Review of Economic Studies (2011) **78**, 613–639 doi: 10.1093/restud/rdq026

Advance access publication 4 February 2011

Other-Regarding Preferences in General Equilibrium

MARTIN DUFWENBERG
University of Arizona and University of Gothenburg
PAUL HEIDHUES
European School of Management and Technology and CEPR
GEORG KIRCHSTEIGER
Université Libre de Bruxelles, ECORE, CEPR and CESifo
FRANK RIEDEL
Bielefeld University
and
JOEL SOBEL
University of California, San Diego

First version received September 2008*; final version accepted November* 2010 (*Eds.*)

We study competitive market outcomes in economies where agents have other-regarding preferences (ORPs). We identify a separability condition on monotone preferences that is necessary and sufficient for one's own demand to be independent of the allocations and characteristics of other agents in the economy. Given separability, it is impossible to identify ORPs from market behaviour: agents behave as if they had classical preferences that depend only on own consumption in competitive equilibrium. If preferences, in addition, depend only on the final allocation of consumption in society, the Second Welfare Theorem holds as long as any increase in resources can be distributed in a way that makes all agents better off. The First Welfare Theorem generally does not hold. Allowing agents to care about their own consumption and the distribution of consumption possibilities in the economy, the competitive equilibria are efficient given prices if and only if there is no Pareto-improving redistribution of income.

Key words: Markets, Other-regarding preferences, Self-interest, Welfare theorems

JEL Codes: D50, D62, D64

1. INTRODUCTION

The standard theory of competitive markets assumes that economic agents are selfish: they attempt to maximize their material well being ignoring the behaviour and opportunities of others. While self interest is an important human trait, (even) classical economists acknowledge that agents are not purely selfish.[1] Recently, the literature on other-regarding preferences (ORPs),

1. To quote Smith (1759, p. 1): "How selfish soever man may be supposed, there are evidently some principles in his nature, which interest him in the fortunes of others, and render their happiness necessary to him, though he derives nothing from it, except the pleasure of seeing it."

which documents and models how decision makers often fail to maximize their narrow self interest, has expanded rapidly.[2] In familiar bargaining and public-goods games, models of ORP yield strikingly different and more accurate predictions about play than standard theory. This paper investigates the extent to which the classic results of general-equilibrium theory hold true in economies with ORP-affected individuals.

We start with a general model where agents' preferences are defined over allocations and agents' opportunity sets instead of only their own consumption bundle. We investigate the hypothesis that individuals with ORP may behave as classical egoistic agents in competitive markets by asking under what conditions the demand function of ORP individuals is independent of the consumption and opportunity sets of other agents. We say that an agent behaves *as-if-classical* if the agent's demand function depends only on her income and prices. Under standard technical assumptions, we show in Section 2 that an as-if-classical demand function exists if and only if the preferences of an agent can be represented by a utility function that is separable between her own consumption bundle and the consumption vectors and opportunity sets of others. As separability is necessary and sufficient, we characterize completely the kind of ORP that do not affect market behaviour.

It is thus possible to compare the outcomes of a general-equilibrium model with ORP to those of a classical model in which each agent maximizes a utility function that depends only on her own consumption. If the separability condition holds, an agent's preferences induce preferences over the own consumption set that are independent of the consumption bundles of the other agents and of the distribution of budget sets. We refer to these preferences over own consumption as *internal preferences*.

Using the results of Section 2, we can relate any economy with separable ORPs to an economy with classical egoistic agents, whose preferences coincide with the internal preferences of the agents of the original economy. Price-taking agents with ORPs behave exactly like their classical counterparts. Consequently, as we observe in Section 3, the equilibria of the other-regarding economy coincide with those of the associated classical economy.

We next present our main results, which concern the extent to which the Fundamental Welfare Theorems extend to our framework. Walrasian equilibrium is efficient with respect to internal preferences, but it need not be efficient with respect to ORPs. To investigate the efficiency properties of equilibrium in more detail, we discuss the domain and structure of ORP. We distinguish two important classes of ORP in Section 4: *well-being externalities*, which can be modelled by utilities that depend on the allocations; and *opportunity-based externalities*, which allow preferences to depend on opportunity sets. While well-being externalities have been widely studied, the more general opportunity-based externalities allow one, *e.g.*, to also capture preference for equal opportunities.

Section 5.1 studies efficient allocations when well-being externalities are present. Efficient allocations need not be equilibria in this case. Indeed, we construct an exchange economy in which efficiency is incompatible with full resource utilization (total consumption equal to total endowment) even when internal preferences are strictly increasing. To rule out this kind of example, we assume that if the resources in the economy increase, then it is possible to make everyone better off. We prove that under this condition, all Pareto-efficient allocations are internally efficient and hence the Second Welfare Theorem holds.

In Section 5.2, we discuss the efficiency of equilibria when agents care only about the consumption opportunities of others. We study a condition we call the *Redistributional Loser Property*. The condition requires that a non-trivial redistribution of income in the population

2. Fehr and Gächter (2000) and Sobel (2005) survey the literature.

must leave someone worse off. The condition therefore places a limit on the importance of distributional concerns. When this condition holds, competitive equilibria satisfy an efficiency property. We show that the condition holds for natural generalizations of some prominent one-dimensional ORP models found in the literature.

Our analysis assumes that the classical general-equilibrium model describes market outcomes. It makes sense to ask whether this is appropriate for a model in which agents exhibit ORP. Section 6 provides conditions under which the core (suitably defined) of the economy is contained in the core of the economy defined by internal preferences. When the core is non-empty, this result provides a generalization of the classical core-equivalence theorem.

2. SEPARABILITY

This section introduces a basic model of competitive equilibrium that is classical except that consumers may have ORP. We then identify a separability condition necessary and sufficient to get a well-defined notion of preferences over own consumption.

Consider an economy with L goods indexed by $l = 1, \ldots, L$. Prices p are normalized such that $p_l \geq 0$ for all $l \in L$, and $\sum_{l=1}^{L} p_l = 1$.

There are J profit-maximizing firms. A typical firm $j \in \{1, \ldots, J\}$ is endowed with a production set $Y_j \subseteq \mathbb{R}^L$, with $y_j \in Y_j$ denoting the production plan implemented by firm j. As usual, negative components of y_j are inputs, and positive ones are outputs. Y_j is closed and bounded from above for all j.[3] The maximum attainable profit of firm j confronted with a price vector p is denoted by $\pi_j(p)$ and, since Y_j is closed and bounded from above, $\pi_j(p)$ exists for all p. Denote by $y = (y_1, \ldots, y_J)$ the implemented production profile and by $Y = \prod_{j=1}^{J} Y_j$ the set of all feasible production profiles.

There are I agents and the consumption set of a typical agent is assumed to be the non-negative orthant $\mathbb{R}^L_+$. The initial endowment of Agent i is denoted by e_i, and the bundle consumed by i, Agent i's *own consumption*, is $x_i = (x_{i1}, \ldots, x_{iL}) \in \mathbb{R}^L_+$. $x = (x_1, \ldots, x_I) \in \mathbb{R}^{L \times I}_+$ is the whole consumption profile, *i.e.*, the allocation of goods. Denote by $\bar{e}$ the aggregate initial endowment, $\sum_{i=1}^{I} e_i$. Firms are owned by the consumers and θ_{ij} denotes i's share of Firm j. The income of Agent i, w_i, is the sum of the value of i's initial endowment and the dividends she earns, $w_i = pe_i + \sum_{j=1}^{J} \theta_{ij} \pi_j(p)$.

Let $B = (B_1, \ldots, B_I)$ be a profile of budget sets, where each B_i is a non-empty compact subset of $\mathbb{R}^L_+$, and denote by $\mathcal{B}$ the set of all profiles of budget sets. Including budget sets in the domain of preferences permits us to describe situations where agents care for what others could have consumed rather than what others actually consume. Section 4 contains a more detailed discussion of the importance of this type of ORP.

To model general ORPs, we assume that each Agent i has a *preference relation* defined over allocations x and over profiles of budget sets B, which we denote by $\succeq_i$. We assume that the agents' preference relations are complete and transitive. To ensure that each agent's preference relation $\succeq_i$ can be represented by a utility function $U_i(x, B)$ defined on the set $\mathbb{R}^{L \times I}_+ \times \mathcal{B}$, we assume that $\succeq_i$ is continuous.[4] We also assume that Agent i's preferences are strictly convex over her own consumption—*i.e.*, for all B, x_{-i} and $x_i \neq x_i'$, $(x_i, x_{-i}, B) \succeq_i (x_i', x_{-i}, B)$ implies that $(\alpha x_i + (1-\alpha) x_i', x_{-i}, B) \succ_i (x_i', x_{-i}, B)$ for all $\alpha \in (0, 1)$. We do not require strict convexity over allocations, which would be far more stringent. For example, strict convexity over allocations would rule out an agent who is only interested in the consumption bundle she receives, because

3. We impose the boundedness assumption for simplicity. Our results go through for unbounded, convex production sets. To this end, one would use the typical compactification argument of general-equilibrium theory.

4. Endow $\mathcal{B}$ with the Hausdorff topology.

an appropriate change in the consumption bundle of a fellow agent would have to make her better off. Indeed, if Agent k is better offer with a convex combination of x_k and x'_k, then a jealous Agent i could prefer (x_i, x_{-i}, B) and (x_i, x'_{-i}, B) to $(x_i, \alpha x_{-i} + (1-\alpha)x'_{-i}, B)$.

We also assume strict monotonicity in own consumption, so that for all B, x_{-i}, $x_i \neq x'_i$ with x_i weakly larger than x'_i in all components we have $(x_i, x_{-i}, B) \succ_i (x'_i, x_{-i}, B)$. With ORP, strict monotonicity rules out, *e.g.,* that an agent wants to reduce her consumption because she feels bad whenever she is much better off than others.

An *economy* $\mathcal{E}$ is described by a tuple $(I, e, (U_i), J, Y, \theta)$ of agents, endowments, utility functions, firms, production sets and ownership shares.[5]

This paper asks whether agents with ORP behave differently from classical agents in perfectly competitive markets. To do so, we study demand behaviour. Since agents' preferences can be represented by a continuous utility function and the budget set is compact, the demand correspondence exists. Because we furthermore assume that an agent's preferences over her own consumption bundles are strictly convex, each agent i has a demand function given by

$$d_i(x_{-i}, B) = \arg\max_{x_i \in B_i} U_i(x, B).$$

Most of our results hold when the domain of the preferences includes budget-set profiles that consist of any non-empty, compact subsets of $\mathbb{R}^L_+$. For some of our results, however, we study budget-set profiles that are induced by a system of incomes and prices, *i.e.,* profiles consisting of sets B_i for which there exists a price $p \in P$ and an income $w_i > 0$ such that

$$B_i = \{x_i \in \mathbb{R}^L_+ : px_i \leq w_i\}. \tag{1}$$

For such budget sets, we write the demand function as $d_i(p, w_i, x_{-i}, B_{-i})$.

In general, the demand function depends on the consumption choice of other agents x_{-i} and the profile of consumption possibility sets of the others, B_{-i}. On the other hand, the demand function of an Agent i with classical preferences is independent of x_{-i} and of B_{-i}. This consideration leads to the following definition:

Definition 1. *Agent i behaves as-if-classical if $d_i(x_{-i}, B)$ is independent of x_{-i} and B_{-i}.*

Observe that even if the consumer's demand behaviour is independent of the budget sets and actions of other consumers, the behaviour of others generally does influence her level of utility. To see when agents behave as-if-classical, we take a closer look at preferences over own consumption. We say that an agent's preferences are separable if her relative evaluation of own consumption bundles is independent of the consumption of others and the profile of budget sets.

