $i$ benefits and $j$ suffers compared to a fair sharing of their total contribution:

$$e_i < \hat{e}_i (e_i + e_j) \text{ and } e_j > \hat{e}_j (e_i + e_j)$$

$$x_i > \varphi_i (\hat{e}_i (e_i + e_j), \hat{e}_j (e_i + e_j), e^{-ij}) \text{ and } x_j < \varphi_j (\hat{e}_i (e_i + e_j), \hat{e}_j (e_i + e_j), e^{-ij})$$

As in Section 2.1, one can also define a notion of weak and a notion of strong exploitation.

It may be one contribution of this paper to show the formal proximity between two notions of exploitation, A-exploitation and D-exploitation, which are generally associated with very different parts of the political spectrum. Typically, the left wing is worried about inequalities in riches and the ensuing inequalities in well-being, which correspond to A-exploitation, whereas the right wing is worried about the public good structure of the welfare state and the fact that those who benefit most from it are those who contribute the least, a concern that can be formalized as D-exploitation. Obviously, the two views rely on exactly opposite opinions about the need for redistribution.

But D-exploitation is not just A-exploitation in reverse. In the context of luck egalitarian theories of justice, there is a clear distinction between circumstances and effort. One can then have A-exploitation on behalf of those who are richly endowed in favorable circumstances, and D-exploitation on behalf of those who make little effort and benefit from social help as a consequence. In terms of social groups, the typical situation envisioned by luck egalitarianism is then that the rich and the undeserving poor are both unfairly advantaged, but on different counts (A-exploitation due to endowments for the former, D-exploitation due to laziness for the latter).

For further reference, this form of exploitation will be called D-exploitation (standing for ‘deviate’, or ‘defect’).

### 3. A general model

Roemer’s 1982 book contained a general model in which the members of the population could be classified (this is to be read literally: distributed between various social classes) as a function of their social role as buyers and sellers of labor. This model has inspired many authors and, with variations, is still the topic of recent investigation (Veneziani, 2007; Yoshihara, 2010; Yoshihara and Veneziani, 2009). The focus of this literature is the unequal exchange of labor in the capitalist economy.

This section introduces a more general and more abstract model in which the various forms of exploitation introduced in the previous section can all appear jointly or separately. The model remains a gross simplification because the description of individuals’ characteristics is kept to a minimum in order to ensure tractability.

#### 3.1. The framework

The model is conceived as follows. Every individual $i$ has a (fixed) capacity $c_i \in \mathbb{R}_+$ and a (variable) effort $e_i \in \mathbb{R}_+$. Capacity and effort can be used in one’s firm or sold
(or rented) to other individuals. Let \( c'_i, e'_i \) denote the quantities of \( i \)'s capacity and effort that are put at the disposal of \( j \). One has \( c_i = \sum_{j \in N} c'_i, \ e_i = \sum_{j \in N} e'_i \). Let \( c' = (c'_j)_{j \in N}, \ e' = (e'_j)_{j \in N} \) be the vectors of capacity and effort at the disposal of \( i \). The payment made by \( i \) to \( j \) as a compensation for \( (c'_j, e'_j) \) is denoted \( p'_i \). The total payment made by \( i \) is denoted \( P_i = \sum_{j \neq i} p'_i \), the total payment received by \( i \) is denoted \( P'_i = \sum_{j \neq i} p'_j \).

There is only one commodity produced in the economy, which is used as numeraire. Production takes place in individual firms and the quantity is determined by a production function: \( F(c', e') \). Everyone who sets up a firm has access to the same technology, so that the production function is the same in all firms.

One might consider the possibility of restricting access to the production sector to some individuals only (the entrepreneurs). It might also be interesting to assume that the function \( F \) exhibits increasing returns to scale, at least initially, so that the rich entrepreneurs have an advantage. But we will not make use of such specific assumptions in this paper.

Individual \( i \)'s gross income is equal to \( y_i = F(c', e') - P_i + P'_i \). There is a transfer system that is described by a transfer function \( T(y) \). Individual \( i \)'s net income is \( x_i = y_i + T(y_i) \). The transfer system is assumed to be balanced: \( \sum_i T(y_i) = 0 \). This ensures that total consumption is equal to total production \( \sum_{i \in N} F(c', e') \).

