Weakening transferable utility: the case of non-intersecting Pareto curves

Article (Accepted Version)

Demuynck, Thomas and Potoms, Tom (2020) Weakening transferable utility: the case of non-intersecting Pareto curves. *Journal of Economic Theory*, 188. a105035 1-37. ISSN 0022-0531

This version is available from Sussex Research Online: <http://sro.sussex.ac.uk/id/eprint/92828/>

This document is made available in accordance with publisher policies and may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the URL above for details on accessing the published version.

Copyright and reuse:

Sussex Research Online is a digital repository of the research output of the University.

Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable, the material made available in SRO has been checked for eligibility before being made available.

Copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

Weakening transferable utility: the case of non-intersecting Pareto curves*

Thomas Demuynck[†] Tom Potoms[‡]

March 19, 2020

Abstract

Transferable utility (TU) is a widely used assumption in economics. In this paper, we weaken the TU property to a setting where distinct Pareto frontiers have empty intersections. We call this the no-intersection property (NIP). We show that the NIP is strictly weaker than TU, but still allows to derive several desirable properties. We discuss the NIP in relation to several models where TU has turned out to be a key assumption: models of assortative matching, principal-agent models with asymmetric information, the Coase Independence Property and Becker's Rotten Kid Theorem.

Keywords: Pareto efficiency, Transferable utility, Assortative matching, Coase Independence Property, Rotten Kid Theorem

JEL: C78; D13; D60; D61; D62.

1 Introduction

Transferable utility (TU) is a widely used assumption in economics. A model is said to have transferable utility when an agent can transfer part of her utility to another agent in a lossless manner. The main attractive feature of the TU assumption stems from its desirable aggregation properties. Under TU, individual members may affect the location of the Pareto frontier but not its shape which is, up to normalization, a straight line with slope minus one. This means that, under Pareto optimality, the distribution of utilities over the agents does not affect the group's decision which will be to pick the Pareto

*We thank 2 anonymous referees, the Associate Editor and the Editor Marciano Siniscalchi for useful comments on an earlier version of this paper.

[†]ECARES, Université Libre de Bruxelles. Avenue F. D. Roosevelt 50, CP 114, B-1050 Brussels, Belgium. E-mail: thomas.demuynck@ulb.be. Thomas Demuynck acknowledges financial support by the Fonds de la Recherche Scientifique-FNRS under grant nr F.4516.18 and EOS project 30544469.

[‡]University of Sussex. Sussex House, Falmer Brighton, BN1 9RH United Kingdom. E-mail: T.Potoms@sussex.ac.uk. Part of this research was conducted while at ECARES, Université Libre de Bruxelles.

frontier farthest from the origin. Results that exploit this property are the famous Coase Independence Property (Coase, 1960) and Becker’s Rotten Kid Theorem (Becker, 1974). The TU assumption is also frequently used in other economic environments like one-to-one matching models.

It is well known that if TU does not hold, Pareto frontiers may intersect. Such intersections lead to several paradoxes, in particular, a potential failure of Coasian Independence and Rotten Kid Theorem, or in the field of welfare economics it can lead to cycles in the Kaldor-Hicks compensation criterion¹. In this respect, there appears to be a discontinuity from the rather restrictive transferable utility setting where aggregation is very easy, and the general setting with intersecting Pareto curves, where it’s very difficult to obtain clear-cut results. In this paper, we weaken the notion of transferable utility, but retain the condition that distinct Pareto frontiers have empty intersections. We call this the No-Intersection-Property (NIP). As we show in the paper, many of the properties attributed to TU also hold under the NIP. To the best of our knowledge, the NIP has not yet been studied as a separate property.

It is a well known fact that if a model has transferable utility then distinct Pareto frontiers have empty intersections. As such, any model that satisfies TU will also satisfy the NIP. It can be hypothesized whether the converse is also true. In this paper, we show that the answer is negative. Towards this result, we first show that the NIP is equivalent to the existence of a *Pareto (aggregation) function* which combines the utility levels of all agents into a single number and a *surplus function* that associates with every Pareto frontier a number that indicates its distance from the origin. A profile of utility values (one for each agent) belongs to a certain Pareto frontier if the value of the Pareto function equals the value of the surplus function. As such, higher numbers for the Pareto function correspond to profiles of utilities that lie on Pareto curves further from the origin. We show that TU is obtained if and only if this Pareto function is *additively separable*.

Next, we revisit some results from the literature that heavily rely on the TU assumption and we look at them from the perspective of the weaker NIP assumption. As a first exercise, we look at models of (two-sided) matching. We refer to Chade, Eeckhout, and Smith (2017) and Chiappori (2017) for a recent overview of matching and search theory. In much of this literature, attention has been devoted to either cases where there are no transfers allowed between agents; the so called non-transferable utility (NTU) case, initiated by Gale and Shapley (1962)), or cases with perfectly transferable utility; the TU case, initiated by Shapley and Shubik (1972). For the latter, the seminal contribution by Becker (1973) showed that, when the match surplus function is supermodular, then any pairwise stable matching exhibits positive assortative matching. The fact that matching patterns under TU can be easily analysed in terms of the surplus function makes it a very attractive model for theoretical and empirical applications.²

¹The Kaldor-Hicks compensation criterion (Hicks, 1939; Kaldor, 1939) states that a change from one situation to another is an improvement if the agents that gain from the switch could, in theory, compensate those who are harmed by the change. It is well-known that the Kaldor-Hicks compensation criterion may lead to inconsistent judgements (cycles) when Pareto curves intersect, see for example Scitovsky (1941).

²See for example the empirical framework of Choo and Siow (2006) or the more complex strategic

The relative simplicity of analysing matching patterns under TU stems from the fact that the problem of distributing the surplus within a couple and maximizing the total surplus over all possible matches can be treated separately. Essentially, our NIP condition can be seen as positing the potential for a similar separation, but then in the context of an imperfectly transferable utility framework, which is an intermediate setting between TU and NTU. We show that under the NIP, if the surplus function satisfies some monotonicity condition with respect to the characteristics of the agents, then the submodularity of the Pareto function and the supermodularity of the surplus function, with at least one of these conditions strict, is sufficient for positive assortative matching patterns to arise in equilibrium. We also provide a differential sufficient conditions for positive assortative matching. This can be contrasted with Legros and Newman (2007), who provide sufficient conditions on the Pareto frontiers for assortative matching to occur, in a more general setting where Pareto frontiers are also allowed to intersect. Building on a recent contribution by Nöldeke and Samuelson (2018), in which the authors draw a close connection between two sided matching models and principal-agent models, we also incorporate this principal-agent setting and show that in our setting the strict single crossing or Spence-Mirrlees condition is satisfied under the same conditions as for our aforementioned assortative matching result: monotonicity and supermodularity of the surplus function and submodularity of the Pareto function. Under this condition, any incentive compatible contract will require higher actions from higher type agents.

As a second exercise, we revisit the Coase Independence Property. We follow the main theoretical contributions in this literature and present a formal setting along the lines of Hurwicz (1995) and Bergstrom (2017). In these papers, the authors introduce a general equilibrium framework in which agents consume a private good and several public goods (or externalities, either positive or negative). The main feature of these models is that the feasible allocations in the economy can be described through a *linear* feasibility restriction. A remarkable result, as discussed in both papers, is that Coasian Independence, which states that the Pareto efficient level of public goods is independent of the allocation of initial endowments, holds if and only if the utility possibility frontiers satisfy TU. In the present paper, we expand this result and provide a more generalized characterization of the Coase Independence Property. In particular we show that, if one allows for a general technology, including non-additive or non-linear technologies, a necessary and sufficient condition for Coasian Independence is that the utility possibility sets satisfy NIP. Therefore, NIP characterizes the settings with Coasian Independence in these more general environments.

The final exercise reconsiders the so-called Becker’s Rotten Kid Theorem (Becker, 1974). The name ‘Rotten Kid’ originates from the usual description of the two-stage game, where after the kids take certain actions which determine the level of the public goods, the parent divides the remaining resources over the kids. The Rotten Kid Theorem states that in such a setting, any Pareto efficient level of public goods can be implemented as a subgame perfect equilibrium of this two-stage game. This result is remarkable and important for the theory

settings in which agents perform pre-match investments in education and then decide on partner choice as in Chiappori, Costa Dias, and Meghir (2018).

of incentives and household economics (e.g. Chiappori and Mazzocco (2017)). Essentially, it provides conditions such that there is no need for complex incentive schemes to implement Pareto optimal allocations. A particular way to retrieve this result is when the kids in the first stage have preferences that lead to Pareto frontiers that are representable as a simplex (i.e. there is TU), joint with a ‘benevolent’ (altruistic) parent (Bergstrom, 1989). We show that the NIP condition, together with a relatively mild assumption on the selection rule of the optimal allocations by the parent is enough to recover the Rotten Kid Theorem.

The structure of the paper is as follows. Section 2 introduces notation, definitions and assumptions. Section 3 introduces the NIP condition and provides the representation theorems. Section 4 relates the NIP with TU and shows that the NIP is a strictly weaker condition. Section 5 discusses the NIP in relation to models of assortative matching and delegation, the Coase Independence Property and Becker’s Rotten Kid Theorem. Section 6 contains a conclusion. The appendix contains the proofs of all the theorems.

2 Notation and definitions

We consider a setting with N agents. The utility level of agent $i \leq N$ is denoted by $u_i \in \mathbb{R}$. A *utility profile* $\mathbf{u} = (u_1, \dots, u_N) \in \mathbb{R}^N$ associates a utility level to each agent i . Utility profiles are denoted by bold letters, e.g. $\mathbf{u}, \mathbf{v}, \mathbf{w}$. We use vector inequalities $\mathbf{u} \geq \mathbf{v}$ if $u_i \geq v_i$ for all $i \leq N$, we write $\mathbf{u} > \mathbf{v}$ if $\mathbf{u} \geq \mathbf{v}$ and $\mathbf{u} \neq \mathbf{v}$ and we write $\mathbf{u} \gg \mathbf{v}$ if $u_i > v_i$ for all i .

A function $g : \mathbb{R}^N \rightarrow \mathbb{R}$ is *weakly increasing (decreasing)* if $\mathbf{u} \gg \mathbf{v}$ implies $g(\mathbf{u}) > (<) g(\mathbf{v})$ and $\mathbf{u} \geq \mathbf{v}$ implies $g(\mathbf{u}) \geq (\leq) g(\mathbf{v})$. We call the function *increasing (decreasing)* if $\mathbf{u} > \mathbf{v}$ implies $g(\mathbf{u}) > (<) g(\mathbf{v})$.

We assume that each agent has an outside option that gives a utility level $\underline{u}_i$ which we normalize, for convenience, to zero, i.e. $\underline{u}_i = 0$. Let $\Phi \subseteq \mathbb{R}^k$ be a set containing the possible values of the environmental variables relevant to the model under consideration. A vector $\phi \in \Phi$ specifies all variables that determine the utility possibility set of the N -agent model. For each value $\phi \in \Phi$, we can associate a *utility possibility set* that contains all feasible and individually rational utility profiles that are attainable when the environmental variables take on the value ϕ . We denote this utility possibility set by $C(\phi)$.

$$C(\phi) = \{\mathbf{u} \in \mathbb{R}_+^N : \mathbf{u} \text{ is attainable in situation } \phi\}.$$

We denote by $\mathcal{D} = \bigcup_{\phi \in \Phi} C(\phi)$ the set of all relevant utility profiles, i.e. the utility profiles that are in some utility possibility set.

Throughout this paper we will impose the following assumption on the collection $(C(\phi))_{\phi \in \Phi}$.

Assumption A. For all $\phi \in \Phi$, $C(\phi)$ is non-empty, closed, bounded and comprehensive, i.e. if $\mathbf{u} \in C(\phi)$ and $\mathbf{0} \leq \mathbf{v} \leq \mathbf{u}$ then $\mathbf{v} \in C(\phi)$.