Definition 2. *Preferences $\succeq_i$ of Agent i are separable if for all allocations $x = (x_1, \ldots, x_I)$ and $x' = (x'_1, \ldots, x'_I)$ and all profiles of budget sets B and B' we have*

$$(x_i, x_{-i}, B) \succeq_i (x'_i, x_{-i}, B)$$

if and only if

$$(x_i, x'_{-i}, B') \succeq_i (x'_i, x'_{-i}, B').$$

Separable preferences can be represented by a utility function of the form $V_i(m_i(x_i), x_{-i}, B)$. Due to monotonicity in own consumption, V_i is strictly increasing in its first argument. Under

5. When we discuss exchange economies, we drop the reference to J, Y and θ.

our assumptions, $m_i : \mathbb{R}^L_+ \to \mathbb{R}$ can be taken to be a continuous, strictly monotone and strictly quasi-concave function. In this case, $m_i(x_i)$ describes Agent i's preferences when the consumption choices and opportunities of the other agents are fixed. We refer to the function $m_i(x_i)$ as a consumer's *internal utility function*. Loosely speaking, this function is a measure of the consumer's well-being absent any social comparisons.

It is intuitive that if an agent has a utility function that is separable in own consumption, then she would choose the same consumption bundle independent of the consumption and characteristics of others. The following theorem also establishes the converse—that if an agent behaves as-if-classical, then her preferences can be represented by a separable utility function—under the assumption of continuously differentiable demand.

Theorem 1. *1. If Agent i's preferences can be represented in the form*

$$V_i(m_i(x_i), x_{-i}, B)$$

for a strictly quasi-concave, continuous function $m_i : \mathbb{R}^L_+ \to \mathbb{R}$ *and a function* $V_i : D \subseteq \mathbb{R} \times \mathbb{R}^{(I-1)L} \times \mathcal{B} \to \mathbb{R}$ *that is increasing in its first argument, then Agent i behaves as-if-classical.*

2. Consider budget-set profiles induced by a system of incomes and prices. Suppose that Agent i's preferences are smooth enough that the demand function $d_i(p, w_i, x_{-i}, B_{-i})$ *is continuously differentiable*[6] *in* (p, w_i)*. If Agent i behaves as-if-classical, then her preferences can be represented in the form*

$$V_i(m_i(x_i), x_{-i}, B)$$

for a strictly quasi-concave continuous function $m_i : \mathbb{R}^L \to \mathbb{R}$ *and a function* $V_i : D \subseteq \mathbb{R} \times \mathbb{R}^{(I-1)L} \times \mathcal{B} \to \mathbb{R}$ *that is increasing in its first argument.*

The proofs of Theorem 1 and all subsequent results are in Appendix A.

The separability requirement is quite strong. The results in this section demonstrate, however, that they are the most general class of preferences that induces a consistent measure of individual utility independent of social comparisons. Furthermore, these preferences include classical preferences and some of the most prominent ORP models as special cases. In particular, if agents have preferences that can be represented by a weighted sum of internal utility functions, our separability assumption holds. Classical utilitarian preferences and the representation of Edgeworth (1881) satisfy the assumption. Recently introduced functional forms, *e.g.*, those of Charness and Rabin (2002) and Fehr and Schmidt (1999), presented to organize experimental results in bargaining and contracting environments also satisfy our separability assumption (see Section 4 below).[7]

3. EQUILIBRIUM EQUIVALENCE

In this section, we analyse the impact of ORP in a general-equilibrium environment. In order to do so, we must adjust the equilibrium definition. We also define a hypothetical economy in

6. A sufficient condition for this is that preferences are $\mathcal{C}^2$ in own consumption without critical points and that the bordered Hessian of U is non-zero at all x. See Mas-Colell (2001, Chapter 2), or Mas-Colell, Whinston and Green (1995, Chapter 3, Appendix).

7. Maccheroni, Marinacci and Rustichinic (2008) and Vostroknutov (2007) present models of ORP that do not satisfy our separability assumption, while Karni and Safra (2002) provide conditions for a separable representation in a context related to ours.

which all agents have corresponding internal preferences in order to investigate the implications of ORP for behaviour and welfare.

A *Walrasian equilibrium* consists of a price vector p^*, a feasible allocation x^*, a production plan y^* and a profile of budget sets B^* such that every firm maximizes its profits for given price p^*, each consumer i chooses her utility maximizing consumption bundle x_i^* for given profile of budget sets B^*, and the profile of budget sets B^* is compatible with p^* and y^*. That is, for all $i = 1, \ldots, I$, $j = 1, \ldots, J$, we have

$$p^* y_j^* \geq p^* y_j' \text{ for all } y_j' \in Y_j$$
$$x_i^* = \arg\max_{x_i \in B_i^*} U_i(x, B^*)$$
$$B_i^* = \left\{ x_i : p^* x_i \leq p^* e_i + \sum_{j=1}^{J} \theta_{ij} p^* y_j^* \right\}.$$

This definition of equilibrium implies that consumers are price takers and producers are profit maximizers. When agents have ORPs, the assumption of profit-maximizing firms is not as straightforward to justify as it is within standard general-equilibrium theory. To illustrate this, consider a firm that is owned by many small shareholders, who together own more than half of the shares, and one big shareholder, who owns the rest. In this case, the firm's profits might be important for the big shareholder's wealth, but negligible for the wealth of the other owners. If the small shareholders envy the big shareholder, a coalition of small shareholders might decide that the firm should not maximize its profits.[8] To exclude such a possibility and to justify profit maximization, we might restrict the analysis to situations where each firm is owned by a single agent.

To understand the role of ORPs, we will compare an economy $\mathcal{E} = (I, e, (U_i), J, Y, \theta)$ to its corresponding *internal economy* $\mathcal{E}^{int} = (I, e, (m_i), J, Y, \theta)$. In an internal economy, each firm has the same production set, and each consumer the same endowment, the same shares and the same internal preferences as in the original economy $\mathcal{E}$. In the internal economy, however, agents care only about their own direct consumption.

Having defined equilibrium and the internal economy, an immediate consequence of Theorem 1 is that

Theorem 2. *If all agents have separable preferences that are strictly monotone in own consumption, the set of Walrasian equilibria of an economy $\mathcal{E}$ coincides with the set of Walrasian equilibria of its corresponding internal economy $\mathcal{E}^{int}$.*[9]

If all agents' preferences are separable in their own consumption bundles and all agents prefer to spend their entire wealth, concerns such as envy, altruism or fairness do not influence market outcomes.[10]

8. The profit-maximizing assumption is also strong in other general-equilibrium contexts. See the discussion of Dierker and Grodal (1995) in the context of oligopolistic firms.

9. After we completed the paper, Rabah Amir pointed out that Dubey and Shubik (1985) contains a result similar to Theorem 2. Dubey and Shubik show that the Nash equilibria of a market game with a continuum of traders with separable, other-regarding preferences depend only on what we call the internal preferences.

10. Theorem 2 implies that if agents have monotone, separable preferences, then there exists no possible comparative static in a perfectly competitive-market setting that would distinguish between selfish and non-selfish preferences. More generally, one may ask under various assumptions on preferences what market data would identify non-selfish preferences. For example, if a ceteris-paribus redistribution of wealth among other members in society changes an agent's apostrophe consumption choice, we could infer that her preferences are not separable and thus also not purely selfish. Our theorem highlights, however, the converse finding in distinguishing between selfish and non-selfish preferences.

4. ORPs IN MULTI-GOOD CONTEXTS

The results in Sections 2 and 3 hold for preferences defined on general domains provided that internal utility can be separated from social concerns. For subsequent results, we limit attention to special classes of preferences that we describe in more detail in this section. We wish to emphasize how and why we include opportunities in utility functions, as we believe that this aspect of our model is central to the study of ORPs. We begin by discussing special cases of standard models of consumption externalities and then introduce budget sets into preferences.

A traditional way to model ORPs is to assume that Agent i's utility is a function of the internal utilities of other agents in the economy. Formally, a well-being externality arises if the utility of Agent i depends on x_i and the internal utility levels $m_k(x_k)$ of agents $k \neq i$. Hence, Agent i's preferences depend non-trivially on the x_{-i}, the consumption of other agents in the economy, but not on B, the set of opportunities. Well-being externalities are thus a subclass of consumption externalities. However, it is the (internal) well-being of your neighbour, independent of its source, that enters into utility.[11]

When there are well-being externalities, Agent i's preferences can be represented by a function $V_i(m_1, \ldots, m_I)$. We make the standard assumptions that $m_k(\cdot)$ is strictly increasing for each k and that $V_i(\cdot)$ is strictly increasing in m_i.

A leading example of well-being externalities is the example of Edgeworth (1881, p. 51) in which

$$V_i(m_1, \ldots, m_I) = m_i + \frac{\beta_i}{I-1}\left(\sum_{k \neq i} m_k\right). \tag{E WB}$$

In Equation (E WB), Agent i cares about his own internal utility and the sum of the utilities of the other agents. If $\beta_i > 0$, then Agent i is altruistic or benevolent. If $\beta_i < 0$ (a case that Edgeworth does not consider), then she is envious or spiteful.[12]

Well-being externalities provide a natural way to generalize existing one-dimensional models of ORPs tailored to allocations of money. Recent literature designed to organize experimental observations in games with monetary outcomes proposes alternative functional forms that can be interpreted as well-being externalities in multi-good settings. For example, the model of Fehr and Schmidt (1999) generalizes to

$$V_i(m_1, \ldots, m_I) =$$

$$m_i - \frac{\alpha_i}{I-1}\sum_k \max\{(m_k - m_i), 0\} - \frac{\beta_i}{I-1}\sum_k \max\{(m_i - m_k), 0\}, \tag{F-S WB}$$

with $\alpha_i \geq \beta_i \geq 0$ and $\beta_i < 1$. The first parameter assumption ensures that agents suffer more from being behind than from being ahead. In the context of a single good, $\beta_i < 1$ ensures that the utility function is monotonically increasing in one's internal utility. Similarly, a simple version of the model of Bolton and Ockenfels (2000) can be written

$$V_i(m_1, \ldots, m_I) = m_i - \beta_i \left| m_i - \frac{\sum_k m_k}{I} \right|, \tag{B-O WB}$$

11. If status is measured according to relative consumption of a particular good, then our model of well-being externalities does not include status concerns that could be captured in a general model of consumption externalities.

12. For other one-good models of envy, see Bolton (1991) and Kirchsteiger (1994); additional examples of one-good models of altruism are provided by Andreoni and Miller (2002) and by Cox and Sadiraj (2006).

where $0 \leq \beta_i < 1$. Finally, the preferences proposed in Charness and Rabin (2002)[13] are

$$V_i(m_1, \ldots, m_I) = m_i + \frac{\beta_i}{I-1}\left[\delta_i \min\{m_1, \ldots, m_I\} + (1-\delta_i)\sum_k m_k\right], \qquad \text{(C-R WB)}$$

where $\beta_i, \delta_i \geq 0$ and $\beta_i \delta_i < 1/(I-1)$. Intuitively, one may think of an agent as maximizing the combination of his own well-being and a given social welfare function. The functional form of Charness and Rabin can be viewed as extending Edgeworth's example (E WB) by adding a Rawlsian-type concern for the worst-off agent to the utility function.

Well-being externalities easily capture the preferences of agents who care about the level of (internal) utility of other agents in the economy. They provide a less compelling model of situations in which an agent's welfare depends on interpersonal comparisons. Consider an economy in which there are two agents, Adam and Eve. Imagine that Adam has ORPs so that he gains or loses utility depending on his position relative to Eve. Adam may, *e.g.*, be jealous of Eve whenever he deems her better off than himself but feel sorry for her when she is worse off. One can try to capture this situation as a well-being externality by assuming that Adam's total utility decreases when his internal utility is less than (some function of) Eve's internal utility or when his internal utility is greater than Eve's. We lack a theory that allows us to make interpersonal comparisons of internal utility, however. Consequently, we have no general way in which to identify when Adam should begin to envy Eve's internal utility.