Individual \( i \)'s preferences depend on income and effort and are represented by the utility function \( U_i(x_i, e_i) \).

An economy is the list of data for the fixed characteristics of the population, i.e. \( E = (c_N, U_N) \). An allocation is fully described by listing trades and payments \( z = (c'_i, e'_i, p'_i)_{i,j \in N} \).

The process by which actions and payments are determined is taken as given and is described by a function that determines the allocation for every economy: \( z(E) \). From this one derives functions \( e_N(E), y_N(E), \) and so on. These functions can be interpreted as describing what the economy does at equilibrium.

The last ingredients to be introduced are fairness rules \( \hat{c}_K(K, c_N), \hat{e}_K(K, e_N) \), which define a fair distribution of \( c \) or \( e \) for the members of a subgroup \( K \), as a function of the current distribution (this is shorthand for the fact that they really depend only on the distribution for the complement \( N \setminus K \) and on the total quantities \( \sum_{i \in K} c_i \) and \( \sum_{i \in K} e_i \)). These rules have to satisfy the constraint \( \sum_{i \in K} \hat{c}_i(K, c_N) = \sum_{i \in K} c_i \) and \( \sum_{i \in K} \hat{e}_i(K, c_N) = \sum_{i \in K} e_i \). Let \( \hat{E}_K = (((\hat{c}_K(K, c_N), c_N|K), U_N) \). We also need fair payments \( (\hat{p}'_i)_{i,j \in N} \) for the detection of unequal exchange.

All functions are assumed to be differentiable.

Compared to the Marxian models, this one is simpler because the multiplicity of commodities has disappeared, eliminating the problem of computing the labor content of any bundle of commodities. But it makes it possible to describe a great variety of behaviors and redistributive policies.
3.2. Forms of exploitation

The purpose of this subsection is to adapt the notions of exploitation introduced in the previous section to this new framework. This exercise is not obvious because this richer model allows for many interactions and trades between the agents.

A-exploitation occurs when some agents would be worse off, and others better off, under a fair redistribution of their endowments in $c$. This means that we focus here on ‘simple’ exploitation, as defined in the previous section.

**A-exploitation:** The members of group $M$ A-exploit the members of group $M'$ in economy $E$ if:

(a) for every $i \in M$, $c_i > \hat{c}_i(M, c_N)$ and

$$U_i(x_i(E), e_i(E)) > U_i\left(x_i\left(\hat{E}_{M\cup M'}\right), e_i\left(\hat{E}_{M\cup M'}\right)\right)$$

(b) for every $i \in M'$, $c_i < \hat{c}_i(M', c_N)$ and

$$U_i(x_i(E), e_i(E)) < U_i\left(x_i\left(\hat{E}_{M\cup M'}\right), e_i\left(\hat{E}_{M\cup M'}\right)\right)$$

Unequal exchange occurs between trade partners when the payment differs from the fair reference. We focus on the net payment, although one could also apply the notion to each payment separately, as done in the previous section in which an isolated trade was considered. These various possibilities all seem relevant, but in the interaction between two agents, it is quite natural to focus on the net situation, because unfairness in one exchange compensated by reciprocal unfairness in another exchange is the essence of trade itself.10

**Unequal exchange:** $i$ benefits from unequal exchange at the expense of $j$ if $p^i_i - p^j_j < \hat{p}^i_i - \hat{p}^j_j$.

M-exploitation occurs when one individual stands to benefit more from an increase in another’s capacity or effort than this other individual. It could be measured at the level of utility, but here we apply it at the level of incomes.

A difficulty for the definition of M-exploitation is that one has to describe what happens when the effort of an individual increases, which requires an idea of what happens out of equilibrium. This can be done by introducing price functions that depend directly on actions, i.e. $p^i_i = f^i_i\left(c^i_j, e^i_i\right)$. Such functions do not exactly define a game form, because $c^i_j, e^i_i$ are joint actions of $i$ and $j$.