We denote the *Pareto frontier* of $C(\phi)$ by $\partial C(\phi)$. It is defined as the upper boundary of the set $C(\phi)$.

$$\partial C(\phi) = \{\mathbf{u} \in C(\phi) : \forall \mathbf{v} \gg \mathbf{u}, \mathbf{v} \notin C(\phi)\}.$$

This set is well defined under Assumption A. Let us illustrate this with the following running example.

Example 1. Consider a set of employees and a set of firms. Every employee has a characteristic δ from a set Δ and every firm has a characteristic ξ from a set Ξ . If an employee and a firm are matched, they produce at most a single unit of output which is sold at a price P . If the good is not produced, then the utility level of the employee and profits of the firm are both zero.

If the good is produced, then the utility of the employee is given by a utility function $u(\ell, w)$ which is increasing in both the leisure ℓ and wage w paid by the firm to the employee. The utility (profit) function of the firm f is equal to $\pi = P - w$. We assume that each employee has a total time endowment of 1 and that the good is produced if and only if the employee provides a high enough level of work:

$$1 - \ell \geq \omega(\delta, \xi).$$

Here, ω is some threshold function that depends on the characteristics of both the firm and the employee. Given this setting, we can define the set of environmental variables as $\Phi = \Delta \times \Xi$. If we invert the utility function $u = u(\ell, w)$ with respect to ℓ , we obtain a function $\ell(u, w)$ which is increasing in u and decreasing in w . Substituting the identity $w = P - \pi$ into this function gives the following formulation of the utility possibility set of a firm-employee match given characteristics $(\delta, \xi) \in \Phi$,

$$C(\delta, \xi) = \{(u, \pi) \in \mathbb{R}_+^2 : \ell(u, P - \pi) \leq 1 - \omega(\delta, \xi)\}.$$

3 The no intersection property

We define a collection of utility possibility sets to satisfy the *no intersection property* (NIP) if distinct Pareto frontiers have empty intersections.

Definition 3.1 (No-Intersection Property (NIP)). Let $(C(\phi))_{\phi \in \Phi}$ be a collection of utility possibility sets satisfying Assumption A. Then $(C(\phi))_{\phi \in \Phi}$ satisfies the no intersection property (NIP), if for all $\phi, \psi \in \Phi$:

$$\text{if } \partial C(\phi) \cap \partial C(\psi) \neq \emptyset, \text{ then } C(\phi) = C(\psi).$$

Figure 1 provides several examples in a 2 agent setting. Subfigure 1a and 1b show cases where the Pareto frontiers don't intersect and hence satisfy the NIP condition. Subfigure 1b is particular in the sense that the Pareto frontiers are straight lines with slope equal to -1 . This is a setting with *transferable utility* (TU). It is clear from this figure that TU implies NIP, but as we will show in section 4, the reverse does not necessarily hold. The two lower panels represent cases where the NIP is not satisfied. In Figure 1c, the frontiers of the utility possibility sets $C(\phi)$ and $C(\psi)$ are touching at point A. Figure 1d on the other hand presents a case where the Pareto frontiers intersect at A. Both cases violate

Figure 1 Each figure represents a collection of Pareto frontiers corresponding to a collection of utility possibility sets, $(C(\phi))_{\phi \in \Phi}$.

Definition 3.1: the point A is in the intersection of both Pareto curves, but the utility possibility sets are not identical.

Before we present our main representation result, let us introduce two additional assumptions. The first requires a richness condition on the collection of the utility possibility sets. It requires that every utility profile in $\mathcal{D} = \cup_{\phi \in \Phi} C(\phi)$ is on some Pareto curve.

Assumption B. For all utility vectors $\mathbf{u} \in \mathcal{D}$, there is a value $\phi \in \Phi$ such that $\mathbf{u} \in \partial C(\phi)$.

A second assumption requires that the utility possibility correspondence is both upper and lower hemicontinuous.

Assumption C. The set Φ is compact and the utility possibility correspondence $C : \Phi \rightarrow \mathbb{R}_+^N$ is continuous, i.e. both upper and lower hemicontinuous at all $\phi \in \Phi$.³

It will often be convenient to use an equivalent definition of the NIP. For two subsets $A, B \subseteq \mathbb{R}_+^N$ we write $A \sqsubset B$ if for all vectors $\mathbf{u} \in A$ there exists a vector $\mathbf{v} \in B$ such that

³In particular, given Assumption A upper hemicontinuity requires that for all $\phi^t \rightarrow \phi$ and $\mathbf{u}^t \in C(\phi^t)$ for all t , then there is a subsequence $(\phi^{t_i})_{i \in \mathbb{N}}$ such that $\mathbf{u}^{t_i} \rightarrow \mathbf{u}$ and $\mathbf{u} \in C(\phi)$. Lower hemicontinuity requires that if $\phi^t \rightarrow \phi$ and $\mathbf{u} \in C(\phi)$ then there exists a number $N \in \mathbb{N}$ and a sequence $(\mathbf{u}^t)_{t \geq N}$ such that $\mathbf{u}^t \rightarrow \mathbf{u}$ and $\mathbf{u}^t \in C(\phi^t)$ for all $t \geq N$.

$\mathbf{v} \gg \mathbf{u}$. We write $A \sqsubseteq B$ if $A = B$ or $A \subset B$. As shown in Lemma A.2 in the appendix, a collection of utility possibility sets satisfies the NIP, if and only if they are completely ordered by the binary relation $\sqsubseteq$, i.e. the relation $\sqsubseteq$ is a linear order on the set $(C(\phi))_{\phi \in \Phi}$. Given this linear order on the utility possibilities sets, one might try to give each Pareto frontier a numerical value that reflects this ordering. The intuition is similar to a setting with indifference curves where the utility values of the indifference curves represent the preference ordering. Drawing the analogy further, in a consumption setting we usually introduce a utility function that associates to any bundle of goods the numerical value of the indifference curve through this bundle. In our setting, bundles are represented by utility profiles $\mathbf{u}$. As such, we would like to associate to every utility profile the value of the Pareto frontier through this profile. The following theorem is a formalization of this intuition.

Theorem 3.2. *Let the collection of utility possibility sets $(C(\phi))_{\phi \in \Phi}$ satisfy Assumption A. Then:*

1. *the collection of utility possibility sets $(C(\phi))_{\phi \in \Phi}$ satisfies the NIP if and only if there exists a function $\rho : \Phi \rightarrow \mathbb{R}$ such that for all $\phi, \psi \in \Phi$:*

$$C(\phi) \sqsubseteq C(\psi) \leftrightarrow \rho(\phi) \leq \rho(\psi).$$

2. *if $(C(\phi))_{\phi \in \Phi}$ also satisfies Assumption B; then $(C(\phi))_{\phi \in \Phi}$ satisfies the NIP if and only if there exists a weakly increasing function $h : \mathcal{D} \rightarrow \mathbb{R}$ and a function $\rho : \Phi \rightarrow \mathbb{R}$ that satisfies Part 1 of the theorem and such that for all $\mathbf{u} \in \mathcal{D}, \phi \in \Phi$:*

$$\mathbf{u} \in C(\phi) \leftrightarrow h(\mathbf{u}) \leq \rho(\phi).$$

and $\mathbf{u} \in \partial C(\phi)$, if and only if:

$$h(\mathbf{u}) = \rho(\phi).$$

3. *if $(C(\phi))_{\phi \in \Phi}$ satisfies also Assumptions B and C; then the functions ρ and h from Part 2 can be chosen to be continuous.*
4. *if there exists a continuous function $h : \mathbb{R}^N \rightarrow \mathbb{R}$ and a function $\rho : \Phi \rightarrow \mathbb{R}$ such that for all $\mathbf{u} \in \mathbb{R}^N$ and $\phi \in \Phi$:*

$$h(\mathbf{u}) \leq \rho(\phi) \leftrightarrow \mathbf{u} \in C(\phi),$$

then $(C(\phi))_{\phi \in \Phi}$ satisfies the NIP.

Parts 1, 2 and 3 of Theorem 3.2 provide various representation results for the NIP condition for various combinations of assumptions on the utility possibility sets. If $(C(\phi))_{\phi \in \Phi}$ satisfies the NIP, then we call the function h of this representation a *Pareto function* and we call $\rho(\phi)$ a *surplus function*.

Utility profiles $\mathbf{u}$ with higher values of the Pareto function $h(\mathbf{u})$ lie on Pareto frontiers further away from the origin. These are Pareto frontiers $\partial C(\phi)$ with larger surplus values

$\rho(\phi)$. Observe that both the Pareto and surplus functions are only unique up to a (continuous) monotone transformation. In other words, if (h, ρ) is a representation of $(C(\phi))_{\phi \in \Phi}$ and if g is a (continuous) increasing function $\mathbb{R} \rightarrow \mathbb{R}$ then $(\tilde{h}, \tilde{\rho})$, where:

$$\tilde{h} = g(h) \text{ and } \tilde{\rho} = g(\rho),$$

is also a representation of $(C(\phi))_{\phi \in \Phi}$. Finally, Part 4 of Theorem 3.2 gives a useful sufficient condition for a collection of utility possibility sets to satisfy the NIP.

Theorem 3.2 shows that, given Assumptions A and B, any Pareto frontier $\partial C(\phi)$ can be formalized as the set of vectors $\mathbf{u} \in \mathcal{D}$ for which:

$$h(\mathbf{u}) = \rho(\phi).$$

If the Pareto frontier is strictly decreasing, it is possible to invert h with respect to u_1 and derive the following representation of the Pareto frontier $\partial C(\phi)$:

$$\mathbf{u} \in \partial C(\phi) \leftrightarrow u_1 = \gamma(\rho(\phi), u_2, \dots, u_N).$$

Here, $\gamma(\rho(\phi), u_2, \dots, u_N) \equiv h^{-1}(\rho(\phi); u_2, \dots, u_N)$ is decreasing in $u_2, \dots, u_N$ and increasing in $\rho(\phi)$. In general, absent the NIP condition, any strictly decreasing Pareto frontier can be represented by a function:

$$u_1 = \tau(\phi, u_2, \dots, u_N),$$

which is also decreasing in $u_2, \dots, u_N$. This shows that for the NIP condition to hold, there should exist a function γ (increasing in its first argument and decreasing in its last $N - 1$ arguments) and a surplus function ρ such that:

$$\tau(\phi, u_2, \dots, u_N) \equiv \gamma(\rho(\phi), u_2, \dots, u_N).$$

This requires that the Pareto frontier function $\tau(\phi, u_2, \dots, u_N)$ is weakly separable in ϕ . A necessary condition for weak separability is that, for all $u_2, \dots, u_N, v_2, \dots, v_N$ and all ϕ and ψ :

$$\text{if } \tau(\phi, u_2, \dots, u_N) \geq \tau(\psi, u_2, \dots, u_N) \text{ then, } \tau(\phi, v_2, \dots, v_N) \geq \tau(\psi, v_2, \dots, v_N).$$

If the function τ is C^2 , i.e. 2 times continuously differentiable, then a local condition is given by the Leontief-Sono conditions (Leontief, 1947; Sono, 1961): for all $\ell, j \leq k$ and all $i = 2, \dots, N$:

$$\frac{\partial}{\partial u_i} \left(\frac{\frac{\partial \tau(\phi, u_2, \dots, u_N)}{\partial \phi_\ell}}{\frac{\partial \tau(\phi, u_2, \dots, u_N)}{\partial \phi_j}} \right) = 0,$$

where ϕ_ℓ and ϕ_j are the i th and j th component of the vector ϕ . This provides an easily testable restriction for the NIP given knowledge of the Pareto frontier function τ . We refer to Blackorby, Primont, and Russell (1978) for an in depth study of the property of (weak) separability. We now return to our previous example and show that the NIP holds.