An alternative approach is to assume that Adam envies Eve if he prefers (according to his internal preferences) her consumption to his.[14] This formulation can be described in models in which preferences depend only on (economy-wide) consumption bundles and does not require interpersonal comparisons of utility. On the other hand, what if Eve, due to differences in endowments, could choose bundles that Adam would love to have, but in fact chooses a bundle that Adam is not interested in at all.[15] For example, she may use her budget to buy apples, while Adam, who is allergic to apples, buys pears instead. If Adam envies Eve because if he had her budget he would have been able to buy more pears, then we must expand the domain of preferences to include these opportunities.[16]

More generally, individuals who desire equality of opportunity have preferences that depend on more than the final allocation of goods. A thought experiment contrasts well-being externalities from opportunity-based externalities. When there are well-being externalities, it is generally possible to change Agent i's utility by changing Agent k's internal preferences (holding allocations fixed). When there are opportunity-based externalities, Agent i's utility need not depend on the internal utility of other agents. On the other hand, when there are opportunity-based externalities, Agent i can be made better off if Agent k's choice set is changed even if the change does not influence Agent k's final allocation. Informally, an individual who prefers that all families can afford child care—whether they choose to use it or not—is consistent with opportunity-based externalities. A childless agent whose preferences exhibit well-being externalities benefits from

13. Charness and Rabin (2002) also include reciprocity concerns in their formulation.

14. For the case of purely selfish agents, this is defined as envy in Varian (1973). Varian (1976) mentions the possibility of envying the possibilities of another agent. Varian, however, does not consider ORP but investigates properties of allocations in a classical environment that are envy-free.

15. The decision-theoretic literature on menu-dependent preferences emphasizes the possibility that Agent i's preferences depend on her own opportunities. This is relevant in order to model preference for flexibility (Dekel, Lipman and Rustichini, 2001; Kreps, 1979), self-control problems (Gul and Pesendorfer, 2001), diversity (Nehring and Puppe, 2002) and freedom (Puppe, 1996).

16. On the other hand, if Adam is jealous because Eve's opportunity set translates into a high level of internal utility for her, this can be described as a well-being externality.

providing more affordable child care facilities only if doing so increases the number of children using the facilities.

For our welfare results, we separate social concerns derived from differences in opportunities from those that derive from concern about the well-being of other agents. In order to do this, we study utility functions that depend on own consumption and the economy-wide budget profile, but not directly on the consumption of other agents. To see the generality of this approach, suppose that Agent i evaluates an opportunity set of Agent k as being the value of the best element within this set. If, furthermore, Agent i selects and evaluates this best element according to Agent k's internal utility function, then well-being externalities can be viewed as a special case of the more general opportunity-based externalities.[17] Opportunity-based externalities are clearly far more general than this. For example (and as illustrated in Theorem 5 below) by assuming that Agent i evaluates the opportunity set of k using his own and not k's internal utility function one may develop opportunity-based versions of the models of Edgeworth, Bolton–Ockenfels, Fehr–Schmidt and Charness–Rabin.

5. WELFARE ANALYSIS

We next examine the extent to which the fundamental welfare theorems hold in our setting. Obviously, the First and the Second Welfare Theorem hold with respect to the internal utility functions. In order to refine our understanding of the welfare properties of equilibria with respect to the full ORP, we separate well-being externalities from opportunity-based externalities and restrict the way in which these externalities enter preferences.

5.1. *Well-being externalities and welfare*

As discussed in Section 4 preferences of Agent i are now represented by the function $V_i(m_1(x_1), \ldots, m_I(x_I))$, with $m_k(\cdot)$ strictly increasing for each k and $V_i(\cdot)$ strictly increasing in its ith argument. Since preferences depend only on the allocation, the usual efficiency definition can be used: An allocation x is called *feasible* if there is a production plan y with $y_j \in Y_j$ for all j and $\sum_{i=1}^{I} x_{il} \leq \bar{e}_l + \sum_{j=1}^{J} y_{jl}$ for all commodities $l = 1, \ldots, L$, and a feasible allocation x is *efficient* if there is no other feasible allocation x' that makes every consumer weakly better off in terms of utility and at least one strictly better off.

Standard references incorporate consumption externalities into general-equilibrium theory. Arrow and Hahn (1971) extend standard existence results. Examples (Edgeworth (1881, p. 51); Hochman and Rodgers (1969)) demonstrate that a Walrasian equilibrium need not be Pareto efficient even when agents have benevolent preferences.[18] More recently, Geanakoplos and Polemarchakis (2008) show (when preferences are separable) that equilibria with consumption externalities are generically inefficient and Noguchi and Zame (2006) observe that equilibria need not be efficient in the presence of consumption externalities. Gersbach and Haller (2001) (see also Gersbach and Haller (2009)) study pure-exchange economies in which the set of agents is partitioned into households. They prove versions of the First and Second Welfare Theorem when agents have ORPs that depend (separably) on the composition of their household but not on the consumption of other agents. The welfare theorems do not hold in their setting when there are consumption externalities. Kranich (1988) studies competitive equilibrium with ORPs that

17. There is a conceptual difference to the usual interpretation of the well-being externalities. If, *e.g.* an agent would not choose the optimal allocation within his budget set, this would not change the social comparison from an opportunity-based perception while it would do so from a well-being interpretation. Such suboptimal choice could arise if social comparisons lead to non-monotonicity in internal utility.

18. For the examples, the preferences can be of the form (E WB) with $\beta_i \in (0, 1)$ for all i.

are weakly increasing in consumption of all commodities of all agents.[19] He permits general non-separable preferences, but assumes that agents' utility is non-decreasing in the internal utility of other agents. He proves existence of equilibrium in a model in which agents can make bilateral transfers. He provides conditions under which equilibrium exists and shows by example that the First Welfare Theorem does not hold. At the end of this subsection, we reinforce this negative result by pointing out that the First Welfare Theorem fails even when we make quite strong assumptions on the nature of well-being externalities.

The literature contains conditions under which the Second Welfare Theorem generalizes. Winter (1969) extends the classical theorem to the case of separable ORP that are increasing in the internal utility of all agents. Borglin (1973) and Rader (1980) generalize the result to the class of separable ORPs that allow for both spitefulness and altruism. We adapt these results to our setting.

It is instructive to begin with an example that demonstrates that Pareto-efficient allocations need not be Walrasian equilibria.

Example 1. ***Hateful society:*** *Consider an exchange economy with two identical agents each with utility function* $V_i = m_i - 2m_k$, *where* $i \neq k$ *and* $m_i(x_i) = h(x_{i1}) + h(x_{i2})$ *for* $h(\cdot)$ *strictly increasing and strictly concave. Let the aggregate endowment be* $\bar{e} = (1, 1)$. *The allocation* $((0, 0), (0, 0))$, *which is obviously not internally efficient as none of the endowment is consumed, is Pareto efficient.*[20] *In this hateful society, it is impossible to make Agent 1 better off without making Agent 2 worse off. Hence, the set of Pareto-efficient allocations is not a subset of the internally efficient allocations and Walrasian equilibria need not be Pareto efficient.*

The preferences in the example exhibit a high degree of spitefulness. We next introduce a condition that rules out such pathological cases. The condition is satisfied by all specific models of ORP discussed above.

Social Monotonicity (SM): For any allocation x and $z \in \mathbb{R}^L_{++}$, there is a $(z_1, \ldots, z_I) \in \mathbb{R}^{L \times I}$ such that $z_i \geq 0$ for all i, $\sum_{i=1}^{I} z_i = z$, and for all i,

$$V_i(m_1(x_1 + z_1), \ldots, m_I(x_I + z_I)) > V_i(m_1(x_1), \ldots, m_I(x_I)).$$

The condition states that any increase in the resources available to the economy can be redistributed to make everyone better off. It is clear that SM fails in the above example with hateful agents. Under SM, Pareto-efficient allocations must be internally efficient. SM ensures that if an outcome is not internally efficient, then in the set of allocations in which all agents are internally better off, there exists an element in which the internal gains are divided between all agents in such a way that everyone is better off.

Theorem 3. *If SM holds, then the set of Pareto-efficient allocations is a subset of the set of internally efficient allocations.*

Benjamin (2008) proves a related result. Using our terminology, he shows that the set of efficient allocations is contained in the set of internally efficient allocations in a two-player contracting game. Benjamin assumes that players have internal payoffs and that one of the players

19. Kranich's paper is the only other paper we have seen that examines the welfare properties of competitive equilibria in a model that permits both well-being and opportunity-based externalities.

20. If the endowment cannot be destroyed, then the edge of the Edgeworth box is the set of Pareto-efficient allocations, while the diagonal is the set of internally efficient allocations.

has ORPs that are a function of the internal payoffs of the two players. It is straightforward to define the set of internally efficient payoffs and the set of efficient payoffs for this game (Benjamin refers to these sets as the materially Pareto efficient and utility Pareto-efficient sets, respectively). Benjamin proves that under an assumption that he calls joint monotonicity, the set of internally efficient payoffs contains the set of efficient payoffs. Joint monotonicity requires that in any neighbourhood of any pair of internal utilities (m_1, m_2), it is always possible to find two larger material utilities $(\hat{m}_1, \hat{m}_2)$ such that $V_i(\hat{m}_1, \hat{m}_2) > V_i(m_1, m_2)$ for $i = 1$ and 2. Assuming preferences are continuous and internal utility functions are strictly monotonic, joint monotonicity is equivalent to SM.[21]

An immediate consequence of Theorem 3 is the Second Welfare Theorem.

Corollary 1 (**Second Welfare Theorem).** *If SM holds, then every Pareto-efficient allocation can be achieved as a Walrasian equilibrium by using suitable lump-sum transfers.*

Corollary 1 guarantees that mandatory redistribution through lump-sum transfers does allow the economy to achieve efficiency. SM does not guarantee that an equilibrium is efficient, however. The next example highlights that there exist economies in which all equilibria are inefficient even when bilateral transfers are feasible.[22] In such economies, efficiency cannot be guaranteed without binding coordination among those willing to give.

Example 2. ***Inefficiency with Bilateral Transfers:*** *Consider an exchange economy with three agents and one good. Let the initial endowment be* $e = (1, 0, 1)$. *Let the utility of Agent 2 be given by* x_2, *i.e., assume that Agent 2 is selfish. Let the utility of Agent 1 be* $x_1 + (2/3)x_2$ *and the utility of Agent 3 be* $x_3 + (2/3)x_2$. *Then, independent of what Agent 3 gives, Agent 1 will never transfer any of the good to Agent 2. Similarly, Agent 3 will not want to transfer any of the good. This allocation, however, is Pareto dominated by the allocation* $(0, 2, 0)$.

5.2. *Opportunity-based externalities and welfare*

We now consider economies in which agents exhibit opportunity-based externalities. We assume in this section that the preferences of Agent i depend non-trivially on i's direct consumption, x_i, and on the budget sets of all agents, but they do not depend directly on the actual consumption of others, x_{-i}. In other words, the utility function can be written as $V_i(m_i(x_i), B)$ for a strictly quasi-concave, monotone and continuous function $m_i : \mathbb{R}^L_+ \to \mathbb{R}$ and a function $V_i : \mathbb{R} \times \mathcal{B} \to \mathbb{R}$ that is increasing in its first variable. In this subsection, we introduce a condition on preferences that is necessary and sufficient for equilibria to satisfy an efficiency property. The efficiency property, which we call efficiency relative to a price, is different from the standard notion of

21. Starting with any pair of internal utilities, (m_1, m_2), a strictly positive increase in available resources z can be reallocated to generate any $(\hat{m}_1, \hat{m}_2) >> (m_1, m_2)$ in a neighbourhood of (m_1, m_2) and hence, by joint monotonicity, can be used to increase the utility of both agents. This shows that joint monotonicity implies SM. To show that SM implies joint monotonicity, note that when $m_i(\cdot)$ and $V_i(\cdot)$ are continuous, we can pick z_i in the definition of SM to be strictly positive. Pick a point $m = (m_1(x_1), m_2(x_2))$. SM implies that for all $z \in \mathbb{R}^L_{++}$, there exists $z_1 + z_2 \leq z$, $z_i >> 0$, such that $V_i(m_1(x_1 + z_1), \ldots, m_I(x_I + z_I)) > V_i(m_1(x_1), \ldots, m_I(x_I))$. Furthermore, by strict monotonicity of $m_i(\cdot)$, $m_i(x_i + z_i) > m_i(x_i)$. By picking z small enough, $(m_1(x_1 + z_1), m_2(x_2 + z_2))$ can be taken to be in any given neighbourhood of m. Hence, joint monotonicity follows from SM.