**M-exploitation:** $i$ M-exploits $j$ with respect to $c^j_j$ if

$$\frac{\partial F}{\partial c^j_j} - \frac{\partial f^j_j}{\partial c^j_j} > \frac{\partial f^j_j}{\partial c^j_j}$$
and with respect to $e_j$ if

$$\frac{\partial F}{\partial e_j^i} - \frac{\partial f_j^i}{\partial e_j^i} > \frac{\partial f_j^j}{\partial e_j^j}$$

One may also consider variants of these definitions, in which the focus is on net income and the marginal tax rate appears. But it does not seem very attractive to entertain the possibility that one agent M-exploits another just because he benefits from a lower tax rate.

Another variant, referring to the more abstract definition of M-exploitation at the end of Section 2.3, would define M-exploitation of $i$ on $j$ with respect to $c_j$ when an increase in $c_j$ benefits $i$ more than $j$ in the equilibrium, i.e. when

$$\sum_{k=1}^{n} \left( \frac{\partial F}{\partial c_k^i} (E) + \frac{\partial F}{\partial e_k^i} (E) \right) - \frac{\partial P_i}{\partial c_j} (E) + \frac{\partial P_i}{\partial c_j} (E) > \sum_{k=1}^{n} \left( \frac{\partial F}{\partial c_k^j} (E) + \frac{\partial F}{\partial e_k^j} (E) \right) - \frac{\partial P_j}{\partial c_j} (E) + \frac{\partial P_j}{\partial c_j} (E)$$

This variant is difficult to analyze because it requires analyzing how the equilibrium depends on endowments. Its drawback is that, under this definition, M-exploitation can occur by a configuration of pecuniary externalities that has little to do with intuitively exploitative relations.

Finally, D-exploitation occurs when some agents deviate from their fair share of effort. The difficulty here, in addition to assessing how prices could change with a change in actions, is that one has to specify what would happen if the agents stuck to their fair share of effort, and this may involve a variety of possible allocations of effort between different firms. Another difficulty is that the transfer system might be unbalanced under a change in the profile of effort, so that one has to redistribute the net balance between the individuals in some way. Let $\tilde{T}$ denote the alternative tax function that would prevail under the reference profile of effort.

**D-exploitation**: Group $M$ D-exploits group $M'$ in economy $E$ if there exists $(\tilde{e}_j^i)_{i \in M \cup M', j \in N}$ such that $\tilde{e}_{M \cup M'} = \hat{e}_{M \cup M'} (M \cup M', e_N)$ and:

(a) for every $i \in M$, $e_i < \tilde{e}_i$ and

$$U_i (E) > U_i \left( \tilde{y}_i + \tilde{T} (\tilde{y}_i), \tilde{e}_i \right)$$

where

$$\tilde{y}_i = F \left( e^i, \tilde{e}_{M \cup M'}, e^i_{-MM'} \right) - \sum_{j \notin M \cup M'} p_i^j - \sum_{j \in M \cup M', j \neq i} f_i^j \left( c_j^i, e_j^i \right) + \sum_{j \neq i} f_i^j \left( c_j^i, e_j^i \right)$$

(ii) for every $i \in M'$, $e_i > \tilde{e}_i$ and

$$U_i (E) < U_i \left( \tilde{y}_i + \tilde{T} (\tilde{y}_i), \tilde{e}_i \right)$$
These various definitions reveal the difficulty of defining such notions in the context of an economic model with a reasonable degree of complexity. One type of difficulty has to do with determining norms of fairness for the distribution of $c$ and $e$ and for payments. Another type of difficulty has to do with the difficulty of defining what would happen if some agents made more effort or changed their trade pattern.

### 3.3. Exploiters and exploited

An important theorem in Roemer’s 1982 book establishes an implication of being a buyer or seller of labor as opposed to being an exploiter or being exploited. It is easy to obtain something similar in this model if one assumes that the price of effort is below the fair price and that the price of capacity is fair. Then $p^j_i < \hat{p}^j_i$ whenever $j$ sells effort to $i$, implying unequal exchange at the expense of $j$ in absence of opposite trade of effort. In the Marxian model, the fact that the price of labor is insufficient is not assumed, but is nevertheless immediately derived from the fact that total production is not consumed by the workers only, so that their wage contains less labor than they themselves deliver.

Note the difference between identifying $i$ as an exploiter and identifying the sale of labor from $j$ to $i$ as an unequal exchange. These observations are recorded in the following proposition.