Example (example 1 continued). *For all values $(\delta, \xi) \in \Phi$ we had that $(u, \pi) \in C(\delta, \xi)$ if and only if,*

$$\ell(u, P - \pi) \leq 1 - \omega(\delta, \xi),$$

If we define the Pareto function $h(u, \pi) \equiv \ell(u, P - \pi)$ and the surplus function $\rho(\delta, \xi) \equiv 1 - \omega(\delta, \xi)$ we obtain that, by part 4 of Theorem 3.2, $(C(\delta, \xi))_{(\delta, \xi) \in \Phi}$ satisfies the NIP.

4 Transferable utility

Transferable utility (TU) is a widely used assumption in economics. A model is said to have TU whenever Pareto frontiers are parallel lines with slope equal to minus one. Of course, utility values are ordinal, so we should allow this property to hold after taking some continuous and increasing transformations of the utilities. The main message of this section is that the NIP is strictly weaker than TU. In other words, although settings with TU satisfy the NIP, the reverse is not necessarily true. The following gives a formal definition of transferable utility.

Definition 4.1 (Transferable utility (TU)). *The collection of utility possibility sets $(C(\phi))_{\phi \in \Phi}$ satisfies transferable utility if and only if there exist continuous and increasing functions $g_i : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ and a function $\kappa : \Phi \rightarrow \mathbb{R}$ such that for all $\phi \in \Phi$,*

$$\mathbf{u} \in C(\phi) \leftrightarrow \sum_{i=1}^N g_i(u_i) \leq \kappa(\phi).$$

In order to see how this relates to the usual definition of TU, let us define $y_i = g_i(u_i)$. Then Definition 4.1 holds if the equation for the Pareto frontier $\partial C(\phi)$ takes the following form:

$$\sum_{i=1}^N y_i = \kappa(\phi), \text{ with } y_i = g_i(u_i).$$

This defines a hyperplane in the transformed utility space, $(y_1, \dots, y_N)$, which is orthogonal to the unit vector. As expected, the TU property implies the NIP condition: if the Pareto curves are parallel hyperplanes, then they do not intersect.⁴ The following counterexample shows that not every setting with NIP also satisfies TU.

Example 2. *This example is borrowed from Bergstrom (2016). Let $a, b \in \mathbb{R}_+$ with $a < b$ and set $\Phi = [a, b]$. Define,*

$$C(\phi) = \{(u_1, u_2) \in \mathbb{R}_+^2 : u_1 + u_2 + u_1(u_2)^2 \leq \phi\}.$$

If we set $h(u_1, u_2) = u_1 + u_2 + u_1(u_2)^2$ and $\rho(\phi) = \phi$, we see that, by part 4 of Theorem 3.2, $(C(\phi))_{\phi \in \Phi}$ satisfies the NIP. However, the collection $(C(\phi))_{\phi \in \Phi}$ does not satisfy TU. If it

⁴This is an easy consequence of part 4 of Theorem 3.2 when defining the Pareto function as $h(\mathbf{u}) \equiv \sum_{i=1}^N g_i(u_i)$ and the surplus function by $\rho(\phi) \equiv \kappa(\phi)$.

does, then there should exist a function κ and two increasing, continuous functions g_1 and g_2 , such that for all $(u_1, u_2) \in \partial C(\phi)$:

$$\kappa(\phi) = \kappa(u_1 + u_2 + u_1(u_2)^2) = \kappa(h(u_1, u_2)) = g_1(u_1) + g_2(u_2).$$

As g_1 and g_2 and h are (weakly) increasing functions, it follows that κ is also an increasing function. By the equalities above, it must be that for all utility values u_1, u_2, v_1, v_2, z_1 and z_2 .⁵

$$\left. \begin{array}{l} h(u_1, u_2) \geq h(v_1, v_2) \\ h(v_1, z_2) \geq h(z_1, u_2) \end{array} \right\} \rightarrow h(u_1, z_2) \geq h(z_1, v_2)$$

This condition is known as the Thomson condition or the double cancellation axiom (see, for example, (Debreu, 1960)). This condition can be used to verify whether a function is additively separable.⁶ For this example, the Thomson condition is violated for the function $h(u_1, u_2) = u_1 + u_2 + u_1(u_2)^2$, e.g. choose $u_1 = 1, u_2 = 1, v_1 = 3, v_2 = 0, z_1 = 8$ and $z_2 = 2$.

Above example suggests that additive separability is important in order to relate the NIP to TU.

Definition 4.2. The Pareto function $h : \mathcal{D} \rightarrow \mathbb{R}$ is additively separable if there exist increasing, continuous functions $g_i : \mathbb{R}_+ \rightarrow \mathbb{R}$ and an increasing function $g : \mathbb{R} \rightarrow \mathbb{R}$ such that for all $\mathbf{u} \in \mathcal{D}$,

$$h(\mathbf{u}) = g \left(\sum_{i=1}^N g_i(u_i) \right).$$

The following theorem shows that TU is satisfied if and only if the NIP holds and the Pareto aggregation function is additively separable.

Theorem 4.3. Let $(C(\phi))_{\phi \in \Phi}$ satisfy Assumptions A and B. If $(C(\phi))_{\phi \in \Phi}$ satisfies the NIP, then TU is satisfied if and only if the Pareto aggregation function $h : \mathcal{D} \rightarrow \mathbb{R}$ is additively separable.

Example (Example 1 continued). We defined $h(u, \pi) \equiv \ell(u, P - \pi)$. A sufficient condition for h to be additively separable is that, $u(\ell, w)$ is additively separable. In this case,

$$u(\ell, w) = k_1(\ell) + k_2(w),$$

⁵In order to see that this condition holds, notice that $h(u_1, u_2) \geq h(v_1, v_2)$ and $h(v_1, z_2) \geq h(z_1, u_2)$ implies $g_1(u_1) + g_2(u_2) \geq g_1(v_1) + g_2(v_2)$ and $g_1(v_1) + g_2(z_2) \geq g_1(z_1) + g_2(u_2)$. Adding these two inequalities together and cancelling the common terms $g_1(v_1) + g_2(u_2)$, gives $g_1(u_1) + g_2(z_2) \geq g_1(z_1) + g_2(v_1)$, which implies $h(u_1, z_2) \geq h(z_1, v_2)$.

⁶An alternative, local differentiable condition is the Sono condition

$$\frac{\partial}{\partial u_2} \left(\ln \left(\frac{\frac{\partial h(u_1, u_2)}{\partial u_1}}{\frac{\partial h(u_1, u_2)}{\partial u_2}} \right) \right) = k(u_2),$$

for some function k of u_2 only.

for some increasing, continuous functions k_1 and k_2 . Then:

$$\ell(u, w) = k_1^{-1}(u - k_2(w)),$$

so:

$$h(u, \pi) = k_1^{-1}(u - k_2(P - \pi)),$$

which is indeed additively separable.

5 Applications

In this section, we look at several models for which TU turned out to be a key assumption, and we discuss their relationship to the NIP property. Section 5.1 looks at models of two-sided matching and principal agent models with asymmetric information. Section 5.2 studies the Coase Independence Property in light of the NIP and Section 5.3 looks at the Rotten Kid Theorem.

5.1 Two-sided matching and principal-agent models with asymmetric information

Two-sided matching models have found a wide range of applications in economics. In particular, there is a close theoretical connection between models of two-sided matching and principal agent models with asymmetric information. As an example⁷, Chade, Eeckhout, and Smith (2017) analyze sorting patterns in two sided matching and draw a direct connection with monotone comparative static results (such as those contained in Milgrom and Shannon (1994) and Topkis (1998)). More recently, Nöldeke and Samuelson (2018) expanded on this nice theoretical connection between models of two-sided matching and principal agent models with asymmetric information. In particular, the conditions that guarantee positive assortative matching in a one to one matching framework is very similar to a single crossing condition in a principal agent model (see Nöldeke and Samuelson (2018) for more details). The main contribution of this section is to show that under NIP, positive assortative matching (and the single crossing condition) holds if some monotonicity condition on the surplus function ρ is satisfied together with a submodularity condition on the Pareto function h and a supermodularity condition on the surplus function ρ .

Two-sided matching Let M and W be two sets of respectively men and women. A matching is a function $\sigma : M \cup W \rightarrow M \cup W$ such that:

- for all men $m \in M$, $\sigma(m) \in W \cup \{m\}$ and for all women $w \in W$, $\sigma(w) \in M \cup \{w\}$.
- if m is matched to w , i.e. $w = \sigma(m)$ then w is matched to m , i.e. $m = \sigma(w)$.

⁷For an earlier application of this relationship, see Tervio (2008), who studied an assignment model of CEO's and firms.

If $\sigma(m) = m$ or $\sigma(w) = w$ we say that the man m or woman w is single. We assume that each man $m \in M$, has a characteristic $\delta \in \Delta$ and each woman $w \in W$ is endowed with a characteristic $\xi \in \Xi$. Let $\Phi = \Delta \times \Xi$ be the set of possible characteristic combinations. If a man with characteristic δ matches to a woman with characteristic ξ , we assume that the possible utility values of the man and the woman are in the utility possibility set $C(\delta, \xi)$. Throughout this section, we assume that the collection $(C(\delta, \xi))_{(\delta, \xi) \in \Phi}$ satisfies the NIP together with Assumptions A and B, so we have a representation in terms of a Pareto function h and a surplus function ρ . Given this, we use the standard notion of stability:

Definition 5.1. *A matching $\sigma : M \cup W \rightarrow M \cup W$ is stable if there are utility values $(u_m, u_w)_{m \in M, w \in W}$ such that for all men $m \in M$ of characteristic δ , and women $w \in W$ of characteristic ξ :*

1. $u_m, u_w \geq 0$,
2. if $\sigma(m) = m$ then $u_m = 0$ and if $\sigma(w) = w$ then $u_w = 0$.
3. if $\sigma(m) = w$, then $h(u_m, u_w) = \rho(\delta, \xi)$.
4. if $\sigma(m) \neq w$, then there is no $(u'_m, u'_w) \gg (u_m, u_w)$ such that $h(u'_m, u'_w) \leq \rho(\delta, \xi)$.

The first two conditions impose individual rationality on the utility profile: condition 1 states that all agents should at least receive the utility that they obtain when being single while condition 2 normalizes the utility of being single to zero. The third condition imposes Pareto optimality on the allocation within a couple: if m is matched to w then their utility profile should be on their Pareto curve. Finally, the last condition requires that if m and w are not matched to each other, then it is impossible for both of them to obtain a higher level of utility by forming a couple. The existence of a stable matching has been established under very general conditions (Kaneko, 1982).

To formalize assortativeness in this matching framework, we assume that there is a ranking $\succ$ on the set of men-characteristics Δ and a ranking $\succ$ on the set of women-characteristics Ξ . For ease of notation, we use the same symbol for the two rankings, but it will be obvious from the context which ranking we are talking about. We are interested in the conditions on the primitives of the model i.e. the Pareto function h and the surplus function ρ , that lead to *positive assortative matching* (PAM) in the sense that women with higher characteristics are matched with men with higher characteristics.

Definition 5.2 (Positive assortative Matching). *A matching σ has positive assortative matching (PAM) if for all men m with characteristic δ , men m' with characteristic δ' , women w with characteristic ξ and women w' with characteristic ξ' : if*

$$w = \sigma(m), \quad w' = \sigma(m'),$$

and

$$\delta \succ \delta',$$

then it is not the case that

$$\xi' \succ \xi$$

The Generalized Increasing Difference (GID) condition of Legros and Newman (2007) characterizes PAM for a very general setting which includes the NIP. In particular, it is also valid in cases where the Pareto curves intersect. In this section, we will look at a set of stronger, but easily verifiable conditions that are sufficient for PAM in the case where NIP holds. As a first condition, we assume that the corresponding utility possibility sets satisfy the following monotonicity assumption:

Assumption D. *For all man-characteristics $\delta, \delta' \in \Delta$ and woman-characteristics $\xi, \xi' \in \Xi$,*

$$\text{if } \delta \succ \delta' \text{ and } \xi \succ \xi' \text{ then } \max\{\rho(\delta, \xi), \rho(\delta', \xi')\} > \max\{\rho(\delta', \xi), \rho(\delta, \xi')\}.$$

This assumption imposes that either the surplus of a higher type man and higher type woman is larger than an alternative match or the surplus of the lower type man going with the lower type woman is greater than the surplus in an alternative match. This condition can be compared with co-ranking as introduced by Legros and Newman (2010, Definition 3) in the study of assortative matching in the non-transferable utility setting.⁸ Next we impose a supermodularity condition on the surplus function ρ together with a submodularity condition on the Pareto function h .