22. Winter (1969) provides an example in which there exists both an efficient and an inefficient equilibrium with bilateral transfers. Goldman (1978) provides an example that demonstrates that competitive equilibria need not be efficient when benevolent agents can voluntarily exchange gifts. Kranich (1988) contains a related example that shows that competitive equilibria with transfers need not be efficient when agents are altruistic.

Pareto efficiency. We discuss its economic significance after formally introducing it below, but before doing so briefly highlight why the standard efficiency notion can fail.

An exact analog to the First Welfare Theorem is unavailable with opportunity-based externalities. Efficiency can fail for trivial reasons. For example, in an economy in which all but one agent have classical preferences and the remaining agent strictly prefers to limit the choices of others, one can improve upon an equilibrium allocation merely by requiring that agents can only choose their equilibrium allocation (rather than letting them choose from a budget set). We will rule out this type of example by requiring choice sets to be budget sets derived from a particular price endowment vector. Even with this restriction, equilibrium allocations will not necessarily be efficient. Consider an exchange economy with two equilibria (with different supporting prices) associated with the same initial endowment. These equilibria will not be Pareto ranked for the internal preferences but could be Pareto ranked when agents have ORP. For example, take a selfish Agent 1 and an altruistic Agent 2. If moving from Equilibrium 1 to Equilibrium 2 makes Agent 1 better off and decreases Agent 2's internal utility, it could still be that Agent 2's overall utility increases because the positive other-regarding effect dominates the negative effect on 2's internal utility. In this example, the two different equilibrium prices create different budget sets for agents even when initial endowments are fixed. Our constrained notion of efficiency avoids these types of examples by only permitting comparisons between budget-set profiles that are consistent with a fixed price vector. Due to the normalization of the price vector, each budget set B_i is consistent only with one particular price vector p and only with one particular income level w. Denote the price vector and the income level consistent with budget set B_i by $p(B_i)$ and $w(B_i)$, respectively. $w(B) = (w(B_1), \ldots, w(B_I))$ denotes the profile of incomes connected with the budget-set profile B. For the rest of this section, we only consider budget-set profiles that are consistent with one price vector, *i.e.*, budget-set profiles B for which $p(B_i) = p$ for all $i = 1, \ldots, I$.

Since budget-set profiles enter the domain of the preferences, we have to define the feasibility of budget sets profiles.

Definition 3. *Let $\mathcal{E}$ be an economy with opportunity-based externalities. The triple $(x, y, B) \in X \times Y \times \mathcal{B}$ is feasible for a price p, if and only if for all $i = 1, \ldots, I$, $j = 1, \ldots, J$, and $l = 1, \ldots, L$:*

i) $y_j \in Y_j$

ii) $\sum_{i=1}^{I} x_{il} \leq \sum_{i=1}^{I} e_{il} + \sum_{j=1}^{J} y_{jl}$

iii) $x_i \in B_i$

iv) $\sum_{i=1} w(B_i) = \sum_{i=1}^{I} pe_i + \sum_{j=1}^{J} py_j.$

In addition to the usual feasibility requirements on the production profile y and the consumption profile x, this feasibility notion also requires some consistency between x, y and B. In particular, each individual consumption bundle must be in the budget set of the respective consumer and that the profile of budget sets is feasible for the amount of income available in the economy.

A triple consisting of a production profile, a consumption profile and a budget-set profile is efficient if there is no other such triple that is also feasible for the same price and that makes all consumers weakly and some consumers strictly better off.

Definition 4. *In an economy $\mathcal{E}$ with opportunity-based externalities, a triple (x, y, B) is efficient with respect to a price vector p if and only if*

1. *(x, y, B) is feasible for p;*

2. *there does not exist another triple (x', y', B'), which is feasible for p and for which*

$$V_i(m_i(x_i'), B') \geq V_i(m_i(x_i), B) \text{ for all } i, \text{ and}$$

$$V_i(m_i(x_i'), B') > V_i(m_i(x_i), B) \text{ for at least one } i.$$

Since prices determine budget sets and budget sets enter preferences, prices enter the efficiency definition. In order for an outcome to be efficient relative to a price, it must be impossible to make everyone better off with another feasible allocation and a new configuration of opportunities, provided that the opportunities are consistent with the given price.[23]

Efficiency with respect to a price is an especially attractive concept if the internal economy is quasi-linear ($m_i(x_i) = h_i(x_{i2}, \ldots, x_{iL}) + x_{i1}$) with strict convexity of preferences in own consumption. In this case, the equilibrium price is unique. Hence, any allocation that is efficient with respect to the equilibrium price is also efficient with respect to redistribution followed by market exchange. This constrained notion of efficiency highlights an important property of a competitive-market equilibrium in more general environments. Consider, *e.g.*, whether we would expect social groups to set up a redistributive mechanisms among themselves. More specifically, we envision a large market economy in which "small" groups can form freely and redistribute endowments among themselves, but in which each agent can thereafter exchange her goods at the anonymous market place. Reminiscent of the small-country assumption in the international trade literature, groups are small in the sense that they cannot affect market prices or do not take their effect on market prices into account. Efficiency given prices implies that no social group can find a redistribution among its members that makes all its members better off.

The following property is crucial for the efficiency of the equilibrium allocation with respect to equilibrium prices.

Redistributional Loser Property (RLP): RLP holds at a budget set profile B if for any other profile of budget sets $B' \neq B$ for which there exists a p such that $p(B_i) = p(B_i') = p$ for all i and $\sum_{i=1}^{I} w(B_i) \geq \sum_{i=1}^{I} w(B_i')$

$$V_k(m_k(d_k(B_k)), B) \leq V_k(m_k(d_k(B_k')), B') \text{ for all } k \implies$$
$$V_k(m_k(d_k(B_k)), B) = V_k(m_k(d_k(B_k')), B') \text{ for all } k. \tag{2}$$

RLP holds if implication (2) holds at all B.

Notice that condition (2) holds if there always exists an Agent r for whom

$$V_r(m_r(d_r(B_r)), B) > V_r(m_r(d_r(B_r')), B'). \tag{3}$$

If inequality (3) holds, then Agent r loses when budget sets change from B to B'. That is, RLP requires that a non-trivial redistribution of income in the population must leave someone worse

23. Other notions of constrained efficiency appear in models of general equilibrium with externalities. Arrow and Hahn (1971, Chapter 6) define conditionally efficiency relative to a price vector in an economy in which prices enter utility functions. Their notion imposes even more constraints than ours, requiring $B' = B$ in Definition 4 (Ellickson (1993, Chapter 7.3) offers a related definition).

off. This is an extremely strong restriction in situations in which agents have ORPs. In particular, it rules out the possibility that a charitable transfer can be beneficial to both the recipient and the donor and not harmful to anyone else. On the other hand, it is precisely the condition needed to describe when equilibria are efficient.

Theorem 4. *The equilibrium outcome (x^*, y^*, B^*) of an economy with opportunity–based externalities is efficient with respect to the equilibrium price vector p^* if preferences satisfy RLP at B^*.*

RLP is thus sufficient for efficiency of competitive markets. Conversely, if the equilibrium outcome (x, y, B) of an economy with opportunity-based externalities is efficient with respect to the equilibrium price vector p, then, by the definition of efficiency, condition (2) holds for all triples (x', y', B') that are feasible with respect to p.

While Theorem 4 implies that the internal economy shares some efficiency properties with a family of economies with ORP, one cannot look only at internal preferences to analyse the distributional impact of changes in opportunities even if RLP holds. For example, a change in the wealth profile from B to B' need not have the same impact on Agent i as it would have on her counterpart i^{int} in the corresponding internal economy: A change beneficial for i^{int} (*i.e.*, a change with $B_i \subseteq B_i'$) may hurt i, and a change beneficial for i might may hurt i^{int}. The theorem only states that all equilibrium outcomes are efficient in the economy with distributional concerns as they are for the corresponding internal economy, provided that RLP holds and that the prices inducing B and B' are the same.

Since RLP implies efficiency, one wonders when RLP holds. Obviously, RLP holds when a consumer has classical preferences. We know that it fails when sufficiently altruistic consumers would want to make bilateral transfers to others. But RLP also holds for prominent specifications of preferences, especially in large economies. To illustrate this point, we will analyse opportunity-based externalities that can be represented by versions of the utility functions (F-S WB), (B-O WB) and (C-R WB) adapted to opportunity-based externalities.

Let $\tilde{m}_i(B_k)$ be the internal utility of Agent i if she could select the item from Agent k's budget set that she prefers most (according to the internal utility function m_i). That is, if B_k represents Agent k's budget set,

$$\tilde{m}_i(B_k) = \max m_i(x_k) \text{ subject to } x_k \in B_k. \tag{4}$$

To simplify notation, we write $\tilde{m}_{ki}$ in place of $\tilde{m}_i(B_k)$. This specification provides a framework in which agents make social comparisons based on the well being that they could derive from the opportunity sets of others. It permits a natural extension of functional forms commonly used to model ORP in environments with monetary payoffs to multi-good environments.

Fix a price vector p. Given $w_i > 0$, let $v_i(w_i) = m_i(d_i(p, w_i))$ be the indirect utility determined by internal preferences. For the remainder of the section, we assume that $v_i(\cdot)$ is differentiable (since $v_i(\cdot)$ is increasing, it will be differentiable almost everywhere) and we denote the derivative of $v_i(\cdot)$ with respect to wealth by $v_i'(\cdot)$.

Given a profile of budget sets B and associated wealth profile $w_i(B)$ for each i, let $\bar{v}_i = \max_{w \leq \sum_k w_k} v_i'(w)$, $\underline{v}_i = \min_{w \leq \sum_k w_k} v_i'(w)$, and $b_i = \underline{v}_i / \bar{v}_i$. The next result shows that RLP holds when the utility function takes on one of several functional forms that generalize utility functions used in one-good models of ORP.

Theorem 5. *Let B be a profile of budget sets and $w(B)$ be the associated wealth profile. RLP holds at B whenever the utility function of Agent i takes one of the following forms:*

1.

$$V_i(m_i(x_i), B) =$$

$$m_i(x_i) - \frac{\alpha_i}{I-1}\sum_k \max\{\tilde{m}_{ki} - \tilde{m}_{ii}, 0\} - \frac{\beta_i}{I-1}\sum_k \max\{\tilde{m}_{ii} - \tilde{m}_{ki}, 0\} \qquad \text{(F-S OB)}$$

with $\alpha_i \geq \beta_i \geq 0$, $\beta_i < b_i$ *and* I *large enough;*

2.

$$V_i(m_i(x_i), B) = m_i(x_i) - \beta_i \left| \tilde{m}_{ii} - \frac{\sum_k \tilde{m}_{ki}}{I} \right| \qquad \text{(B-O OB)}$$

with $0 \leq \beta_i < b_i$*;*

3.

$$V_i(m_i(x_i), B)$$

$$m_i(x_i) + \frac{\beta_i}{I-1}\left[\delta_i \min\{\tilde{m}_{1i}, \ldots, \tilde{m}_{Ii}\} + (1-\delta_i)\sum_k \tilde{m}_{ki}\right] \qquad \text{(C-R OB)}$$

with

$$\frac{b_i}{(1-\delta_i)(b_i-1)+\delta_i} > \beta_i > -b_i \text{ and } I \text{ large enough.}$$

The proof of Theorem 5 demonstrates that for each of the functional forms inequality (3) holds. Furthermore, the Agent r in inequality (3) can always be taken to be one of the agents who loses most from the redistribution of income induced by the changes in opportunity sets (*i.e.*, $w(B_r) - w(B_r') = \max_i\{w(B_i) - w(B_i')\}$).

The functional forms in Theorem 5 are generalizations of standard functional forms adapted to opportunity-based externalities. Expression (F-S OB) is an analog to equation (F-S WB); expression (B-O OB) modifies equation (B-O WB) and expression (C-R OB) modifies equation (C-R WB). We conclude that Theorem 5 demonstrates that RLP is satisfied in a wide range of functional forms found in the literature, including standard preferences exhibiting the possibility of both altruism and spite.