**Proposition 3.1.** Assume that, at the allocation $z(E)$, for all $i, j$, $p^j_i < \hat{p}^j_i$ iff $e^j_i > 0$. Then $i$ benefits from unequal exchange at the expense of $j$ whenever $e^j_i > 0$ and $e^j_j = 0$.

The obvious proof is omitted. Note that the assumption implicitly means that the price of $c$ is fair, or, at least, is not so unfair as to generate unequal exchange by itself.

We now turn to the case in which $c$ is overpaid, while $e$ is still underpaid. A ‘cross-trade’ is defined as a situation in which the same commodity ($c$ or $e$) is traded in the two directions by two agents between themselves (e.g. $c^j_i > 0$). Assuming that there is no cross-trade, the following result identifies the cases in which $i$ benefits from unequal exchange at the expense of $j$: either $i$ sells capital to $j$ and buys $j$’s labor (with one of the quantities being strictly positive), or $j$ sells capital and labor to $i$ with a great proportion of the latter, or $i$ sells capital and labor to $j$ with a great proportion of the former.

**Proposition 3.2.** Assume that, at the allocation $z(E)$, there are prices $q, s$ such that $p^j_i = qc^j_i + se^j_i$ for all $i, j$. Assume that there are fair prices $\hat{q}, \hat{s}$ such that $q > \hat{q}$ and $s < \hat{s}$, and that there are no cross-trades between any two agents. Then $i$ benefits from unequal exchange at the expense of $j$ iff one of the following cases holds:

(i) $e^j_i \geq 0$, $c^j_i \geq 0$, $e^j_i c^j_i > 0$

(ii) $e^j_i > 0$, $c^j_j > 0$,

\[
\frac{e^j_i}{c^j_j} > \frac{q - \hat{q}}{\hat{s} - s}
\]
(iii) $e^i_j > 0, c^i_j > 0$, 
\[ \frac{e^i_j}{c^i_j} < \frac{q - \hat{q}}{\hat{s} - s} \]

**Proof.** $p^i_j - p^i_j < \hat{p}^i_j - \hat{p}^i_j$ is equivalent to
\[ (q - \hat{q}) (c^i_j - c^j_i) + (s - \hat{s}) (e^j_i - e^i_j) < 0 \]
which occurs only in the three cases listed in the proposition, given that $q > \hat{q}$ and $s < \hat{s}$. ■

Another key theorem in Roemer's analysis connects wealth and social roles. Assuming that individuals seek to maximize leisure subject to meeting their subsistence needs, one finds that wealthier individuals work less and hire more. Therefore, the exploiters are found among the wealthiest and the exploited among the least wealthy. One can obtain similar results in the model introduced here by assuming that a greater $c_i$ implies a greater net sale of $c$ and a lower net sale of $e$. An individual can be considered a net exploiter (in the sense of unequal exchange) if $P_i - P^i < \hat{P}_i - \hat{P}^i$ (where $\hat{P}_i = \sum_{j \neq i} \hat{p}^j_i$ and $\hat{p}^i = \sum_{j \neq i} \hat{p}_j^i$).

**Proposition 3.3.** Assume that, at the allocation $z(E)$, there are prices $q, s$ such that $p^j_i = qc^j_i + se^j_i$ for all $i, j$. Assume that there are fair prices $\hat{q}, \hat{s}$ such that $q \geq \hat{q}$ and $s < \hat{s}$. Moreover, assume that at $z(E)$, for all $i, j$ such that $c_i > c_j, \sum_{k \neq i} (c^k_i - c^k_j) > \sum_{k \neq i} (c^k_j - c^k_i)$ and $\sum_{k \neq i} (e^k_i - e^k_j) < \sum_{k \neq i} (e^k_j - e^k_i)$. Then there is $c^* > c$ such that for all $i \in N$, 
\[ c_i > c^* \iff P_i - P^i < \hat{P}_i - \hat{P}^i \]

**Proof.** $P_i - P^i < \hat{P}_i - \hat{P}^i$ is equivalent to
\[ \sum_{j \neq i} (p^j_i - p^j_i) < \sum_{j \neq i} (\hat{p}^j_i - \hat{p}^j_i) \]
which can also be written as
\[ (q - \hat{q}) \sum_{j \neq i} (c^j_i - c^j_i) + (\hat{s} - s) \sum_{j \neq i} (e^j_i - e^j_i) < 0 \]
The first term is non-increasing in $c$ and the second one is decreasing in $c$, under the assumptions. ■