Assumption E. *The surplus function ρ is supermodular and the Pareto function h is submodular with one of the two strict. Formally, for all characteristics, $\delta, \delta' \in \Delta, \xi, \xi' \in \Xi$ with $\delta \succ \delta'$ and $\xi \succ \xi'$:*

$$\rho(\delta, \xi) + \rho(\delta', \xi') \geq \rho(\delta, \xi') + \rho(\delta', \xi),$$

and for all $u_m > u_{m'}$ and $u_w > u_{w'}$,

$$h(u_m, u_w) + h(u_{m'}, u_{w'}) \leq h(u_{m'}, u_w) + h(u_m, u_{w'}).$$

where one of these inequalities is strict.

We remark that the properties of sub- and supermodularity are not invariant up to ordinal (i.e. monotone) transformations. On the other hand, as noticed in the previous section, the Pareto and surplus function are only defined up to a common ordinal transformation. Given this, it suffices that the desired modularity properties are satisfied for at least one of these transformations.

The following proposition shows that under NIP, assumptions D and E are sufficient for a stable matching to exhibit PAM.

Proposition 5.3. *If the Pareto and surplus functions satisfy Assumptions D and E, then any stable matching will satisfy PAM.*

⁸Co-ranking is a strict weakening of monotonicity as introduced by Becker (1973) and is equivalent to GID as introduced by Legros and Newman (2007) in the limiting case when utilities are strictly non-transferable.

The proof of Proposition 5.3 is simple enough, so we present it here in the main text. Towards a contradiction, let us assume that Assumptions D and E are satisfied and that there are men m, m' with characteristics δ, δ' and women w, w' with characteristics ξ, ξ' and a stable matching, σ such that $w' = \sigma(m)$, $w = \sigma(m')$, $\delta \succ \delta'$ and $\xi \succ \xi'$. Let $u_m, u_{m'}, u_w$ and $u_{w'}$ be the utility values assigned to the agents in this stable matching. Given that the matching is stable, we must have that:

$$\begin{aligned} h(u_m, u_{w'}) &= \rho(\delta, \xi'), & h(u_{m'}, u_w) &= \rho(\delta', \xi), \\ h(u_m, u_w) &\geq \rho(\delta, \xi), & h(u_{m'}, u_{w'}) &\geq \rho(\delta', \xi'). \end{aligned}$$

Without loss of generality, assume that $\max\{\rho(\delta, \xi), \rho(\delta', \xi')\} = \rho(\delta, \xi)$.⁹ Then condition D gives:

$$\begin{aligned} h(u_m, u_w) &\geq \rho(\delta, \xi) > \rho(\delta', \xi) = h(u_{m'}, u_w), \\ h(u_m, u_w) &\geq \rho(\delta, \xi) > \rho(\delta, \xi') = h(u_m, u_{w'}) \end{aligned}$$

Given that h is weakly increasing, we have that $u_m > u_{m'}$ and $u_w > u_{w'}$. But then, by (strict) supermodularity of ρ :

$$\begin{aligned} h(u_{m'}, u_w) + h(u_m, u_{w'}) &= \rho(\delta', \xi) + \rho(\delta, \xi'), \\ &\leq (<) \rho(\delta, \xi) + \rho(\delta', \xi'), \\ &\leq h(u_m, u_w) + h(u_{m'}, u_{w'}), \end{aligned}$$

which contradicts (strict) submodularity of h . A proof along the same lines also shows that any stable matching exhibits negative assortative matching whenever Assumption D is replaced by:

$$\min\{\rho(\delta, \xi), \rho(\delta', \xi')\} < \min\{\rho(\delta', \xi), \rho(\delta, \xi')\},$$

and ρ is (strictly) submodular while h is (strictly) supermodular.

It is interesting to contrast the result in Proposition 5.3 with the special case where TU holds. In this case, the Pareto aggregation function takes the form $h(u_m, u_w) = u_m + u_w$, which is obviously submodular. Then, PAM is obtained whenever ρ is strictly supermodular.¹⁰

If the characteristics, δ and ξ are real numbers and if the surplus and Pareto functions are C^2 then PAM can be characterized by a differential condition (see Chade, Eeckhout, and Smith, 2017).¹¹ Towards this end, we start from the following condition for a profile (u_m, u_w) on the Pareto frontier $\partial C(\delta, \xi)$:

$$h(u_m, u_w) = \rho(\delta, \xi).$$

⁹If the maximum is equal to $\rho(\delta', \xi')$ then we obtain $u_m < u_{m'}$ and $u_w < u_{w'}$, which leads to the same contradiction.

¹⁰It can be shown that in this case, Assumption D is superfluous.

¹¹We thank an anonymous referee for pointing this out to us.

Inverting h gives an expression for the utility of the man as a function of the characteristics and the utility of the woman:

$$u_m = \gamma(\rho(\delta, \xi); u_w) \equiv \mu(\delta, \xi; u_w).$$

In the differential case, the GID condition takes on the following form (Chade, Eeckhout, and Smith, 2017):

$$\mu_{\delta, \xi} > \frac{\mu_{\xi}}{\mu_{u_w}} \mu_{\delta, u_w},$$

where we use sub-indices to denote partial derivatives. Using the identity $\gamma(\rho(\delta, \xi); u_w) \equiv \mu(\delta, \xi; u_w)$, this gives:

$$\gamma_{\rho, \rho} \rho_{\delta} \rho_{\xi} + \gamma_{\rho} \rho_{\delta, \xi} > \frac{\gamma_{\rho} \rho_{\xi}}{\gamma_{u_w}} \gamma_{\rho, u_w} \rho_{\delta}.$$

Using the implicit function theorem on the identity $h(\gamma(\rho; u_w), u_w) \equiv \rho$, we can substitute out the partial derivatives of γ to obtain the following condition:

$$\rho_{\delta, \xi} > \frac{h_{u_m, u_w}}{h_{u_w} h_{u_m}} \rho_{\xi} \rho_{\delta}. \quad (1)$$

In order to see the connection between this condition and Proposition 5.3 above, notice that Assumption D is related to the setting where either $\rho_{\xi}, \rho_{\delta} > 0$ or $\rho_{\xi}, \rho_{\delta} < 0$. Also, for Assumption E, supermodularity of ρ relates to $\rho_{\delta, \xi} > 0$, while submodularity of h relates to $h_{u_m, u_w} < 0$. Next, if h is monotone, then $h_{u_m}, h_{u_w} > 0$. Given these sign restrictions, the left hand side of (1) is positive while its right hand side is negative. As such, the differential condition of PAM will be satisfied.

Also, as noticed above, Assumptions (D) and (E) only need to be satisfied for one particular ordinal transformation of the functions h and ρ . To see the implication of this, consider transforming h and ρ by a common increasing, C^2 function g :

$$\tilde{h}(u_m, u_w) = g(h(u_m, u_w)), \text{ and } \tilde{\rho}(\delta, \xi) = g(\rho(\delta, \xi)).$$

Then taking derivatives gives:

$$\begin{aligned} \tilde{h}_{u_m, u_w} &= g'' h_{u_m} h_{u_w} + g' h_{u_m, u_w}, \\ \tilde{\rho}_{\delta, \xi} &= g'' \rho_{\delta} \rho_{\xi} + g' \rho_{\delta, \xi}. \end{aligned}$$

where g' is the first derivative of g and g'' is its second order derivative. The function $\tilde{h}$ is submodular if $\tilde{h}_{u_m, u_w} \leq 0$ while $\tilde{\rho}$ is supermodular if $\tilde{\rho}_{\delta, \xi} \geq 0$. Given that $g' > 0$ and if $\rho_{\delta} \cdot \rho_{\xi} > 0$, this gives the condition:

$$\frac{h_{u_m, u_w}}{h_{u_m} h_{u_w}} \leq -\frac{g''}{g'} \leq \frac{\rho_{\delta, \xi}}{\rho_{\delta} \rho_{\xi}}.$$

By taking a transformation g where $g'' < 0$, this allows the Pareto function h to be non submodular as long as $\rho_{\delta, \xi}$ is sufficiently large. On the other hand, if $g'' > 0$ then ρ is allowed to be non-supermodular as long as h_{u_m, u_w} is sufficiently negative.

Example (Example 1 continued). We have that $h(u, \pi) = \ell(u, P - \pi)$ and $\rho(\delta, \xi) = 1 - \omega(\delta, \xi)$. The function $\rho(\delta, \xi)$ is supermodular if $\omega(\delta, \xi)$ is submodular. Next, the function h is submodular if,

$$h_{u, \pi} \leq 0.$$

It can be verified that this is equal to the following condition on the utility function,

$$u_{\ell, w} \leq u_{\ell, \ell} \frac{u_w}{u_\ell}.$$

If the utility function is concave, then the right hand side will be negative. As such, this condition holds only if the cross partial derivative of $u(\ell, w)$ is sufficiently negative. A particular class of utility functions that satisfies this condition is:

$$u(\ell, w) = \frac{w^{1-\gamma}}{1-\gamma} \exp((\gamma-1)k(\ell)),$$

where $\gamma > 1$ and $k', k'' < 0$. It can be shown that u is increasing in both ℓ and w . Also, u is concave if $\gamma k'' > (1-\gamma)(k')^2$. Similar types of utility functions can be found in, for example, King, Plosser, and Rebelo (2002) or Blundell, Costa-Dias, Meghir, and Shaw (2016). Notice that u is not consistent with TU, even though as we have shown, it does fit our NIP framework.

Principal-agent models with adverse selection Formally, there is a close connection between the two sided matching model and the principal-agent model with adverse selection (see Chade et al. (2017) and the recent contribution of Nöldeke and Samuelson (2018)). Preserving the notation from the previous part, consider a principal-agent model where agents are endowed with a type $\delta \in \Delta$, and can take an action $\xi \in \Xi$. There is the possibility of a (monetary) transfer v from the agent to the principal. The utility function of the agent is a function of her type δ , her action ξ , and the transfer v to the principal:

$$u = \mu(\delta, \xi, v).$$

It is assumed that μ is strictly decreasing in v . This fits the NIP framework if μ satisfies the following functional specification:

$$u = \mu(\delta, \xi, v) = \gamma(\rho(\delta, \xi), v). \quad (2)$$

Indeed, if we invert γ with respect to its last argument and define $h \equiv \gamma^{-1}$, we obtain $h(u, v) = \rho(\delta, \xi)$, which gives the NIP setting. One particular case is the quasi-linear specification:

$$u = \rho(\delta, \xi) - v,$$

which gives $h(u, v) \equiv u + v$. This corresponds to the TU case.

An important property in models with adverse selection is the strict single crossing condition.

Definition 5.4. *The utility function $\gamma(\rho(\delta, \xi); v)$ satisfies the strict single crossing property if for all $\delta, \delta' \in \Delta$ and $\xi, \xi' \in \Xi$, if $\delta \succ \delta'$ and $\xi \succ \xi'$, then for all v, v' ,*

$$\gamma(\rho(\delta', \xi), v) \geq \gamma(\rho(\delta', \xi'), v'),$$

implies:

$$\gamma(\rho(\delta, \xi), v) > \gamma(\rho(\delta, \xi'), v').$$

The strict single crossing condition states that if a ‘low’ type agent weakly prefers the ‘high’ action contract (ξ, v) over the ‘low’ action contract (ξ', v') then a ‘high’ type agent strictly prefer the ‘high’ action contract (ξ, v) over the ‘low’ action contract (ξ', v') .