The parameter restrictions in Theorem 5 depend in an intuitive way on the variability of the marginal utility of income. Consider the special case in which $v_r(\cdot)$ is linear. This would be the case if agents had quasi-linear internal utility functions. RLP is most likely to hold in this case since the internal cost of a transfer does not depend on the level of wealth. In this case, $\bar{v}_r = \underline{v}_r$ and so $b_r = 1$ and the ranges for β_r in the theorem agree with those found in the one-dimensional version of the models (designed for risk-neutral agents). On the other hand, when the marginal utility of income is variable, the parameter restrictions in Theorem 5 become more stringent, leaving only the self-regarding versions of the functional forms in the limit $b_r = 0$. Example 3, below, confirms that the conclusions of Theorem 5 need not hold when the marginal internal utility of income approaches infinity.

A common feature of the preferences in Theorem 5 is that, in large economies, the opportunities of a particular other agent have a small impact on the utility of a decision maker. This assumption seems appropriate when agents have preferences that take into account the opportunity sets of all other agents symmetrically. As a consequence of this assumption, the parameter restrictions sufficient for RLP are weaker in equations (F-S OB) and (C-R OB) as the economy grows larger. For these functional forms, the power of an individual to make a meaningful change to the distribution of income of the economy decreases as the economy grows large.

RLP limits the extent to which an agent can take the opportunities of others into account. It strikes us as incompatible with much real-world charity. A rich agent with low marginal utility of

income sacrifices little internal utility when she makes a transfer to a poor agent. RLP assumes that this transfer is unattractive. The following example makes this point concretely. It further demonstrates that when the marginal internal utility of income is unbounded, there may be scope for efficiency enhancing redistribution-even if the economy is large.

Example 3. *Inefficiency and no RLP: We consider a one-good exchange economy with two groups of size $n \geq 2$, rich and poor agents. Poor agents are selfish, while rich agents are altruistic. All agents have the same internal utility function m. We assume that utility takes the form: $m(x_i) = x_i^{\alpha}$ for $\alpha \in (0,1)$.*

The preferences of any rich agent $i = 1, \ldots, n$ are

$$V_i(m_i(x_i), B) = m_i(x_i) + \frac{\beta}{I-1} \sum_{k \neq i} \tilde{m}_{ki}$$

for some $\beta > 0$ and $\tilde{m}_{ki}$ described following equation (4). Let aggregate endowment be $\bar{e} = n(1+\eta)$ and $\eta > 0$ be sufficiently small (specified exactly in Appendix A).

The initial endowment is given by $e_i = 1$ for the rich agents $i = 1, \ldots, n$ and $e_i = \eta$ for the poor agents $i = n+1, \ldots, 2n$. The allocation e is internally efficient and, hence, the unique Walrasian equilibrium of an economy where e is the initial endowment. The corresponding equilibrium price p^* is 1, and the corresponding income levels are 1 for $i = 1, \ldots, n$, and η for $i = n+1, \ldots, 2n$. The corresponding profile of budget sets is denoted by B^*.

We now construct another tuple (x', B') that is feasible for price p^* and dominates (e, B^*). The idea is that every rich agent gives $\epsilon > 0$ of his income to some poor agent, so that income levels are given by $1-\varepsilon$ for $i = 1, \ldots, n$, and by $\eta+\varepsilon$ for $i = n+1, \ldots, 2n$. Denote the corresponding profile of budget sets by B'. The redistribution of incomes lead to optimal consumption bundles of $x_i' = 1-\varepsilon$ for $i = 1, \ldots, n$ and $x_i' = \eta+\varepsilon$. In Appendix A, we show that (x', B') dominates (e, B^*) for all n. Hence, for all n, the Walrasian equilibrium is inefficient with respect to the price vector p.

6. CORE EQUIVALENCE

Throughout the paper, we have assumed the classical general-equilibrium model describes market outcomes. Is competitive equilibrium the appropriate way to model behaviour of agents with ORPs? This question is valid in the standard models, although it perhaps has greater force in our setting with ORP because market outcomes need not be efficient. There are two possible approaches to this problem. In an earlier version of the paper, we use arguments of Roberts and Postlewaite (1976) to show that, as in classical economies, price taking is approximately optimal in large economies.

In this section, we examine the classical core-equivalence theorem, which asserts that the set of core allocations shrinks to the set of competitive equilibria as the number of agents grows. Specifically, we consider the Debreu–Scarf thought experiment in which agents' internal preferences are replicated. We show that a generalization of the SM condition introduced in Section 5.1 implies that the core is contained in the internal core. We focus on well-being externalities because it is unclear how to extend the opportunity-based preferences we have introduced above to cases in which a coalition of agents jointly determines the use of a given set of resources.

Our results depend on an extension of the SM condition that we introduced in Section 5.1. We require that any subgroup of agents can find a way to distribute extra endowments among themselves in such a way that every member of the subgroup is better off. Under this *GSM* assumption, the core of the original economy is a subset of the core of the internal economy. In particular, we get the equal treatment property: agents with the same internal preferences

and endowments get the same consumption bundle in every core allocation. The Debreu–Scarf theorem then implies that the core of the limit economy is a subset of the set of Walrasian equilibria. We do not get full core equivalence in general because the core can be empty. We then give a simple sufficient condition for a non-empty core.

The first conceptual problem that we encounter is that of defining the core. An allocation belongs to the core if it can be viewed as the outcome of cooperation among agents. Classically, x is in the core if there is no coalition $C \subseteq I$ that can improve upon or block x. *C improves upon x* if there is a C-allocation $x' = (x'_i)_{i\in C}$ such that x' is feasible when the coalition C is autarkic and every member of C prefers x' to x. Feasibility is easy to define formally: x' is C-feasible if $\sum_{i\in C}(x'_i - e_i) \leq 0$. Making precise the requirement that every agent in C prefers x' to x is more subtle: given that preferences depend on others' consumption choices, how should we evaluate the actions of agents outside of a coalition once the coalition forms? This problem does not arise in the classical case when Agent i's preferences depend only on x_i, but it raises important issues in our context.

We focus on a notion of core in which improvements are relatively easy for coalitions to find.

Definition 5. *A coalition $C \subseteq \{1,\ldots,I\}$ can improve upon an allocation x if there exists a C-feasible allocation x' such that*

$$U_i((x'_k)_{k\in C}, (x_k)_{k\notin C}) > U_i(x) \quad \textit{for all } i \in C.$$

A feasible allocation x is in the core if there is no coalition C that can improve upon x.

This definition is, in spirit, a generalization of the test for deviations in Nash equilibrium: holding the allocations of the other agents fixed, a coalition can improve itself if it is able to reallocate its resources in a way that makes all members of the coalition better off. The strong Nash equilibrium concept by Aumann (1959), defined for non-cooperative games, makes the same assumption about the behaviour of non-coalition members.

Our definition of the core is a natural generalization of the definition of competitive behaviour. In competitive equilibrium, individuals assume that opponents do not change their consumption when they consider deviating from their equilibrium consumption. In our definition of the core, coalitions maintain a similar assumption about the complementary coalition. In both cases, when an agent decides whether to make a demand different from that specified, he does not take into account that markets will not clear (making the actions of the rest of the economy infeasible). With this notion of stability, we can generalize the results in Section 5.1.

In general, we cannot expect to have equality of core and equilibria even in the continuum limit because we know that Walrasian equilibria can be inefficient. On the other hand, we have shown that SM implies a version of the Second Welfare Theorem. It thus seems plausible that a suitable strengthening of the SM condition yields that the core of large economies is a subset of the set of Walrasian equilibria.

Group Social Monotonicity (GSM) Let $C \subseteq I$ be a coalition. For any allocation x and $z \in \mathbb{R}^L_{++}$, there is a redistribution $(z_j)_{j\in C} \geq 0$ with $\sum_{j\in C} z_j = z$ such that the members of C prefer

$$y_j = \begin{cases} x_j + z_j, & j \in C, \\ x_j, & j \notin C, \end{cases}$$

to x, *i.e.*,

$$U_i(y) > U_i(x) \quad (i \in C).$$

Lemma 1. *Under GSM, the core is a subset of the internal core.*

Lemma 1 is a generalization of Theorem 3. If an outcome x is not in the internal core, then there exists a coalition that can achieve a higher level of internal utility for its members using only the resources of its members. Hence, every member of this coalition can get the same internal utility as in the allocation x even if we reduce the resources available to the coalition by a small amount z. GSM guarantees that there is a way to redistribute z to coalition members in a way that makes every member of the coalition strictly better off, which implies the conclusion of Lemma 1.

We now perform the Debreu–Scarf thought experiment by replicating an economy many times. Note that replication is not a trivial task with ORPs. Suppose that Adam is altruistic and benefits from Eve's well-being in a two-person economy. Now replicate them. How does Adam feel about Eve 1 and/or Eve 2? There are several more or less natural choices to formalize Adam's preferences in the replicated economy. Somewhat fortunately, our results do not depend on the way the replicated Adams care about the replicated Eves' consumption choices. Let us start with I agents with separable preferences $V_i(m_i(x_i), x_{-i})$. The n-replica of the economy $\mathcal{E}_n$ has nI agents. Let us denote by $x_{i,n}$ the consumption choice of the nth copy of agent i, $n = 1, \ldots, N$. We suppose that preferences of Agent i, n can be represented by $V_{i,n}(m_i(x_{i,n}), x_{-(i,n)})$ that is monotone in $m_i(x_{i,n})$ for all $x_{-(i,n)}$. Note that all copies of Agent i have the same internal utility function. We leave the way the utility of Agent i, n depends on others' consumption choices completely general. Lemma 1 tells us that the core of the Nth replica is a subset of the core of the internal Nth replica economy. As a consequence of the classical equal-treatment lemma, core allocations treat all agents of type i equally. By the theorem of Debreu and Scarf (1963), the internal core shrinks to the set of Walrasian equilibria as n grows large. Let $\mathcal{C}_N$ denote the core of $\mathcal{E}_N$ and let $WE(\mathcal{E})$ be the set of Walrasian equilibria of an economy $\mathcal{E}$. We thus get

Theorem 6. *Under GSM*

$$\bigcap_{N \in \mathbb{N}} \mathcal{C}_N \subseteq WE(\mathcal{E}).$$

As we remarked above, one cannot get equality of the limit core and the set of Walrasian equilibria, as these equilibria can be inefficient in general and, in this case, the grand coalition could improve. Indeed, while we do have existence of Walrasian equilibria with separable preferences, the core may be empty.

Example 4. *Let there be three agents and one consumption good. Let the internal utility functions of all agents be linear. There are three units of the consumption good available and individuals each have unit endowment. The utility functions are* $U_1 = m_1 + 2m_2$, $U_2 = m_2 + 2m_3$, $U_3 = m_3 + 2m_1$. *No agent wants to destroy any of the endowment. Thus, we can restrict attention to allocations* $(a_1, a_2, 3 - a_1 - a_2)$. *The allocation* $(1, 1, 1)$ *is efficient.*

Observe that any allocation in which $a_1 > 0$ *is blocked by the coalition* $C = \{1, 2\}$, *which prefers to allocate the good to Agent 2. Any allocation in which* $a_2 > 0$ *is blocked by* $C = \{2, 3\}$ *and any allocation in which* $3 - a_1 - a_2 > 0$ *is blocked by* $C = \{3, 1\}$. *Thus, the core is empty.*

In the preceding example, an outcome fails to be in the core because an agent gains from making a unilateral transfer. This kind of altruism, plus a disagreement across agents about who should receive transfers, destroys the core. The next result provides a sufficient condition for the non-emptiness of the core. The condition requires that any coalition that improves utility of its members must also improve the internal utility of its members. In particular, no agent would

gain from making a unilateral transfer. This condition implies that any allocation in the internal core is in the core. Since the internal core is non-empty, the core of the economy is non-empty.