This brings us to the link between A-exploitation and unequal exchange. This model makes it clear that A-exploitation can occur even in absence of unequal exchange. One only needs to have utility $U_i$ increase with $c_i$, which is easy to obtain when $c_i$ brings income. When effort is underpaid, and is decreasing in $c_i$, then the less wealthy may be
A-exploited and simultaneously suffer from unequal exchange. But it need not be the case that they are A-exploited just because of unequal exchange.

Moreover, as explained by Roemer, the connection between wealth and labor supply and demand, and therefore the connection between wealth and unequal exchange, vanishes when individuals may have heterogenous preferences over consumption and leisure, which corresponds, in the present model, to the case in which effort and endowment are not perfectly correlated in the population. But breaking the link between wealth and unequal exchange need not affect the link between wealth and A-exploitation.

The next proposition, however, shows that the link between wealth and A-exploitation is complicated even when there is no transfer paradox, i.e. when transferring some endowment between two agents makes the recipient better off and penalizes the donor.

**Proposition 3.4.** Assume that the mapping \( z(E) \) is such that \( U_i \) increases and \( U_j \) decreases when a transfer in \( c \) is made from \( j \) to \( i \), other things equal. Let \( \hat{c}_i(K, c_N) = \frac{1}{n} \sum_{j \in K} c_j \) for all \( K \) and all \( i \in K \). Then whenever \( c_i > c_j \), \( i \) A-exploits \( j \). But in a subgroup of size \( k > 2 \), it is not always true that those with greater than average endowment A-exploit those with less than average endowment. It holds true when a transfer in \( c \) between \( i \) and \( j \) does not affect the others’ utility.

**Proof.** Let \( c_i > c_j \). Moving from \((c_i, c_j)\) to \((\frac{c_i + c_j}{2}, \frac{c_i + c_j}{2})\) decreases \( U_i \) and raises \( U_j \), which proves that \( i \) A-exploits \( j \).

If there are more than two agents in a group, transfers between two agents may affect the utility of the others. Those whose endowment is close to the average are likely to be more affected by the changes in prices induced by the equalization of endowments in all the group than by their change in endowment.

When transfers have no externality on the others, one can move from \( c_N \) to \( \hat{c}_N(N, c_N) = \left( \frac{1}{n} \sum_{i \in K} c_i, \ldots, \frac{1}{n} \sum_{i \in K} c_i \right) \) by a sequence of transfers such that every \( i \) such that \( c_i > \frac{1}{n} \sum_{j \in K} c_j \) is always a donor in a transfer and every \( i \) such that \( c_i < \frac{1}{n} \sum_{j \in K} c_j \) is always a recipient (see Bossert and Fleurbaey, 2002). Therefore, necessarily, every \( i \) such that \( c_i > \frac{1}{n} \sum_{j \in K} c_j \) has \( U_i(x_i(E), e_i(E)) > U_i \left(x_i \left( \hat{E}_N \right), e_i \left( \hat{E}_N \right) \right) \) because \( i \)’s utility goes down at every transfer in which \( i \) is involved and stays put at every other transfer; similarly, every \( i \) such that \( c_i < \frac{1}{n} \sum_{j \in K} c_j \) has \( U_i(x_i(E), e_i(E)) < U_i \left(x_i \left( \hat{E}_N \right), e_i \left( \hat{E}_N \right) \right) \).

This result means that even though the rich typically A-exploit the poor, one cannot generally say that all those with greater-than-average endowment exploit all those with less-than-average endowment. It may happen that the upper middle class suffers from the inequalities, or that the lower middle class benefits from the inequalities.

M-exploitation is not directly connected to A-exploitation and unequal exchange in this model. What happens is the following. The first point is that, if the buyer is rational and is not submitted to quantity constraints, the quantity that is traded is never the occasion of M-exploitation because it is paid at a marginal price that is equalized to its marginal productivity. Therefore the seller gets the full marginal productivity and the buyer would not benefit at all from an increase in the quantity exchanged. This is true
independently of the market power of the buyer, not just in a competitive context. This is summarized in the following proposition.