Introducing the utility values $u' = \gamma(\rho(\delta', \xi'), v')$ and $u = \gamma(\rho(\delta, \xi), v)$, we have that the strict single crossing condition can be rewritten in the following way.

Definition 5.5 (Strict single crossing condition). *The functions (h, ρ) satisfy the strict single crossing property if for all $\delta \succ \delta'$ and $\xi \succ \xi'$, if:*

$$h(u', v') = \rho(\delta', \xi') \text{ and } h(u, v) = \rho(\delta, \xi),$$

then,

$$\rho(\delta', \xi) \geq h(u', v) \text{ implies } h(u, v') > \rho(\delta, \xi').$$

Similar to the proof of Proposition 5.3, one can easily show that this strict single-crossing property is satisfied if Assumptions D and E hold.

Proposition 5.6. *If h and ρ satisfy Assumptions (D) and (E), then the strict single crossing condition is satisfied and consequently, a ‘high’ type agent will choose a ‘high’ action contract.*

This result handles settings beyond the quasi-linear specification as shown in the following example.

Example 3. *Assume that the utility function of the agent can be written in the following way:*

$$\beta(-e^{-u} + 1) = e^{\alpha v} \cdot (\rho(\delta, \xi) - v). \quad (3)$$

where $\alpha \geq 0$, $\beta > 0$ and $\alpha\beta < 1$. Observe that the left hand side is a monotone transformation of u so its ordering perfectly reflects the utility ranking. Also, notice that for $\alpha = 0$, the specification corresponds to the quasi-linear TU case. However, if $\alpha > 0$ the income effect is not constant. Inverting the right hand side (3) gives:

$$h(u, v) \equiv \beta(-e^{-u} + 1)e^{-\alpha v} + v.$$

One can easily verify that $h_u > 0$, $h_v > 0$ and that $h_{u,v} < 0$, so h is a submodular and monotone function. As such, in order for the strict single-crossing condition to be satisfied, it suffices that ρ is supermodular and satisfies Assumption D.

To show that ‘high’ type agents will choose a ‘high’ action contract, we combine the strict single crossing condition with incentive compatibility constraints. In particular, consider a setting with two agents of types δ and δ' . Then the incentive compatibility constraints for the contracts (ξ', v') , (ξ, v) , offered to agents of types δ' and δ respectively, take the form:

$$\gamma(\rho(\delta, \xi), v) \geq \gamma(\rho(\delta, \xi'), v'), \quad (4)$$

$$\gamma(\rho(\delta', \xi'), v') \geq \gamma(\rho(\delta', \xi), v). \quad (5)$$

Condition (4) states that the δ -type agent weakly prefers the contract (ξ, v) over the contract (ξ', v') while condition (5) states that the δ' -type agent weakly prefers the contract (ξ', v') over the contract (ξ, v) . By defining the utility values $u = \gamma(\rho(\delta, \xi), v)$ and $u' = \gamma(\rho(\delta', \xi'), v')$ and inverting γ , we obtain the equivalent conditions:

$$h(u, v') \geq \rho(\delta, \xi'), \quad (6)$$

$$h(u', v) \geq \rho(\delta', \xi). \quad (7)$$

It is well known that if the strict single crossing condition holds, then the incentive compatibility constraints require that higher types must take contracts with higher actions. If we take a setting where Assumptions D and E are satisfied, this can be shown in a very simple way. Towards a contradiction, assume that $\delta \succ \delta'$ but $\xi' \succ \xi$ (i.e. higher type agents take lower actions) and the incentive compatibility constraints are satisfied. Concerning Assumption D, let us consider the case where $\max\{\rho(\delta, \xi'), \rho(\delta', \xi)\} = \rho(\delta, \xi')$. The other case is analogous. Then, by (6):

$$\begin{aligned} h(u, v') &\geq \rho(\delta, \xi') > \rho(\delta', \xi') = h(u', v') \text{ and } , \\ h(u, v') &\geq \rho(\delta, \xi') > \rho(\delta, \xi) = h(u, v), \end{aligned}$$

which, by monotonicity of h implies that $u > u'$ and $v' > v$. However, then:

$$\begin{aligned} h(u, v) + h(u', v') &\geq (>) h(u', v) + h(u, v'), \\ &\geq \rho(\delta, \xi') + \rho(\delta', \xi), \\ &\geq (>) \rho(\delta, \xi) + \rho(\delta', \xi'). \end{aligned}$$

The first inequality follows from the (strict) submodularity of the function h . The second inequality follows from the incentive compatibility constraints (6) and (7). Finally the last inequality is due to the (strict) supermodularity of ρ . By definition of u and u' , however, $h(u, v) = \rho(\delta, \xi)$ and $h(u', v') = \rho(\delta', \xi')$, which gives the desired contradiction.

5.2 The Coase independence property

One of the key economic conjectures, first suggested in the article by Coase (1960), is that when agents are allowed to bargain freely, and in the absence of transaction costs, they will reach a Pareto efficient outcome and the Pareto efficient amount of externalities

(either positive or negative) will be independent of the particular way bargaining rights or endowments are allocated. In this section, we will mainly focus on the second part of this conjecture, saying that the particular way liabilities or rights are given to the various agents in the economy should not influence the Pareto efficient level of the externality. This result was never formalized by Coase in his original article, but since then, there have been several attempts to provide formal interpretations to this Coase Independence Property.

Special interest has been given to a *characterization* of the set of preferences that allow for the Coasian Independence to be satisfied. Important contributions in this direction have been given by Hurwicz (1995) and Bergstrom (2017). These papers start from a general equilibrium framework in which N agents consume a private good and a vector of public goods.¹² Their setting also assumes that there exists a cost function $c(\mathbf{y})$ that specifies the amount of money necessary to produce the bundle of public goods $\mathbf{y}$, and that there is a fixed amount of aggregate resources, I , which can be allocated to either private consumption or the production of the public goods. Given this, they specify the following feasibility constraint for the economy as a whole:

$$\sum_{i=1}^N x_i + c(\mathbf{y}) \leq I, \quad (8)$$

which has the particular feature that it is linear in the private goods x_i and additively separable between the private and public goods.

The results in Hurwicz (1995) and Bergstrom (2017) show that TU is necessary and sufficient to obtain the Coase Independence Property. In other words, in this type of economy, the Pareto efficient level of public goods is independent of the way endowments are allocated among the agents if and only if there is transferable utility among the agents.

In the present paper, we are interested in studying the Coase independence property in settings where the feasibility constraint is more general than this linear-separable specification. We show that for more general feasibility constraints, it is not TU but rather NIP that characterizes the set of economies for which Coasian Independence holds.

Setting We consider a general equilibrium model with N agents who consume one private good and $m \geq 2$ public goods.¹³ Agent i has an initial endowment $\omega_i > 0$ of the private good. We denote the final amount of the private good going to individual i by x_i and we represent by $\mathbf{x} \in \mathbb{R}_+^N$, the allocation of the private good over the N agents. The vector $\mathbf{y} \in \mathbb{R}_+^m$ specifies the bundle of public goods.

Each agent $i \leq N$ has a utility function $u_i(x_i, \mathbf{y})$ that depends on her own consumption of the private good and the total consumption of public goods. For an allocation $(\mathbf{x}, \mathbf{y}) \in$

¹²(Hurwicz, 1995) allows for positive and negative externalities, whereas (Bergstrom, 2017) considers a more standard public goods setting. From a theoretical viewpoint, externalities and public goods can be treated in a very similar way. In this paper, we focus on the public goods setting.

¹³The result easily generalizes to the case where agents are allowed to consume more than one private good.

$\mathbb{R}_+^{N+m}$ we denote by $\mathbf{u}(\mathbf{x}, \mathbf{y})$ the utility profile $\mathbf{v}$ where for all agents i : $v_i = u_i(x_i, \mathbf{y})$. In other words, $\mathbf{u}(\mathbf{x}, \mathbf{y})$ is the utility profile that results from the allocation $(\mathbf{x}, \mathbf{y})$.

In contrast to the aforementioned literature on the Coase Independence Property, we allow for a general feasibility constraint, which takes the following form:

$$f(\mathbf{x}, \mathbf{y}) \leq 0. \quad (9)$$

The standard linear-separable setting given by (8) is a special case of (9) where $f(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^N x_i + c(\mathbf{y}) - I$. We denote by $\mathcal{Y}$ the feasible set values of the public goods bundle that satisfy individual rationality:

$$\mathcal{Y} = \{\mathbf{y} \in \mathbb{R}^m : \exists \mathbf{x} \in \mathbb{R}_+^N, f(\mathbf{x}, \mathbf{y}) \leq 0 \text{ and } \mathbf{u}(\mathbf{x}, \mathbf{y}) \geq \mathbf{0}\}.$$

For each $\mathbf{y} \in \mathcal{Y}$, we can define the utility possibility set $C(\mathbf{y})$ as the set of all feasible and individual rational utility profiles, provided that the vector of public goods is equal to $\mathbf{y}$:

$$C(\mathbf{y}) = \{\mathbf{v} \in \mathbb{R}_+^N : \exists \mathbf{x} \in \mathbb{R}_+^N, f(\mathbf{x}, \mathbf{y}) \leq 0 \text{ and } \mathbf{v} \leq \mathbf{u}(\mathbf{x}, \mathbf{y})\}.$$

Observe that individual rationality is guaranteed by the assumption that $\mathbf{v} \geq \mathbf{0}$. Finally, we denote by $\mathcal{P}|_Y$ the set of all Pareto efficient allocations when the public goods are restricted to lie in the set $Y \subseteq \mathcal{Y}$.

$$\mathcal{P}|_Y = \left\{ (\mathbf{x}, \mathbf{y}) \in \mathbb{R}_+^{N+m} : \begin{array}{l} \mathbf{y} \in Y, \\ \mathbf{u}(\mathbf{x}, \mathbf{y}) \in C(\mathbf{y}), \\ \nexists \mathbf{v} \in \bigcup_{\mathbf{y} \in Y} C(\mathbf{y}) : \mathbf{v} \gg \mathbf{u}(\mathbf{x}, \mathbf{y}) \end{array} \right\}.$$

We say that the Coase Independence Property holds if every Pareto efficient allocation involves the same levels of public goods.

Definition 5.7. *The collection of utility possibility sets $(C(\mathbf{y}))_{\mathbf{y} \in \mathcal{Y}}$ satisfies the Coase Independence Property if for all subsets $Y \subseteq \mathcal{Y}$ and all bundles of public goods $\mathbf{y} \in Y$, if $(\mathbf{x}, \mathbf{y}) \in \mathcal{P}|_Y$ and $(\mathbf{x}', \mathbf{y}) \in \mathcal{P}|_{\{\mathbf{y}\}}$, then $(\mathbf{x}', \mathbf{y}) \in \mathcal{P}|_Y$.*

The definition has the following intuition. Let Y be a subset of $\mathcal{Y}$ and consider an allocation $(\mathbf{x}, \mathbf{y})$ that is Pareto efficient in the set of all allocations for which the public goods are restricted to the set Y . Now, take any other allocation $(\mathbf{x}', \mathbf{y})$ with the same amount of public goods which is Pareto efficient but now only over the set of feasible allocations where the level of public goods is restricted to be equal to $\mathbf{y}$. Then if the Coase Independence Property holds, we have that $(\mathbf{x}', \mathbf{y})$ is also Pareto efficient over the bigger set of allocations.

Negating this condition, we obtain that if an allocation $(\mathbf{x}, \mathbf{y}) \in \mathcal{P}|_{\{\mathbf{y}\}}$ is not Pareto efficient over the set $\bigcup_{\mathbf{y} \in Y} C(\mathbf{y})$, then none of the allocations with public goods equal to $\mathbf{y}$ is Pareto efficient over this set. Effectively, every Pareto efficient allocation should have the same amounts of public goods.

Before stating our main result, we impose some further conditions on the utility functions and the feasibility constraint.

Assumption F.