Theorem 7. *Let x be an internal core allocation. Assume that no coalition $C \subseteq \{1, \ldots, I\}$ can find a C-feasible allocation x' in which $m_i(x'_i) < m_i(x_i)$ and $U_i(x') > U_i(x)$ for some i. Then, x belongs to the core of the original economy.*

Appendix B reviews alternative definitions of the core for games with externalities.

7. CONCLUSION

We have shown that under standard technical assumptions, ORPs induce consistent preferences over own consumption if and only if the ORPs satisfy a separability condition. In this case, associated with any economy, there is an economy in which agents have classical preferences. When the separability condition holds, equilibria in economies with ORPs coincide with those in the associated classical economy. Hence, agents who care directly about the welfare and opportunities of others cannot be distinguished from selfish agents in market settings. Sobel (2010) establishes related results in a simple trading environment in which players have market power. He assumes that agents have ORPs and identifies necessary and sufficient conditions on these preferences under which market equilibria will be identical to competitive equilibria. When agents have market power, separability is not sufficient for this result. In addition, he shows that as the economy grows, market equilibria are approximately competitive under weak assumptions on preferences.

The fact that market behaviour may not be affected by ORPs does not mean that we can ignore the existence of ORP in markets. First, market outcomes need not be efficient. Second, even when market equilibria are efficient—and we have given conditions that imply a form of efficiency—the agents who gain and lose from interventions will depend on the precise nature of preferences.

The paper makes several contributions. From a technical point of view, we demonstrate that some classic results hold under more general assumptions about preferences. We contribute to behavioural economics by identifying aspects of classical general-equilibrium theory that are robust to relaxing the empirically questionable assumptions of purely selfish preferences. We give some guidance to welfare economists interested in the performance of markets in which agents have ORP. Finally, we describe an identification problem which cautions empiricists who observe classical competitive behaviour in markets from concluding that agents have classical selfish preferences.

Acknowledgement. This paper combines and extends portions of three independent papers written by subsets of the authors: "Do Social Preferences Matter in Competitive Markets?" by Heidhues and Riedel, "Classical Market Outcomes with Non-Classical Preferences" by Kirchsteiger and Dufwenberg and "Do Markets Make People Selfish?" by Sobel. We thank Geir Asheim, Robert Aumann, Dan Benjamin, Ted Bergstrom, Antoni Calvó-Armengol, Miguel Costa-Gomes, Vincent Crawford, Egbert Dierker, Armin Falk, Victor Ginsburg, Claus–Jochen Haake, Martin Hellwig, Matthew Jackson, James Jordan, Botond Kőszegi, John List, Mark Machina, Manfred Nermuth, Clemens Puppe, John Quah, Joachim Rosenmüller, Herbert Scarf, Klaus Schmidt, Uzi Segal, Thomas Sterner, Jean-Marc Tallon, Walter Trockel, Hal Varian, Bill Zame, Klaus Zauner and referees for helpful comments. We also acknowledge comments from seminar audiences at the Autonomous University of Barcelona, University of California, Berkeley, Karlsruhe, Ludwig-Maximilians-Universität Munich, Nottingham, Paris I, Penn State, University of Southern California, the 2008 European Economic Association meetings in Milan, the General Equilibrium Workshop in Rio 2008, the 2008 Southwestern Economic Theory conference at University of California, Santa Barbara and the 2007 Stony Brook Gale Feast for comments. Heidhues thanks the Deutsche Forschungsgemeinschaft for financial support through SFB/TR 15. Kirchsteiger thanks the Banque Nationale de Belgique for financial support. Riedel thanks the Deutsche Forschungsgemeinschaft for financial support through grant Ri–1128–3–1 and Universities Paris I and IX for their hospitality and support. Sobel thanks the

Guggenheim Foundation, National Science Foundation, and the Secretaría de Estado de Universidades e Investigacíon del Ministerio de Educación y Ciencia (Spain) for financial support and is grateful to the Departament d'Economia i d'Història Econòmica and Institut d'Anàlisi Econòmica of the Universitat Autònoma de Barcelona for hospitality and administrative support.

APPENDIX A

Proof of Theorem 1. We prove Part (1) first. As m_i is continuous and strictly quasi-concave, the standard utility maximization problem

$$\max_{x_i \geq 0, px_i \leq w_i} m_i(x_i)$$

has a unique solution, which we denote, in a slight abuse of notation, $d_i(p, w_i)$ for $p \gg 0$ and $w_i > 0$. This demand function does not depend on x_{-i} or B_{-i}. Now take any x_{-i} and B. We have for all budget-feasible x_i

$$m_i(x_i) < m_i(d_i(p, w_i))$$

whenever $x_i \neq d_i(p, w_i)$. As $V_i(m, x_{-i}, B)$ is increasing in m, it follows that

$$V_i(m_i(x_i), x_{-i}, B) < V_i(m_i(d_i(p, w_i)), x_{-i}, B)$$

whenever $x_i \neq d_i(p, w_i)$. Thus, $d_i(p, w_i)$ also uniquely maximizes utility for Agent i. In particular, her demand function is independent of x_{-i}; in other words, she behaves as if selfish.

Now consider Part (2). Let $d_i(p, w_i)$ be the demand function of Agent i which, by assumption, does not depend on x_{-i} and B_{-i}. In a first step, we construct an internal utility function on the consumption set $\mathbb{R}^L_+$ of Agent i. This is a standard integrability problem. Such a function $m_i(x_i)$ exists if d_i is continuously differentiable, homogeneous of degree zero, d_i has a symmetric and negative semi-definite Slutsky substitution matrix and d_i satisfies Walras's law: $pd_i(p, w_i) = w$ for all $p \gg 0$ and $w_i > 0$.

By assumption, d_i is continuously differentiable. As demand d_i is derived from utility maximization (albeit with the additional parameters x_{-i} and B), homogeneity of degree zero and negative semi-definiteness of the substitution matrix hold true as well. Walras's law follows from monotonicity. We can then apply the integrability theorem of Hurwitz and Uzawa (1971) to obtain a utility function $m_i(x_i)$ that rationalizes x_i. In particular, we have for all x_{-i} that

$$U_i(x_i, x_{-i}, B) \geq U_i(z_i, x_{-i}, B) \Leftrightarrow m_i(x_i) \geq m_i(z_i) \quad (x_i, z_i \in \mathbb{R}^L_+). \tag{A.1}$$

We can thus define a function $V_i(\mu, x_{-i}, B)$ on the image of m and $\mathbb{R}^{(I-1)L}_+$ by setting

$$V_i(\mu, x_{-i}, B) = U_i(x_i, x_{-i}, B)$$

for some x_i with $m_i(x_i) = \mu$. This definition does not depend on the particular x_i chosen as we have $U(x_i, x_{-i}, B) = U(z_i, x_{-i}, B)$ for all x_i, z_i with $m_i(x_i) = m_i(z_i)$ by condition (A.1).

Finally, we must show that V_i is increasing in μ. Let $\mu > \nu$ for two numbers μ, ν in the image of m_i. Choose x_i, z_i with $\mu = m_i(x_i)$ and $\nu = m_i(z_i)$. We then get from $m_i(x_i) > m_i(z_i)$ and condition (A.1) that

$$U_i(x_i, x_{-i}, B) > U_i(z_i, x_{-i}, B).$$

By definition of V_i, this is equivalent to

$$V_i(\mu, x_{-i}, B) > V_i(\nu, x_{-i}, B).$$

Thus, V_i is increasing in its first variable. ‖

Proof of Theorem 3. Assume that there exists a Pareto-efficient allocation x that is not internally efficient. Hence, there exists a feasible allocation x' such that $m_i(x'_i) > m_i(x_i)$ for all i. It follows from monotonicity that there exists $\tilde{x}'$

with $\tilde{x}'_{il} < x'_{il}$ for all $i = 1, \ldots, I$ and $l = 1, \ldots L$ such that $m_i(\tilde{x}'_i) = m_i(x_i)$ for all i. The SM condition guarantees that it is possible to make all agents better off by some distribution of $x' - \tilde{x}'$. ‖

Proof of Corollary 1. The result follows because every internally efficient allocation can be implemented under SM (Theorem 2) and the set of Pareto-efficient allocations is a subset of the set of internally efficient payoffs. ‖

Proof of Theorem 4. Since in equilibrium, each agent i chooses a utility maximizing consumption bundle in B_i^*, $V_i(d_i(B_i^*), B^*) \geq V_i(x'_i, B^*)$, for all $x'_i \epsilon B_i^*$. If a change from the equilibrium outcome (x^*, y^*, B^*) to outcome (x', y', B') constitutes a Pareto improvement, it must therefore be that $B^* \neq B'$. The profile of budget sets B^* induces a profile of incomes $w(B^*)$. Since $p(B_i^*) = p(B'_i) = p^*$ for all i, $B^* \neq B'$ implies that $w(B^*) \neq w(B')$. Because in equilibrium each firm is profit maximizing, it must hold that

$$\sum_{i=1}^{I} p^* e_i + \sum_{i=1}^{I}\sum_{j=1}^{J} \theta_{ij}\, p^* y_j^* \geq \sum_{i=1}^{I} p^* e_i + \sum_{i=1}^{I}\sum_{j=1}^{J} \theta_{ij}\, p^* y'_j,$$

it follows that $\sum_{i=1}^{I} w(B_i^*) \geq \sum_{i=1}^{I} w(B'_i)$. A change from (x^*, y^*, B^*) to (x', y', B') is a Pareto improvement if and only if $V_i(d_i(B_i^*), B^*) \geq V_i(d_i(B'_i), B')$, for all i, with one inequality strict. This is not possible if RLP holds since if $V_i(d_i(B_i^*), B^*) \geq V_i(d_i(B'_i), B')$, for all i, then (2) implies that $V_i(d_i(B_i^*), B^*) = V_i(d_i(B'_i), B')$, for all i. ‖

Proof of Theorem 5. Let B and B' be two profiles of budget sets with $B \neq B'$, $\sum_{i \in I} w(B_i) \geq \sum_{i=1}^{I} w(B'_i)$, and $p(B_i) = p(B'_k)$ for all $i, k = 1, \ldots, I$. Let r be a consumer who loses most in terms of income by a change from B to B'; *i.e.*, for all i

$$w(B_r) - w(B'_r) \geq w(B_i) - w(B'_i). \tag{A.2}$$

Let $w_k = w(B_k)$, $w'_k = w(B'_k)$. Note that

$$w_r - w'_r > 0 \tag{A.3}$$

and

$$\sum_i w_i \geq \sum_i w'_i. \tag{A.4}$$

We let $v_r(w_r) = m_r(d_r(p, w_r))$ and $v'_r(\cdot)$ be the associated derivative. Let $\bar{v} = \max_{w \leq \sum_{i=1}^{I} w_i} v'_r(w)$ and $\underline{v} = \min_{w \leq \sum_{i=1}^{I} w_i} v'_r(w)$.[24] Let

$$V_r = V_r(m_r(d_r(p, w_r)), m_r(d_r(p, w_1)), \ldots, m_r(d_r(p, w_I)))$$

and

$$V'_r = V_r(m_r(d_r(p, w'_r)), m_r(d_r(p, w'_1)), \ldots, m_r(d_r(p, w'_I))).$$

Finally, let

$$\mu = \frac{\sum_k v_r(w_k)}{I} \quad \text{and} \quad \mu' = \frac{\sum_k v_r(w'_k)}{I}.$$

We need two related preliminary facts.

Lemma 2. $\sum_{k \neq r}(v_r(w_k) - v_r(w'_k)) \geq -((\bar{v} - \underline{v})(I-1) + \underline{v})(w_r - w'_r)$.

Proof.

$$\begin{aligned}
\sum_{k \neq r}(v_r(w_k) - v_r(w'_k)) &\geq \sum_{k: w_k < w'_k} \bar{v}(w_k - w'_k) + \sum_{k: k \neq r, w_k > w'_k} \underline{v}(w_k - w'_k) \\
&\geq \sum_{k: w_k < w'_k} (\bar{v} - \underline{v})(w_k - w'_k) - \underline{v}(w_r - w'_r) \\
&\geq -((\bar{v} - \underline{v})(I-1) + \underline{v})(w_r - w'_r),
\end{aligned}$$

24. We drop the subscript r on $\underline{v}$ and $\bar{v}$ to simplify notation.

where the second inequality follows from inequality (A.4) and the third inequality follows from inequalities (A.2) and (A.4). ‖

Lemma 3. $\mu - \mu' \geq -(\bar{v} - \underline{v})(w_r - w_r')$.