**Proposition 3.5.** If the buyer $i$ is rational and unconstrained, there is no M-exploitation of $c_j^i$ or $e_j^i$.

**Proof.** The optimal choice of $i$ is max $U_i(x_i, e_i)$ s.t. $x_i = y_i + T(y_i), y_i = F(c^i, e^i) - P_i + P^i$, $P_i = \sum_{j \neq i} f_j^i\left(e_j^i, c_j^i\right), P^i = \sum_{j \neq i} f_j^i\left(e_j^i, c_j^i\right)$. The FOC of this program includes

$$\frac{\partial F}{\partial e_j^i} = \frac{\partial f_j^i}{\partial e_j^i}$$

so that

$$\frac{\partial P_i}{\partial e_j^i} > \frac{1}{2} \frac{\partial F}{\partial e_j^i}$$

which proves that $e_j^i$ is not M-exploited. The same holds true for $c_j^i$.

The second point is that M-exploitation occurs for sure on non-contracted elements of the trade. Suppose for instance that a certain quantity $\bar{e}_j^i$ is the object of the contract, but on the spot $i$ can seek to obtain more from $j$ without paying. Then one has M-exploitation to the greatest extent. A variant of this situation is when labor has two elements, time and effort, and only time is the object of contract, while effort is not part of an explicit contract. In the current model, this can be captured by letting $e$ measure the intensity of effort and $c$ measure time. Then, if only $c$ is contracted upon, M-exploitation does not occur for $c$ but it does plague the delivery of $e$.

This analysis sheds light on social conflict about wages. One aspect is the fair wage issue, which has to do with unequal exchange, when workers feel that their wage is unduly low because it would be greater in a society with a more equal distribution of wealth or a better distribution of market power. The other aspect is when workers feel that certain components of their effort are not properly acknowledged and rewarded, and feel cheated and used as an exploitable resource from which as much as possible is extracted. The two aspects are often mixed and blurred, but they are analytically quite distinct and can occur separately and independently. Highly paid workers can still be M-exploited, and workers on piecework can avoid M-exploitation but nevertheless be insufficiently paid.

Adapting D-exploitation to this model reveals a key difference from A-exploitation that was not apparent in the previous section. A-exploitation can be analyzed, as in Roe-mer’s book, by simply looking at equilibrium allocations. In contrast, D-exploitation refers to deviations from norms of action, and therefore requires a description of what happens out of equilibrium.

Another thing that is special about D-exploitation is that it highlights the importance of preferences and behaviors. The exploiters are those whose lazy preferences lead them not to contribute to the general effort. They may not be advantaged in any particular sense, and therefore there is no systematic connection with the other forms of exploitation.
However, a weak connection can occur if one observes that, for given preferences (or behavioral patterns), the rich tend to deliver less effort than the poor. In this case, the group of D-exploiters includes a mixture of rich with standard behavior, who simply use their wealth to buy extra leisure, and poor with lazy tendencies. A strong status of A-exploiter may then push into the group of D-exploiters as well.

This adaptation of the notion of D-exploitation to this model may not perfectly capture the usual character of the ‘surfer’ who does not contribute to the funding of the social system of solidarity and who lives on income support. The common view of free riders of the welfare state excludes the rich who do contribute by paying taxes (even if they do not work) and who do not depend on income support. To better capture this idea, one should define D-exploitation in terms of a norm of income $\hat{y}_i$ rather than a norm of effort $\hat{e}_i$.

Another noteworthy phenomenon is the following. If the norm of effort is rather low, and if individuals are rational and are not constrained to work more than they want, it may happen that there is no exploited group. Even though there can be a group of ‘feckless’ individuals who are better off than if they worked according to the norm, it may happen that there is no disadvantaged counterpart that they exploit, because those who work more than the norm do so to their own interest, and, like the lazy individuals, would also be worse off than at their current situation if they worked no more than the norm. The hardworking can be counted as exploited only if the norm is close to what they do, so that reducing their work to the level of the norm would not be too harmful to them, and would be more than compensated by the tax externality due to the greater contribution of the lazy subgroup.