1. The utility functions $u_i : \mathbb{R}_+^{1+m} \rightarrow \mathbb{R}$ are continuous, increasing and are normalized by setting $u_i(\omega_i, \mathbf{0}) = \underline{u}_i = 0$.
2. The feasibility function $f : \mathbb{R}_+^{N+m} \rightarrow \mathbb{R}$ of equation (9) is continuous and increasing.
3. For all agents $i \leq N$, there is an amount $\bar{x}_i$ of private goods such that $f(\bar{x}_i \mathbf{e}_i, \mathbf{0}) > 0$, where $\mathbf{e}_i$ is the N dimensional vector with a 1 at place i and zeros everywhere else.
4. For all $\mathbf{y} \in \mathcal{Y}$, $u_i(0, \mathbf{y}) < u_i(\omega_i, \mathbf{0}) = 0$.

Assumptions F.1 and F.2 impose some regularity conditions on the utility and transformation functions. Assumption F.3 requires that there is a level of private consumption for agent i that is not feasible, even when the consumption of all other goods is zero. It guarantees that the amounts of private goods $\mathbf{x}$ that are feasible are bounded from above. Finally, Assumption F.4 states that if an agent receives no private goods, then she prefers to break down the negotiations. Also, the condition implies that every feasible and individual rational allocation $(\mathbf{x}, \mathbf{y})$ (i.e. with $f(\mathbf{x}, \mathbf{y}) \leq 0$ and $\mathbf{u}(\mathbf{x}, \mathbf{y}) \geq \mathbf{0}$) must have strictly positive amounts of the private goods. This allows us to exclude Pareto efficient allocations where some agents have zero private consumption.

The following theorem shows that the Coase Independence Property holds if and only if the NIP holds.

Theorem 5.8. *Assume that Assumption F is satisfied. Then, the collection of utility possibility sets $(C(\mathbf{y}))_{\mathbf{y} \in \mathcal{Y}}$ satisfies the NIP if and only if the Coase Independence Property is satisfied.*

From Theorem 4.3, we know that if $(C(\mathbf{y}))_{\mathbf{y} \in \mathcal{Y}}$ satisfies NIP, it does not necessarily satisfy TU. Therefore, the result contained in Theorem 5.8 provides a novel characterization for the Coase Independence Property, for cases where the feasibility constraint is more general than the linear-separable case (8).

That the NIP implies Coasian Independence is quite intuitive. The sufficiency proof of Theorem 5.8 is illustrated in Figure 2. The argument is by contradiction. Assume that the Coase Independence Property holds but NIP is violated. In the figure, this is illustrated by showing two distinct Pareto curves $\partial C(\mathbf{y}')$ and $\partial C(\mathbf{y})$ that intersect at $\mathbf{u}(\mathbf{x}, \mathbf{y})$. As this utility profile is on both Pareto curves, it follows that:

$$(\mathbf{x}, \mathbf{y}) \in \mathcal{P}|_{\{\mathbf{y}, \mathbf{y}'\}}. \quad (10)$$

As the two Pareto curves are distinct, it is possible to show that there is an allocation $(\mathbf{x}', \mathbf{y})$ such that $\mathbf{u}(\mathbf{x}', \mathbf{y}) \in \partial C(\mathbf{y})$ but $\mathbf{u}(\mathbf{x}', \mathbf{y})$ is strictly below the Pareto curve $\partial C(\mathbf{y}')$.¹⁴ From this, it follows that:

$$(\mathbf{x}', \mathbf{y}) \in \mathcal{P}|_{\{\mathbf{y}\}} \text{ and } (\mathbf{x}', \mathbf{y}) \notin \mathcal{P}|_{\{\mathbf{y}'\}}. \quad (11)$$

¹⁴If the two Pareto curves don't intersect, but only touch, the argument is basically the same, but one might have to exchange $\mathbf{y}$ with $\mathbf{y}'$.

Figure 2 Illustration of the proof of Theorem 5.8

However, from the Coase Independence Property, condition (10) and the first part of (11), it follows that:

$$(\mathbf{x}, \mathbf{y}) \in \mathcal{P}|_{\{\mathbf{y}, \mathbf{y}'\}} \text{ and } (\mathbf{x}', \mathbf{y}) \in \mathcal{P}|_{\{\mathbf{y}\}} \rightarrow (\mathbf{x}', \mathbf{y}) \in \mathcal{P}|_{\{\mathbf{y}, \mathbf{y}'\}}.$$

which contradicts the second part of (11).

We remark that Theorem 5.8 provides an easy way to compute the set of Pareto optimal levels of public goods under the NIP setting. From Theorem 3.2, we know that NIP guarantees the existence of a surplus function $\rho : \mathcal{Y} \rightarrow \mathbb{R}$ that ranks the Pareto curves.¹⁵ In addition, it is easy to show that, given Assumption F, the utility possibility correspondence $C(\cdot)$ is upper hemicontinuous¹⁶ which allows us to choose $\rho(\cdot)$ to be upper semicontinuous. Given this, for any compact subset $Y \subseteq \mathcal{Y}$, we can define the set:

$$Y' = \arg \max_{\mathbf{y} \in Y} \rho(\mathbf{y}).$$

It then follows that:

$$\mathcal{P}|_Y = \bigcup_{\mathbf{y} \in Y'} \mathcal{P}|_{\{\mathbf{y}\}} = \mathcal{P}|_{Y'}.$$

In other words, the Pareto-efficient levels of public goods can be obtained by maximising the surplus function $\rho(\mathbf{y})$.

5.3 Becker's Rotten Kid Theorem

The Rotten Kid Theorem (Becker, 1974) is a variation on the Coasian Independence result. It describes the following two stage game. In the first period, each individual from a group

¹⁵Lemma A.4 in the appendix shows that for our model, the collection $(C(\mathbf{y}))_{\mathbf{y} \in \mathcal{Y}}$ satisfies Assumption A which explains why we don't need to list it as a separate assumption in Theorem 5.8.

¹⁶A proof of this is available from the authors upon request.

of agents (kids) takes an action. This profile of actions determines the level of public goods. In stage two of the game, another agent, called the parent, takes the actions of the kids as given and allocates to each kid a feasible amount of the private good. The objective function of the parent is given by a welfare function which is strictly increasing in the utility levels of all the kids. The Rotten Kid Theorem describes a situation where each allocation (over private and public goods) that is welfare maximizing for the parent can be implemented as a subgame perfect Nash equilibrium of this two stage game. In other words, the level of public goods, collectively chosen by the kids, coincides with the optimal level of public goods desired by the parent. This result is highly important given that it provides precise conditions to obtain a first best allocation (level) of public and private goods, even in the absence of any sort of pre-commitment. Furthermore, the Rotten Kid Theorem also has implications for the analysis of household decisions (see for example (Becker, 1974) and Chiappori and Mazzocco (2017)). In particular, it can be used as a justification for ‘income pooling’, which implies that a redistribution of resources within the household doesn’t affect the household’s choices, thereby providing a foundation for the standard ‘unitary’ model of the household.¹⁷ Previous theoretical literature on the subject showed that under some minor conditions on the preferences of the parent and if the utility possibility sets of the kids satisfy TU, then the Rotten Kid Theorem is satisfied (Bergstrom, 1989; Cornes and Silva, 1999; Chiappori and Werning, 2002). All of these papers departed from the linear-separable specification (8) to determine the set of feasible allocations. Furthermore, in such an environment TU is also necessary to obtain a Rotten Kid Theorem.¹⁸ In this section, we will show that the Rotten Kid Theorem holds under the weaker NIP condition, even if we relax the feasibility restriction to (9).

In order to formalize things, we borrow the notation and concepts developed in the previous section. We assume that there are N kids and there is an agent, the parent, whose objective function is given by a welfare function $W : \mathbb{R}_+^N \rightarrow \mathbb{R}$ that takes as arguments the utility profiles for the kids. It is assumed that W strictly increases in all arguments. The rotten kid game has the following structure. First, each kid selfishly chooses an action $a_i \in A_i$, where A_i is a set of feasible actions for kid i (which coincides with their strategy sets). Let us denote the chosen profile of actions by $\mathbf{a} \in \prod_{i=1}^N A_i$. The action profile where kid i chooses a_i and all other kids choose according to the profile $\mathbf{a}$ is denoted by $(a_i, \mathbf{a}_{-i})$. We assume that the profile of actions uniquely determines the level of public goods $\mathbf{y}$ via some surjective function $g : \prod_{i=1}^N A_i \rightarrow \mathcal{Y}$,

$$g(\mathbf{a}) = \mathbf{y}.$$

For a given profile of actions $\mathbf{a}$ and an associated level of public goods $\mathbf{y} = g(\mathbf{a})$, a strategy for the parent consists in choosing a feasible vector of private quantities for every possible action profile of the kids. This can be modelled by a function $s : \prod_{i=1}^N A_i \rightarrow \mathbb{R}_+^N$, where

¹⁷In this argument, the interaction between kids and the parent are replaced by the interaction between the spouses.

¹⁸More precisely, Bergstrom (1989) uses the term ‘conditional TU’, which means that the Pareto frontiers are linear with slope -1 conditional on the set of actions taken by the kids.

for all action profiles $\mathbf{a}$:

$$\mathbf{u}(s(\mathbf{a}), g(\mathbf{a})) \geq \mathbf{0} \text{ and } f(s(\mathbf{a}), g(\mathbf{a})) \leq 0.$$

The first constraint imposes individual rationality while the second requires the vector $s(\mathbf{a})$ to be technologically feasible given the level of public goods $g(\mathbf{a})$. Let us denote by $\mathcal{S}$ the set of all strategies feasible for the parent (i.e. satisfying individual rationality and technological feasibility).

In the second stage of the game, the parent chooses from $\mathcal{S}$ a strategy that maximizes her utility:

$$s^*(\mathbf{a}) \in \arg \max_{\mathbf{x}} W(\mathbf{u}(\mathbf{x}, g(\mathbf{a}))) \text{ s.t. } f(\mathbf{x}, g(\mathbf{a})) \leq 0, \text{ and } \mathbf{u}(\mathbf{x}, g(\mathbf{a})) \geq \mathbf{0}.$$

Notice that our technological feasibility constraint subsumes the ‘standard’ case in which the action profile determines the aggregate income to be distributed by the parent, that is, $f(\mathbf{x}, g(\mathbf{a})) = \sum_{i=1}^N x_i - I(\mathbf{a}) \leq 0$. Now, let $\mathcal{S}^*$ be the set of all optimal strategies of the parent.

A subgame perfect Nash equilibrium for the rotten kid game then consists of a strategy for the parent $s^* \in \mathcal{S}^*$ together with a profile of actions $\mathbf{a}^*$ for the kids such that each kid chooses an action that maximizes her utility, given $s^*(\cdot)$ and given the actions chosen by the other kids, $\mathbf{a}_{-i}^*$:

$$\forall i \leq N, \forall a_i \in A_i : u_i(s_i^*(\mathbf{a}^*), g(\mathbf{a}^*)) \geq u_i(s_i^*(a_i, \mathbf{a}_{-i}^*), g(a_i, \mathbf{a}_{-i}^*)),$$

where $s_i^*(\mathbf{a})$ is the i -th component of $s^*(\mathbf{a})$. For later reference, we will denote with SPN the set of all subgame perfect Nash equilibria $(s^*, \mathbf{a}^*)$ of the rotten kid game. A feature of a subgame perfect equilibrium in this setting is that the parent takes the level of public goods as given. Ideally, however, the parent would like to decide on both the allocation of private goods, $\mathbf{x}$ and the level of public goods (or, equivalently, the actions chosen by the kids). This ‘first best solution’ for the parent is given by the solution to the following optimization program:

$$\max_{\mathbf{x}, \mathbf{a}} W(\mathbf{u}(\mathbf{x}, g(\mathbf{a}))) \text{ s.t. } f(\mathbf{x}, g(\mathbf{a})) \leq 0 \text{ and } \mathbf{u}(\mathbf{x}, g(\mathbf{a})) \geq \mathbf{0}.$$

Denote by FB the set of all first best allocations of private goods and kids’ actions. We say that the Rotten Kid Property holds if any first best allocation can also be reached as an outcome of a subgame perfect Nash equilibrium. In our setting, this leads to the following definition.