Proof.

$$\begin{aligned} I(\mu - \mu') &= \sum_{k \neq r} (v_r(w_k) - v_r(w_k')) + v_r(w_r) - v_r(w_r') \\ &\geq -((\bar{v} - \underline{v})(I-1) + \underline{v})(w_r - w_r') + \underline{v}(w_r - w_r') \\ &\geq -(I-1)(\bar{v} - \underline{v})(w_r - w_r'), \end{aligned}$$

where the first inequality uses Lemma 2. ‖

1. It suffices to show that

$$\begin{aligned} &v_r(w_r) - \frac{\alpha_r}{I-1} \sum_{k=1}^{I} \max\{v_r(w_k) - v_r(w_r), 0\} - \frac{\beta_r}{I-1} \sum_{k=1}^{I} \max\{v_r(w_r) - v_r(w_k), 0\} \\ &> v_r(w_r') - \frac{\alpha_r}{I-1} \sum_{k=1}^{I} \max\{v_r(w_k') - v_r(w_r'), 0\} - \frac{\beta_r}{I-1} \sum_{k=1}^{I} \max\{v_r(w_r') - v_r(w_k'), 0\}. \end{aligned} \tag{A.5}$$

For any $w_k > w_r$ such that $w_r - w_r' > w_k - w_k'$, decreasing w_k' to $w_k + w_r' - w_r$ increases the right-hand side of inequality (A.5) without violating inequalities (A.2) or (A.4). So, since $w_r - w_r' \geq w_k - w_k'$ by inequality (A.2), in order to prove the result it is sufficient to show that inequality (A.5) holds when $w_k > w_r$ implies that $w_r - w_r' = w_k - w_k'$. Hence, we take $w_r = \max_{1 \leq k \leq I} w_k$. Consequently, it suffices to show that

$$\begin{aligned} &v_r(w_r) - \frac{\beta_r}{I-1} \sum_{k=1}^{I} (v_r(w_r) - v_r(w_k)) \\ &> v_r(w_r') - \frac{\alpha_r}{I-1} \sum_{w_k' > w_r'} (v_r(w_k') - v_r(w_r')) - \frac{\beta_r}{I-1} \sum_{w_r' > w_k'} (v_r(w_r') - v_r(w_k')). \end{aligned} \tag{A.6}$$

Since the right-hand side of inequality (A.6) is no greater than

$$v_r(w_r') - \frac{\beta_r}{I-1} \sum_{k=1}^{I} (v_r(w_r') - v_r(w_k')),$$

inequality (A.6) holds whenever

$$v_r(w_r) - \frac{\beta_r}{I-1} \sum_{k=1}^{I} (v_r(w_r) - v_r(w_k)) > v_r(w_r') - \frac{\beta_r}{I-1} \sum_{k=1}^{I} (v_r(w_r') - v_r(w_k')). \tag{A.7}$$

To complete the proof it suffices to show that

$$(1 - \beta_r)(v_r(w_r) - v_r(w_r')) > \frac{\beta_r}{I-1} \sum_{k \neq r} (v_r(w_k') - v_r(w_k)). \tag{A.8}$$

Since $v_r(w_r) - v_r(w'_r) \geq \underline{v}(w_r - w'_r)$, it follows from Lemma 2 that inequality (A.8) holds provided that

$$(1-\beta_r)\underline{v} > \beta_r\left((\bar{v}-\underline{v}) + \frac{\underline{v}}{I-1}\right) \quad \text{or} \quad \beta_r < \frac{1}{\bar{v}/\underline{v} + 1/(I-1)}.$$

2. By the triangle inequality,

$$|v_r(w_r) - v_r(w'_r) - (\mu - \mu')| \geq |v_r(w_r) - \mu| - |v_r(w'_r) - \mu'|. \tag{A.9}$$

If $v_r(w_r) - v_r(w'_r) \geq \mu - \mu' > 0$, then the result follows from inequality (A.9).
If $v_r(w_r) - v_r(w'_r) \geq 0 > \mu - \mu'$, then let

$$\beta_r \leq \underline{v}/\bar{v}. \tag{A.10}$$

It follows that

$$\begin{aligned}(1-\beta_r)(v_r(w_k) - v_r(w'_k)) &\geq (1-\beta_r)\underline{v}(w_r - w'_r)\\ &\geq \beta_r(\bar{v}-\underline{v})(w_r - w'_r)\end{aligned}$$

and therefore

$$v_r(w_r) - v_r(w'_r) \geq \beta_r((v_r(w_r) - v_r(w'_r)) + (\bar{v}-\underline{v})(w_r - w'_r)). \tag{A.11}$$

Also we have

$$\begin{aligned}v_r(w_r) - v_r(w'_r) + (\bar{v}-\underline{v})(w_r - w'_r) &\geq v_r(w_r) - v_r(w'_r) - (\mu - \mu')\\ &> |(v_r(w_r) - \mu)| - |(v_r(w'_r) - \mu')|,\end{aligned}$$

where the first inequality follows from Lemma 3 and the second inequality follows from inequality (A.9). It follows that if inequality (A.10) holds, then RLP holds.
Finally, when $v_r(w_r) - v_r(w'_r) < \mu - \mu'$, it follows from inequality (A.9) that it suffices to show that

$$v_r(w_r) - v_r(w'_r) > \beta_r(\mu - \mu' + v_r(w_r) - v_r(w'_r)). \tag{A.12}$$

Inequality (A.12) holds if $\beta_r < \underline{v}/(\bar{v}-\underline{v})$ since $(w_r - w'_r)\bar{v} \geq \mu - \mu'$ and $v_r(w_r) - v_r(w'_r) \geq \underline{v}(w_r - w'_r)$. The result follows because

$$\frac{\underline{v}}{\bar{v}-\underline{v}} > \frac{\underline{v}}{\bar{v}}.$$

3. Note that if $v_r(w_i) = \min_k v_r(w_k)$, then

$$\begin{aligned}\min\{v_r(w_1), \ldots, v_r(w_I)\} - \min\{v_r(w'_1), \ldots, v_r(w'_I)\} &\geq v_r(w_i) - v_r(w'_i)\\ &\geq \bar{v}(w_i - w'_i)\\ &\geq -(I-1)(w_r - w'_r).\end{aligned} \tag{A.13}$$

The third inequality follows from inequality (A.2).
When $\beta_r > 0$,

$$V_r - V'_r \geq (w_r - w'_r)(\underline{v} - \beta_r(1-\delta_r)(\bar{v} - \underline{v} + \frac{\underline{v}}{I-1}) - \beta_r\delta_r\bar{v}), \tag{A.14}$$

where the inequality follows from inequality (A.13) and Lemma 2. Consequently, RLP holds whenever

$$\frac{\underline{v}}{(1-\delta_r)(\bar{v} - \underline{v} + \frac{\underline{v}}{I-1}) + \delta_r\bar{v}} > \beta_r > 0.$$

When $\beta_r < 0$,

$$\begin{aligned}V_r - V'_r &\geq v_r(w_r) - v_r(w'_r) + \frac{\beta_r(1-\delta_r)}{I-1}\left(\sum_{k\neq r}(v_r(w_k) - v_r(w'_k))\right)\\ &\quad + \beta_r\delta_r(v_r(w_i) - v_r(w'_i))\\ &\geq (v_r(w_r) - v_r(w'_r))\left(1 + \frac{\beta_r\delta_r}{I-1}\right) + \frac{\beta_r(1-\delta_r)}{I-1}\left(\sum_{k:w_k<w'_k}\bar{v}(w_k - w'_k)\right)\\ &\geq \underline{v}\left(1 + \frac{\beta_r\delta_r}{I-1}\right)(w_r - w'_r) + \beta_r\bar{v}(w_r - w'_r).\end{aligned}$$

The first inequality holds when $v_r(w_i') = \min_k v_r(w_k')$. One obtains the second inequality by discarding positive terms and using the definition of $\bar{v}$. Provided that

$$1 + \frac{\beta_r \delta_r}{I-1} > 0$$

(which holds for sufficiently large I), the third inequality follows from inequalities (A.2) and (A.4). Hence, RLP holds provided that

$$\beta_r > -\frac{\underline{v}}{\bar{v} + \frac{\beta_r \delta_r \underline{v}}{I-1}} \quad \|$$

Computation for Example 3. We do the calculation for the rich agent. Let $i \in \{1, \ldots, n\}$. Note that

$$\frac{n}{2n-1} \geq \frac{1}{2} \geq \frac{n-1}{2n-1}. \tag{A.15}$$

Then we have

$$\begin{aligned}
&V_i(x_i', B') - V_i(e_i, B^*) \\
&= (1-\varepsilon)^\alpha + \frac{\beta}{2n-1}((n-1)(1-\varepsilon)^\alpha + n(\eta+\varepsilon)^\alpha) - 1 - \frac{\beta}{2n-1}((n-1) + n\eta^\alpha) \\
&= \left(1 + \frac{\beta(n-1)}{2n-1}\right)((1-\varepsilon)^\alpha - 1) + \beta\frac{n}{2n-1}((\eta+\varepsilon)^\alpha - \eta^\alpha) \\
&\geq (1+\beta/2)((1-\varepsilon)^\alpha - 1) + \beta/2((\eta+\varepsilon)^\alpha - \eta^\alpha),
\end{aligned}$$

where the inequality follows from equation (A.15) since $(1-\varepsilon)^\alpha - 1 < 0$ and $(\eta+\varepsilon)^\alpha - \eta^\alpha > 0$. Hence, it suffices to show that the last expression is strictly positive for ϵ sufficiently close to 0. As the expression is zero for $\varepsilon = 0$, it suffices to show that the right derivative with respect to ϵ is positive at 0. Taking the derivative and setting $\varepsilon = 0$, one has $-(1+\beta/2)\alpha + \beta/2\alpha\eta^{\alpha-1}$. For $\alpha < 1$, and $\eta < (\beta/(2+\beta))^{1/(1-\alpha)}$, the above expression is positive and, thus, the altruistic agents are better off after the redistribution. As the poor agents are selfish, they benefit from the redistribution. Thus, we have robust inefficiency even if $n \to \infty$.

Proof of Lemma 1. If x is not in the internal core, then there is a coalition C and a C-feasible allocation $x' = (x_k')_{k \in C}$ such that $m_i(x_i') > m_i(x_i)$ for $i \in C$. From monotonicity and continuity of $m_i(\cdot)$, we can find $\eta_i \ll x_i', i \in C$ such that $m_i(\eta_i) = m_i(x_i)$ for $i \in C$. Let

$$z = \sum_{i \in C}(x_i' - \eta_i) \geq 0, z \neq 0.$$

GSM implies that the coalition C can improve upon x. Hence, x is not in the core. $\|$

Proof of Theorem 7. The core of the internal economy is not empty because internal preferences are convex (Scarf, 1967). We want to show that x belongs to the core of the original economy. If not, there is a coalition C and a C-feasible allocation x' such that all members in C prefer x' to x. As x belongs to the internal core, some members of C must have a lower internal utility. This contradicts the assumption. $\|$

APPENDIX B

When a game has externalities, a coalition must take into account the reaction of the complementary coalition in order to decide whether a defection is attractive. Different models of how the complementary coalition reacts lead to different notions of the core because they generate different conditions under which a coalition can improve upon a given allocation. In the text, we took the view that agents outside of the coalition do not change their behaviour. In this appendix, we review other notions that have appeared in the literature.