A configuration with lazy individuals living on income support without exploiting the tax payers who globally benefit from their above-the-norm behavior is especially plausible if the latter actually benefit from A-exploitation. Their greater capacities give them strong incentives to work hard, so that they can reap the benefits of their advantageous endowment.

More generally, in political debates about the ‘excessive generosity’ of the welfare state, it may be interesting to jointly analyze D-exploitation and A-exploitation. When those who feel D-exploited are also A-exploiters, the transfers made by the welfare system may be described as feeding the former exploitation but also reducing the latter.

This is an area where the analysis of exploitation comes nicely close to recent developments of theories of justice as equality of opportunity. Transfers that reduce inequalities due to circumstances for which individuals are not responsible may distort incentives. Insofar as the apparent laziness of the disadvantaged is influenced by the incentive structure, they should not be held responsible for living below the working norm. A proper notion of D-exploitation should then be formulated not in terms of ordinary working behavior but in terms of a deeper notion of effort that takes account of the circumstances. Roemer’s (1998) suggestion to measure effort as one’s position in the distribution of people with similar circumstances is precisely meant to go in this direction.
4. Conclusion

Exploitation is a multifaceted notion. Analyzing its various forms may help understand social conflicts, because the protesters often voice complaints motivated by the feeling of being at the wrong end of some form of exploitation.

One form of exploitation, called A-exploitation in this paper, is unfair disadvantage due to an unjust distribution of endowments. The Occupy movement focuses on the distribution of wealth and income and points to the increasing gap between the top 1% and the rest of the population. What motivates the movement is not just the distribution itself, but the fact that it generates multiple unfair advantages for the richest, which the 99% at the bottom are deprived of. Redistributing wealth and income would also improve the distribution of well-being, which is the essence of A-exploitation.

Another form of exploitation is unequal exchange. Social conflicts about prices and wages are common, but it has been shown here that they may bear on two different issues connected to two different forms of exploitation. One is the problem of the level itself, which may be deemed unfair in reference to some notion of fair price or wage. Only this is really the unequal exchange problem. Another potential problem is the fact that the payment covers only a part of what is delivered, so that the protesters feel a lack of recognition through the payment system. This has to do with what has been called M-exploitation here.11

The wage labor relation is typically plagued by both issues. The level of wages may appear low due to monopsony configurations or due to the excessive amount of wealth poverty which makes too many workers desperate to earn a living. But workers also generally feel that the pressure to effort imposed on them by managers is meant to extract their ‘life substance’ with only a dubious promise of reward. This feeling is especially strong when, after efforts contributed to keep a firm afloat, layoffs fall on them with a conspicuous lack of humanity.

The exploitation of the hardworking by the lazy through the welfare state is not generally voiced by the same side of the political spectrum, but it is important to acknowledge that it too connects to the idea of exploitation as some form of unfair advantage. It often appears wise to retort that the exploiters in this case are rather badly off, whereas the hardworking are free to choose their ‘virtuous’ level of effort and do so for their own interest. Formally, however, the similarity between A-exploitation and D-exploitation shows that the problem may be real. It may happen that the lazy genuinely benefit from their situation as compared to one in which they would work the norm, whereas the hardworking genuinely pay a cost because their taxes are greater than they would be otherwise. A serious analysis of the well-being of the so-called lazy, however, is likely to reveal that most of them would in fact be better off if they could be brought to adopt the standard lifestyle. This is an empirical issue over which facts, not prejudice, should decide.

Historically, the Marxian analysis of exploitation has focused on unequal exchange in the hope of proving that labor is not fully paid. Roemer’s classical work has shown that this exclusive focus was not well justified, first because labor accounting is not compelling (not to mention the mistaken labor theory of value), second because A-exploitation, which he proposed as an alternative, is probably a more important issue. What has been argued in this paper is that all forms of exploitation, and there are at least four of them, deserve to be studied and identified in social organizations.12
An important analytical insight is that it is impossible to analyze exploitation without referring to norms. A-exploitation needs a norm of distribution of endowments. Unequal exchange needs a notion of fair price. D-exploitation needs a notion of fair contribution. It may seem that M-exploitation is an exception but it is not really one. As defined in this paper, it occurs when the buyer gains more from a marginal increase in trade than the seller. This involves an implicit 50% norm of sharing. One could adopt another norm.