Definition 5.9 (Rotten Kid Property). *We say that the Rotten Kid Property holds if, for all $(\mathbf{x}, \mathbf{a}) \in FB$, there is a subgame perfect Nash equilibrium $(s^*, \mathbf{a}^*) \in SPNE$ such that $\mathbf{a}^* = \mathbf{a}$ and $\mathbf{x} = s^*(\mathbf{a}^*)$.*

Most of the papers in the literature studying the Rotten Kid Theorem assume a TU setting. Under TU, the utility possibility sets $C(g(\mathbf{a}))$ are hyperplanes with slope -1 . In

Figure 3 Illustration of the proof of Proposition 5.10

such settings, it is found to be crucial that the utility of each kid is a normal good for the parent (e.g. Bergstrom (1989); Cornes and Silva (1999) and Chiappori and Werning (2002)). In the present setting, we drop TU, but also impose an assumption akin to normality with respect to the kids' utilities:

Assumption G. For all optimal strategies of the Parent $s^* \in \mathcal{S}^*$, and all action profiles $\mathbf{a}, \mathbf{a}' \in \prod_{i=1}^N A_i$:

$$\text{if } C(g(\mathbf{a})) \subseteq C(g(\mathbf{a}')), \text{ then } \mathbf{u}(s^*(\mathbf{a}'), g(\mathbf{a}')) \geq \mathbf{u}(s^*(\mathbf{a}), g(\mathbf{a})).$$

This assumption states that if the utility possibility set expands then the optimal choice of the parent is to (weakly) increase the utility level of all kids. The following shows that under this condition, NIP leads to the Rotten Kid Property.

Proposition 5.10. If Assumption G is satisfied and if the collection $(C(\mathbf{y}))_{\mathbf{y} \in \mathcal{Y}}$ satisfies the NIP, then the Rotten Kid Property holds.

The proof of Proposition 5.10 is illustrated in Figure 3 for a setting with two kids. Consider a first best allocation $(\mathbf{x}, \mathbf{a})$. Then clearly $\mathbf{x}$ is a best response for the parent conditional on kid's action profile $\mathbf{a}$. As such, there is a strategy $s^* \in \mathcal{S}^*$ such that:

$$s^*(\mathbf{a}) = \mathbf{x}.$$

The resulting profile $\mathbf{u}(s^*(\mathbf{a}), g(\mathbf{a}))$ is indicated in Figure 3. Observe that, by Assumption G and the NIP, we have that for any other action profile $\mathbf{a}'$ either $\mathbf{u}(s^*(\mathbf{a}'), g(\mathbf{a}'))$ is above the utility profile $\mathbf{u}(s^*(\mathbf{a}), g(\mathbf{a}))$, i.e. in the grey region, or it is below the utility profile $\mathbf{u}(s^*(\mathbf{a}), g(\mathbf{a}))$, i.e. in the arched region.

Now, towards a contradiction, if $(s^*, \mathbf{a})$ is not a subgame perfect Nash equilibrium, then there is at least one kid i that has a profitable deviation, say $a'_i \in A_i$ and:

$$u_i(s^*(\mathbf{a}), g(\mathbf{a})) < u_i(s^*(a'_i, \mathbf{a}_{-i}), g(a'_i, \mathbf{a}_{-i})).$$

Relating back to Figure 3, this means that the utility profile $\mathbf{u}(s^*(a'_i, \mathbf{a}_{-i}), g(a'_i, \mathbf{a}_{-i}))$ cannot be in the arched region. Given that it is either in the arched or grey region, it should therefore be in the grey region and, hence, dominate the profile $\mathbf{u}(s^*(\mathbf{a}), g(\mathbf{a}))$. However, this means that $(\mathbf{x}, \mathbf{a})$ was not a first best solution, which gives the desired contradiction.

6 Conclusion

This paper introduces a weakening of the well-known transferable utility property, namely the case where any two distinct Pareto curves have an empty intersection. We call this the no-intersection property or NIP. We showed that under some mild regularity condition any setting with NIP has a representation in the form of a Pareto aggregation function and a surplus function. We showed that NIP is strictly weaker than transferable utility (TU). In particular, the TU property holds if and only if the Pareto function is additively separable.

Next, we looked at some models where the TU property has turned out to be an important. We showed that in a two-sided matching model, positive assortative matching is obtained if some monotonicity condition on the surplus function holds, if the Pareto aggregation function is submodular and the surplus function is supermodular. We then showed that the NIP is useful to simplify the study of the analytical relationships between two-sided matching models and principal agent models. Finally, we revisited the Coase Independence Property and the Rotten Kid Theorem. In contrast to the previous literature, we allowed for a more general environment in which the technology (set of feasible allocations) is not necessarily linear (additively separable). We showed that in this generalized case the Coase Independence Property is characterized by NIP and that the Rotten Kid Theorem can be retrieved under the weaker condition of NIP (compared to the usual TU assumption.)

References

- Becker, G. S., 1973. A theory of marriage: Part I. *Journal of Political Economy* 81, 813–846.
- Becker, G. S., 1974. A theory of social interactions. *Journal of Political Economy* 82, 1063–1093.
- Bergstrom, T., 2016. Lecture notes on separable preferences. Tech. rep.
- Bergstrom, T. C., 1989. A fresh look at the rotten-kid theorem and other household mysteries. *Journal of Political Economy* 97, 1138–1159.
- Bergstrom, T. C., 2017. When was coase right? Tech. rep.
- Blackorby, C., Primont, D., Russell, R. R., 1978. *Duality, Separability and Functional Structure: Theory and Economic Applications*. Elsevier, New York.

- Blundell, Costa-Dias, Meghir, Shaw, 2016. Female labor supply, human capital and welfare reform. *Econometrica* 84, 1705–1753.
- Chade, H., Eeckhout, J., Smith, L., 2017. Sorting through search and matching models in economics. *Journal of Economic Literature* 55, 493–544.
- Chiappori, P.-A., 2017. Matching with transfers. Princeton University Press.
- Chiappori, P.-A., Costa Dias, M., Meghir, C., 2018. The marriage market, labor supply and education choice. *Journal of Political Economy* 126, S1, S26–S72.
- Chiappori, P.-A., Mazzocco, M., 2017. Static and intertemporal household decisions. *Journal of Economic Literature* 55, 985–1045.
- Chiappori, P.-A., Werning, I., 2002. Comment on "rotten kids, purity, and perfection". *Journal of Political Economy* 110, 475–480.
- Choo, E., Siow, A., 2006. Who marries whom and why. *Journal of Political Economy* 114, 175–201.
- Coase, R. H., 1960. The problem of social cost. *Journal of Law and Economics* 3, 1–44.
- Cornes, R. C., Silva, E. C. D., 1999. Rotten kids, purity and perfection. *Journal of Political Economy* 107, 1034–1040.
- de Clippel, G., 2008. An axiomatization of the inner core using appropriate reduced games. *Journal of Mathematical Economics* 44, 316–323.
- Debreu, G., 1960. *Mathematical Methods in the Social Sciences*. Stanford University Press, Ch. Topological Methods in Cardinal Utility Theory, pp. 16–26.
- Gale, D., Shapley, L. S., 1962. College admission and the stability of marriage. *The American Mathematical Monthly* 69, 9–15.
- Hicks, J. R., 1939. The foundations of welfare economics. *The Economic Journal* 49, 696–712.
- Hurwicz, L., 1995. What is the coase theorem? *Japan and the World Economy* 7, 49–74.
- Kaldor, N., 1939. Welfare propositions of economics and interpersonal comparisons of utility. *The Economic Journal* 49, 549–552.
- Kaneko, M., 1982. The central assignment game and the assignment market. *Journal of Mathematical Economics* 10, 205–232.
- King, R. G., Plosser, C. I., Rebelo, S. T., 2002. Production, growth and business cycles: Technical appendix. *Computational Economics* 20, 87–116.

- Legros, P., Newman, A., 2007. Beauty is a beast, frog is a prince: Assortative matching with nontransferabilities. *Econometrica* 75, 1073–1102.
- Legros, P., Newman, A., 2010. Co-ranking mates: Assortative matching in marriage markets. *Economics Letters* 106, 177–179.
- Leontief, W., 1947. Introduction to a theory of the internal structure of functional relationships. *Econometrica* 15, 361–373.
- Milgrom, P., Shannon, C., 1994. Monotone comparative statics. *Econometrica* 62, 157–180.
- Nöldeke, G., Samuelson, L., 2018. The implementation duality. *Econometrica* 86, 1283–1324.
- Scitovsky, T., 1941. A note on welfare propositions in economics. *Review of Economic Studies* 9, 77–88.
- Shapley, L., Shubik, M., 1972. The assignment game I: The core. *International Journal of Game Theory* 1, 111–130.
- Sono, M., 1961. The effect of price changes on the demand and supply of separable goods. *International Economic Review* 2, 239–271.
- Tervio, M., 2008. The difference that ceos make: An assignment model approach. *American Economic Review* 98, 642–668.
- Topkis, D. M., 1998. *Supermodularity and Complementarity*. Princeton University Press.

A Proofs of the Theorems

Proof of Theorem 3.2

We first proof the following two lemmata

Lemma A.1. *Given Assumption A, a collection of utility possibility sets $(C(\phi))_{\phi \in \Phi}$ satisfies the NIP if and only if for all environmental variables $\phi, \psi \in \Phi$:*

$$C(\phi) \setminus C(\psi) \neq \emptyset \text{ implies } C(\psi) \sqsubset C(\phi).$$

Proof. ($\rightarrow$) The proof is by contradiction. Assume that $(C(\phi))_{\phi \in \Phi}$ satisfies the NIP but $C(\phi) \setminus C(\psi) \neq \emptyset$ and $C(\psi) \not\sqsubset C(\phi)$.

Let us first show that there exist utility profiles $\mathbf{u}$ and $\mathbf{v}$ such that:

$$\mathbf{u} \in C(\phi) \setminus C(\psi), \tag{12}$$

$$\mathbf{v} \in C(\psi) \setminus C(\phi). \tag{13}$$

As $C(\phi) \setminus C(\psi) \neq \emptyset$ we know there is an $\mathbf{u} \in C(\phi) \setminus C(\psi)$ which shows (12). Also, as $C(\psi) \not\subseteq C(\phi)$ we know that there exists a profile $\mathbf{w} \in C(\psi)$ such that for all $\mathbf{z} \gg \mathbf{w}$, $\mathbf{z} \notin C(\phi)$. There are several cases to consider:

- $\mathbf{w} \notin C(\phi)$.

In this case, (13) is satisfied by setting $\mathbf{v} = \mathbf{w}$.

- $\mathbf{w} \in C(\phi)$.

– If there is a $\mathbf{z} \gg \mathbf{w}$ such that $\mathbf{z} \in C(\psi)$, then also $\mathbf{z} \notin C(\phi)$, so in this case (13) is satisfied with $\mathbf{v} = \mathbf{z}$.

– If for all $\mathbf{z} \gg \mathbf{w}$, $\mathbf{z} \notin C(\psi)$.

Then $\mathbf{w} \in \partial C(\phi)$ and $\mathbf{w} \in \partial C(\psi)$. As such, from the NIP, we have that $C(\phi) = C(\psi)$ which contradicts (12).

For a vector $\mathbf{w} > \mathbf{0}$ and a $\chi \in \Phi$, define the correspondence:

$$\Gamma(\mathbf{w}, \chi) = \{\gamma \in \mathbb{R}_+ : \gamma \mathbf{w} \in \partial C(\chi)\}.$$

In Appendix B, we show that $\Gamma(\mathbf{w}, \chi)$, as a function of $\mathbf{w}$, is non-empty, convex valued and upper hemicontinuous on $\mathcal{D} \setminus \{\mathbf{0}\}$.