One possibility is that a coalition C can improve upon x if there is a feasible reallocation within the coalition that ensures a social state preferred by all the agents in C regardless of the strategies the other agents outside the coalition may choose. Formally, we say that C can α-improve upon x if there is an $(x_k')_{k \in C}$ that is feasible for C such that

$$U_i((x_k')_{k \in C}, (x_k')_{k \notin C}) > U_i(x) \quad \text{for all } i \in C$$

and for all $(x_k')_{k \notin C}$ that are feasible for the complement of C.

In this definition, a coalition can improve upon an allocation only if it can find a reallocation of its resources that increases the utility of its members for any feasible behaviour of the individuals outside of the coalition. This definition makes it difficult for a coalition to improve upon an allocation when there are externalities. For example, suppose that everyone in the economy cares about the well being of a particular, poor agent, Agent 0. Take an allocation in which Agent 0 receives an adequate allocation. No coalition that excludes Agent 0 can improve upon the allocation because the complementary coalition can threaten to "starve" Agent 0. Informally, one would expect small coalitions to have limited opportunities to make α-improvements because potential improvements must be tested against coordinated responses by the rest of the economy.

The literature considers two other variations of the core concept that replace α-improvement with other assumptions about how agents outside of a coalition respond to a deviation. Aumann and Peleg (1960) introduce the β-core, which consists of those allocations x in which a coalition cannot β-improve upon x. For a coalition to β-improve upon x it must be that for all $(x_k)_{k\notin C}$ that are feasible for the complement of C there is an $(x'_k)_{k\in C}$ that is feasible for C that makes every agent in C better off relative to x. Chander and Tulkens (1995) introduce the γ-core. Translated to our framework, outsiders consume their endowments when a coalition is formed. It is straightforward to show that the γ-core is contained in the β-core which is in turn contained in the α-core. All these cores will generally be "large" in the sense that conclusions of Lemma 1 and Theorem 6 will not hold for these definitions of the core. For example, the γ-core need not be a subset of the internal core. Since the α- and β-cores contain the γ-core, the following example demonstrates that all three of these cores may be quite large.

Example 5. *Let there be three agents and two goods. Let $m_i(x,y) = xy$ for $i = 1,2,3$. Suppose that Agents 1 and 3 are egoistic, so that for $i = 1$ and 3*

$$U_i(m_1,m_2) = m_i$$

and that $U_2(m_1,m_2) = \min\{m_1,m_2\}$. Let endowments be $\omega_1 = (0,1/3)$, $\omega_2 = (2/3,2/3)$ and $\omega_3 = (1/3,0)$.

The internally efficient allocations are

$$\{((z_1,z_1),(z_2,z_2),(z_3,z_3)) : z_i \geq 0,\ z_1+z_2+z_3 = 1\}.$$

The internal core consists of those internally efficient allocations that are also both internally individually rational and that cannot be improved upon by two-agent coalitions. The internal core is equal to

$$\left\{\left((z_1,z_1),\left(\frac{2}{3},\frac{2}{3}\right),(z_3,z_3)\right) : z_1+z_3 \geq 1/3, z_1, z_3 \geq \sqrt{\frac{2}{3}}\left(1-\frac{\sqrt{6}}{3}\right)\right\}. \tag{B.1}$$

To verify this description of the internal core, first note that efficiency requires that Agent i must consume equal quantities of the two goods. Individual rationality guarantees that Agent 2 receives at least $2/3$ of each good. If Agent 2 received more than $2/3$, then the coalition $\{1,3\}$ could improve itself. Finally, if either Agent 1 or Agent 3 received less than the lower bound in the set (B.1), then she could join with Agent 2 and improve herself.

The γ-core is large. Consider, e.g., the extreme allocation

$$x_1 = (1,1), \quad x_2 = (0,0), \quad x_3 = (0,0).$$

This allocation gives a utility of 0 to Agent 2. Now suppose that Agent 2 wants to deviate, say to his endowment $\omega_2 = (2/3,2/3)$. When Agent 1 consumes her endowment, Agent 2 receives utility 0 and hence does not improvement her payoff. Similarly, the coalition containing Agents 2 and 3 also cannot improve itself. It follows that elements of the γ-core need not be internally individually rational.

The α-, β- and γ-cores and the core defined in Definition 5 are equivalent when U_i depends only on x_i. Definition 5 makes it relatively easy to block a proposed allocation and, thus, creates existence problems as we show Example 4. The other core notions have less trouble with existence. In particular, the α- and β-cores are non-empty for the economy of Example 4.[25] While the γ-core for the economy of Example 4 is empty (for the same reasons that our core is empty), it is not hard to construct examples in which our core is empty while the γ-core is non-empty.

Theorem 7 gives conditions under which our core is non-empty. The α-, β- and γ-cores are also non-empty under these conditions. We give conditions under which the core of a large economy is contained in the set of competitive allocations in Theorem 6. Example 5 demonstrates that the γ-core will generally not be an element of the set of competitive allocations. It is straightforward to modify the example so that the conclusion of Theorem 6 also fails to hold for

25. Agent 2 would not join the coalition $C = \{1,2\}$ if Agent 3 can destroy his entire endowment.

the γ-core. Hence, the core-equivalence result will not hold for the α-, β- and γ-cores without further assumptions on preferences.

REFERENCES

ANDREONI, J. and MILLER, J. H. (2002), "Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism", *Econometrica*, **70**, 737–753.

ARROW, K. J. and HAHN, F. H. (1971), *General Competitive Analysis* (San Francisco: Holden-Day).

AUMANN, R. J. (1959), "Acceptable Points in General Cooperative n-person Games", in Tucker, A. W. and Luce, R. D. (eds) *Contributions to the Theory of Games IV*. (Princeton: Princeton University Press (287–324).

AUMANN, R. J. and PELEG, B. (1960), "Von Neumann-Morgenstern Solutions to Cooperative Games without Side Payments", *Bulletin of the American Mathematical Society*, **66**, 173–179.

BENJAMIN, D. (2008), "Social Preferences and the Efficiency of Bilateral Exchange" (Technical report, Cornell University).

BOLTON, G. (1991), "A Comparative Model of Bargaining: Theory and Evidence", *American Economic Review*, **81**, 1096–1135.

BOLTON, G. and OCKENFELS, A. (2000), "ERC: A Theory of Equity, Reciprocity and Competition", *American Economic Review*, **90**, 166–193.

BORGLIN, A. (1973), "Price Characterization of Stable Allocations in Exchange Economies with Externalities", *Joumal of Economic Theory*, **6**, 483–494.

CHANDER, P. and TULKENS, H. (1995), "A Core-Theoretic Solution for the Design of Cooperative Agreements on Transfrontier Pollution", *International Tax and Public Finance*, **2**, 279–293.

CHARNESS, G. and RABIN, M. (2002), "Understanding Social Preferences with Simple Tests", *Quarterly Journal of Economics*, **117**, 817–869.

COX, J. C. and SADIRAJ, V. (2006), "Direct Tests of Models of Social Preferences and Introduction of a New Model" (Working Paper No. 06-13, Georgia State University).

DEBREU, G. and SCARF, H. E. (1963), "A Limit Theorem on the Core of an Economy", *International Economic Review*, **4**, 235–246.

DEKEL, E., LIPMAN, B. L. and RUSTICHINI, A. (2001), "Representing Preferences with a Unique Subjective State Space", *Econometrica*, **69**, 891–934.

DIERKER, E. and GRODAL, B. (1995), "Profit Maximization, Relative Prices, and the Maximization of Shareholders Real Wealth" (Discussion Paper No. 95-07, University of Copenhagen).

EDGEWORTH, F. Y. (1881), *Mathematical Psychics* (New York: Augustus M. Kelley).

DUBEY, P. and SHUBIK, M. (1985), "Perfect Competition in Strategic Market Games with Interlinked Preferences", *Economics Letters*, **17**, 3–4.

ELLICKSON, B. (1993), *Competitive Equilibrium: Theory and Applications* (Cambridge: Cambridge University Press).

FEHR, E. and GÄCHTER, S. (2000), "Fairness and Retaliation: The Economics of Reciprocity", *Journal of Economic Perspectives*, **14**, 159–181.

FEHR, E. and SCHMIDT, K. (1999), "A Theory of Fairness, Competition, and Cooperation", *Quarterly Journal of Economics*, **114**, 817–868.

GEANAKOPLOS, J. and POLEMARCHAKIS, H. (2008), "Pareto Improving Taxes", *Journal of Mathematical Economics*, **44**, 682–692.

GERSBACH, H. and HALLER, H. (2001), "Collective Decisions and Competitive Markets", *Review of Economic Studies*, **68**, 347–368.

GERSBACH, H. and HALLER, H. (2009), "Club Theory and Household Formation" (Technical Report, Virginia Polytechnic University).

GOLDMAN, S. M. (1978), "Gift Equilibria and Pareto Optimality", *Journal of Economic Theory*, **18**, 368–370.

GUL, F. and PESENDORFER, W. (2001), "Temptation and Self-control", *Econometrica*, **69**, 1403–1435.

HOCHMAN, H. M. and RODGERS, J. D. (1969), "Pareto Optimal Redistribution", *American Economic Review*, **59**, 542–557.

HURWITZ, L. and UZAWA, H. (1971), "On the Integratability of Demand Functions", in Chipman, J. S., Hurwicz, L., Richter, M. K. and Sonnenschein, H. F. (eds) *Preferences, Utility and Demand* (New York: Harcourt Brace, Jovanovich) 174–214.

KARNI, E. and SAFRA, Z. (2002), "Individual Sense of Justice: A Utility Representation", *Econometrica*, **70**, 263–284.

KIRCHSTEIGER, G. (1994), "The Role of Envy in Ultimatum Games", *Journal of Economic Behavior and Organization*, **25**, 373–390.

KRANICH, L. J. (1988), "Altruism and Efficiency: A Welfare Analysis of the Walrasian Mechanism with Transfers", *Journal of Public Economics*, **36**, 369–386.

KREPS, D. M. (1979), "A Representation Theorem for 'Preference for Flexibility"', *Econometrica*, **47**, 565–578.

MACCHERONI, F., MARINACCI, M. and RUSTICHINIC, A. (2008), "Social Decision Theory: Choosing within and between Groups" (Technical Report, University of Minnesota).

MAS-COLELL, A. (2001), *The Theory of General Economic Equilibrium: A Differentiable Approach*. (Cambridge: Cambridge University Press).

MAS-COLELL, A., WHINSTON, M. D. and GREEN, J. R. (1995), *Microeconomic Theory* (New York: Oxford University Press).
NEHRING, K. and PUPPE, C. (2002), "A Theory of Diversity", *Econometrica*, **70**, 1155–1198.
NOGUCHI, M. and ZAME, W. R. (2006), "Competitive Markets with Externalities", *Theoretical Economics*, **1**, 143–166.
PUPPE, C. (1996), "An Axiomatic Approach to "Preference for Freedom of Choice", *Journal of Economic Theory*, **68**, 174–199.
RADER, T. (1980), "The Second Theorem of Welfare Economics when Utilities are Interdependent,", *Joumal of Economic Theory*, **23**, 420–424.
ROBERTS, J. and POSTLEWAITE, A. (1976), "The Incentives for Price-Taking Behavior in Large Exchange Economies", *Econometrica*, **44**, 115–127.
SCARF, H. E. (1967), "The Core of an n-person Game", *Econometrica*, **35**, 50–69.
SMITH, A. (1759), *The Theory of Moral Sentiment* (London: A. Millar).
SOBEL, J. (2005), "Interdependent Preferences and Reciprocity", *Journal of Economic Literature*, **43**, 396–440.
SOBEL, J. (2010), "Do Markets Make People Look Selfish?" (Technical Report, University of California, San Diego).
VARIAN, H. R. (1973), "Equity, Envy, and Efficiency", *Journal of Economic Theory*, **9**, 63–91.
VARIAN, H. R. (1976), "Two Problems in the Theory of Fairness", *Journal of Public Economics*, **5**, 249–260.
VOSTROKNUTOV, A. (2007), "Preferences over Consumption and Status" (Technical Report, University of Minnesota).
WINTER, S. (1969), "A Simple Remark on the Second Optimality Theorem of Welfare Economics", *Journal of Economic Theory*, **1**, 99–103.