The fact that norms underlie all notions of exploitation suggests that the concept of exploitation unavoidably depends on a more comprehensive conception of social justice. Even if one does not accept Roemer’s (1985) thesis that the concern for exploitation should be abandoned for the study of social justice, the key point that studying social justice is essential is incontrovertible. Exploitation is at most a part of the social justice problem.

Funding

The author(s) received no financial support for the research and/or authorship of this article.

Acknowledgements

This paper is written for the thirtieth anniversary of the publication of John Roemer’s General Theory of Exploitation and Class. It has benefitted from discussions with N Gravel, detailed comments by R Veneziani and two referees, and reactions from the audience at the ASSA conference in Chicago.

Notes

1. This is also the perspective defended in Yoshihara and Veneziani (2009).
2. Obviously, this assumes that the only source of unfairness in the pricing mechanism is the distribution of c. Otherwise, one should refer to an alternative, fair pricing mechanism.
3. This possibility is excluded from the models by the assumptions made, but it is easy to conceive variants of the models in which this could occur.
4. Here an increase in $e_j$ is the counterpart in the model of the increase in the slave’s capacity in the introductory paragraph. This is perhaps imperfect because one might want to disentangle labor hours and productivity, and restrict M-exploitation to productivity, viewed as an intrinsic characteristic of the worker. If $e_j = w_j l_j$, with $w_j$ a skill parameter and $l_j$ denoting labor hours, M-exploitation would occur if an increase in $w_j$ benefits $i$ more than $j$. But this is equivalent to the definition in terms of $e_j$ increase for the functional forms adopted in the current model.

Another complication may occur when $e_j = e_j^* + e_j^{**}$, where $e_j^*$ is an unpaid corvée whereas the product of $e_j^{**}$ goes to the worker. Then one has

$$
\varphi_i(c_N) = F(e_j) - p_i(e_j^{**}, c_i, c_j)
$$

and

$$
\varphi_j(c_N) = p_i(e_j^{**}, c_i, c_j) - U(e_j)
$$

and M-exploitation occurs when $e_j^*$ increases.

5. This range could be a single point (the current situation), in which case the notion of M-exploitation would be strictly local. There are many possibilities here.
6. The same can in fact occur with unequal exchange when \(i\) sells \(e_j\) to \(j\) at too low a price and 
\(j\) sells \(e_i\) to \(i\) at too low a price. Such a configuration might justify looking at the net balance, 
though, whereas nothing like a net balance seems available for M-exploitation.

7. Again, these norms of fairness are not to be viewed as global; they are relative to the given 
total \(E\) for the two individuals.

8. The \(F\) function is also a characteristic of the economy, but is not listed in \(E\) for simplicity, as 
it is not used as a variable.

9. For some parts of the analysis, it would be straightforward to introduce multi-dimensional 
\(c_i, e_i\). Norms of fairness would then be multi-dimensional as well (e.g. equal sharing in 
every dimension). Each dimension of \(c\) or \(e\) could be the object of a separate analysis of 
exploitation.

10. The model allows the individuals to use part of \(c_i\) and \(e_i\) for themselves. No matter how one 
defines the fair price, though, an individual cannot exploit himself according to the definition 
because the net payment (fair or actual) is zero.

11. The two issues are independent. The payment level can be correct even if only a part of the 
service is officially acknowledged (this part is then overpaid), and the payment can be unfair 
even if all dimensions of the service are considered.

12. In fact, Roemer himself admitted, in the debate raised by his book, that unequal exchange of 
labor should not be ignored and should be considered along exploitation in terms of unequal 
endowments (see Roemer, 1989a,b). I thank R Veneziani for drawing my attention to these 
references.

References


10: 283–345.


and Beyond. Cambridge: Cambridge University Press, pp. 159–186.

University Press.

14: 30–65.

Roemer JE (1989a) Second thoughts on property relations and exploitation. Canadian Journal of 
Philosophy 15(Suppl): 257–266.

18: 90–97.


Roemer JE and Silvestre J (1993) The proportional solution for economies with both private and 

Sample RJ (2003) Exploitation: What It Is and Why It’s Wrong. Lanham, MD: Rowman and 
Littlefield.