From (12) and (13), we have that, $\mathbf{u}, \mathbf{v} > \mathbf{0}$ and that:

$$1 \leq \Gamma(\mathbf{u}, \phi), \quad \Gamma(\mathbf{v}, \phi) \leq 1, \quad (14)$$

$$1 \leq \Gamma(\mathbf{v}, \psi), \quad \Gamma(\mathbf{u}, \psi) \leq 1 \quad (15)$$

Let:

$$G(\theta) = \Gamma((1 - \theta)\mathbf{u} + \theta\mathbf{v}, \phi) - \Gamma((1 - \theta)\mathbf{u} + \theta\mathbf{v}, \psi).$$

Here, for two sets A and B , the set $A - B$ contains all elements $c = a - b$ where $a \in A$ and $b \in B$. The correspondence $G(\theta)$ is non-empty, upper hemicontinuous and convex valued on $[0, 1]$. Also, by (14) and (15), $G(0)$ is bounded from below by 0 and $G(1)$ is bounded from above by 0. By the intermediate value theorem for correspondences (see de Clippel (2008, Lemma 2)), there should be a θ^* such that:

$$0 \in G(\theta^*).$$

Let $\mathbf{w} = (1 - \theta^*)\mathbf{u} + \theta^*\mathbf{v} > \mathbf{0}$. Then:

$$0 \in \Gamma(\mathbf{w}, \psi) - \Gamma(\mathbf{w}, \phi).$$

By the definition of Γ , we have that there is a γ such that:

$$\gamma \mathbf{w} \in \partial C(\phi) \cap \partial C(\psi).$$

From the NIP, it follows that $C(\phi) = C(\psi)$, a contradiction.

($\leftarrow$) The proof is again by contradiction. Assume that $\{C(\phi)\}_{\phi \in \Phi}$ satisfies the second part of the lemma but the NIP does not hold. This means that there is a utility profile:

$$\mathbf{u} \in \partial C(\phi) \cap \partial C(\psi) \text{ and } C(\phi) \neq C(\psi).$$

Without loss of generality, assume that $C(\phi) \setminus C(\psi) \neq \emptyset$. Then, the second part of the lemma tells us that $C(\psi) \sqsubset C(\phi)$. As $\mathbf{u} \in C(\psi)$, there must be a profile $\mathbf{w} \gg \mathbf{u}$ such that $\mathbf{w} \in C(\phi)$. This contradicts the fact that $\mathbf{u} \in \partial C(\phi)$. $\square$

Lemma A.2. *Let the collection of utility possibility sets $(C(\phi))_{\phi \in \Phi}$ satisfy Assumption A. Then $(C(\phi))_{\phi \in \Phi}$ satisfies the NIP if and only if the binary relation $\sqsubseteq$ is a complete ordering (reflexive and transitive) on $(C(\phi))_{\phi \in \Phi}$ with asymmetric part $\sqsubset$, i.e. $A \sqsubset B$ if and only if $A \sqsubseteq B$ and not $B \sqsubseteq A$.*

Proof. ($\rightarrow$) Let us first show that $\sqsubseteq$ is complete. Let $(C(\phi))_{\phi \in \Phi}$ satisfy the NIP and let $\phi, \psi \in \Phi$.

- If $C(\phi) = C(\psi)$, then immediately, $C(\phi) \sqsubseteq C(\psi)$.
- if $C(\phi) \neq C(\psi)$ then either $C(\phi) \setminus C(\psi) \neq \emptyset$ or $C(\psi) \setminus C(\phi) \neq \emptyset$. In the first case, by Lemma A.1, $C(\psi) \sqsubseteq C(\phi)$. In the second, again by Lemma A.1, $C(\psi) \sqsubseteq C(\phi)$.

Next, let us show that $\sqsubseteq$ is transitive. Let $C(\phi) \sqsubseteq C(\psi) \sqsubseteq C(\xi)$. If either the first or the second comparison is an equality, then immediately $C(\phi) \sqsubseteq C(\xi)$. Else, we have:

$$C(\phi) \sqsubset C(\psi) \sqsubset C(\xi).$$

From this, we see that for all $\mathbf{u} \in C(\phi)$ there is a $\mathbf{v} \gg \mathbf{u}$ such that $\mathbf{v} \in C(\psi)$ and for all $\mathbf{v} \in C(\psi)$ there is a $\mathbf{w} \gg \mathbf{v}$ such that $\mathbf{w} \in C(\xi)$. This immediately implies that for all $\mathbf{u} \in C(\phi)$ there is a $\mathbf{w} \gg \mathbf{u}$ such that $\mathbf{w} \in C(\xi)$. Conclude that $C(\phi) \sqsubset C(\xi)$, what we needed to show.

To show that $\sqsubset$ is the asymmetric part of $\sqsubseteq$, assume, towards a contradiction, that:

$$C(\phi) \sqsubseteq C(\psi), \tag{16}$$

and:

$$C(\psi) \sqsubset C(\phi). \tag{17}$$

Condition (17) implies that for all $\mathbf{v} \in \partial C(\psi)$ there is an $\mathbf{u} \gg \mathbf{v}$ such that $\mathbf{u} \in C(\phi)$. This shows that $C(\phi) \neq C(\psi)$. As such, (16) can be strengthened to $C(\phi) \sqsubset C(\psi)$. But then there is a $\mathbf{w} \gg \mathbf{u} \gg \mathbf{v}$ such that $\mathbf{w} \in C(\psi)$. This contradiction with $\mathbf{v} \in \partial C(\psi)$.

($\leftarrow$) The proof is by contradiction. Let $\sqsubseteq$ be an ordering on $(C(\phi))_{\phi \in \Phi}$ with asymmetric part $\sqsubset$. Let:

$$\mathbf{u} \in \partial C(\phi) \cap \partial C(\psi),$$

and assume, towards a contradiction, that $C(\phi) \neq C(\psi)$. Without loss of generality, let $C(\phi) \setminus C(\psi) \neq \emptyset$. Then, as $\sqsubset$ is complete, $C(\psi) \sqsubset C(\phi)$. Given that $\mathbf{u} \in C(\psi)$ there should be a $\mathbf{v} \gg \mathbf{u}$ such that $\mathbf{v} \in C(\phi)$. However, this contradicts $\mathbf{u} \in \partial C(\phi)$. $\square$

We are ready for the proof of Theorem 3.2.

Part 1

Assume that $(C(\phi))_{\phi \in \Phi}$ satisfies Assumption A.

($\rightarrow$) Assume that $(C(\phi))_{\phi \in \Phi}$ satisfies the NIP. We need to construct a function $\rho : \Phi \rightarrow \mathbb{R}$ such that:

$$C(\phi) \sqsubseteq C(\psi) \leftrightarrow \rho(\phi) \leq \rho(\psi).$$

Define the function:

$$\rho(\phi) = \max\{\|\mathbf{u}\| : \mathbf{u} \in C(\phi)\}, \quad (18)$$

where $\|\cdot\|$ is the usual L^2 norm: $\|\mathbf{u}\| = \sqrt{\sum_{i=1}^N (u_i)^2}$. The function $\rho(\phi)$ is well defined as $C(\phi)$ is compact and non-empty for all $\phi \in \Phi$ (by Assumption A).

Let us first show that $\rho(\phi) \leq \rho(\psi)$ implies $C(\phi) \sqsubseteq C(\psi)$. The proof is by contradiction. Let:

$$\rho(\phi) \leq \rho(\psi), \quad (19)$$

and assume, towards a contradiction, that:

$$C(\phi) \not\sqsubseteq C(\psi). \quad (20)$$

By the NIP condition and Lemma A.2, we have that (20) is equivalent to $C(\psi) \sqsubset C(\phi)$. Take a utility profile $\mathbf{u} \in C(\psi)$ such that $\rho(\psi) = \|\mathbf{u}\|$. The existence of $\mathbf{u}$ is guaranteed by definition (18). Then as $C(\psi) \sqsubset C(\phi)$ there is a $\mathbf{v} \gg \mathbf{u}$ with $\mathbf{v} \in C(\phi)$. However, this implies that:

$$\rho(\psi) = \|\mathbf{u}\| < \|\mathbf{v}\| \leq \rho(\phi),$$

which contradicts (19).

Next, we show that $C(\phi) \sqsubseteq C(\psi)$ implies $\rho(\phi) \leq \rho(\psi)$. The proof is by contradiction. Assume that:

$$C(\phi) \sqsubseteq C(\psi) \quad (21)$$

and, towards a contradiction:

$$\rho(\phi) > \rho(\psi). \quad (22)$$

Consider a profile $\mathbf{v} \in C(\phi)$ such that $\rho(\phi) = \|\mathbf{v}\|$. Then, by (22), we know that $\rho(\psi) < \|\mathbf{v}\|$. This implies that $\mathbf{v} \in C(\phi) \setminus C(\psi)$. By the NIP and Lemma A.1 it follows that $C(\psi) \sqsubset C(\phi)$, which contradicts the fact that $\sqsubseteq$ is an ordering with asymmetric part $\sqsubset$ (see Lemma A.2).

($\leftarrow$) Assume that:

$$\rho(\phi) \leq \rho(\psi) \leftrightarrow C(\phi) \sqsubseteq C(\psi). \quad (23)$$

We need to show that $(C(\phi))_{\phi \in \Phi}$ satisfies the NIP. By Lemma A.1, it suffices to show that:

$$C(\phi) \setminus C(\psi) \neq \emptyset \text{ implies } C(\psi) \sqsubseteq C(\phi). \quad (24)$$

Towards this end, let $\mathbf{u} \in C(\phi) \setminus C(\psi)$. Clearly $C(\psi) \neq C(\phi)$. First observe that $C(\phi) \not\sqsubseteq C(\psi)$. As such, by contrapositive of (23) $\rho(\phi) > \rho(\psi)$. Again by (23), it follows that $C(\psi) \sqsubseteq C(\phi)$. Conclude that $C(\phi) \neq C(\psi)$ and $C(\psi) \sqsubseteq C(\phi)$ which demonstrates (24).

Part 2

($\rightarrow$) Let $(C(\phi))_{\phi \in \Phi}$ satisfy Assumptions A and B. We need to show that there exists a weakly increasing function $h : \mathcal{D} \rightarrow \mathbb{R}$ and a function $\rho : \Phi \rightarrow \mathbb{R}$ that satisfies:

$$C(\phi) \subseteq C(\psi) \leftrightarrow \rho(\phi) \leq \rho(\psi) \quad (25)$$

and such that for all $\mathbf{u} \in \mathcal{D}, \phi \in \Phi$:

$$\mathbf{u} \in C(\phi) \leftrightarrow h(\mathbf{u}) \leq \rho(\phi). \quad (26)$$

We consider the function $\rho : \Phi \rightarrow \mathbb{R}$ as defined in (18) above. From Part 1, we know that (25) is satisfied, so we only need to establish (26). Let $\mathbf{u} \in \mathcal{D}$. To define the value $h(\mathbf{u})$, notice that, by Assumption B, we know that there exists a $\phi \in \Phi$ such that $\mathbf{u} \in \partial C(\phi)$. Define:

$$h(\mathbf{u}) = \rho(\phi) \text{ where } \mathbf{u} \in \partial C(\phi). \quad (27)$$

In order to show that h is well defined, we need to show that (27) cannot assign two distinct values to $h(\mathbf{u})$. Indeed, if:

$$\mathbf{u} \in \partial C(\phi) \cap \partial C(\psi).$$

then by the NIP, $C(\phi) = C(\psi)$ and therefore, $\rho(\phi) = \rho(\psi)$. This shows that $h : \mathcal{D} \rightarrow \mathbb{R}$ is well defined.

Now, in order to establish (26), let us first show that for these specifications of ρ and h :

$$\mathbf{u} \in C(\phi) \rightarrow h(\mathbf{u}) \leq \rho(\phi).$$

Let $\mathbf{u} \in C(\phi)$. If $\mathbf{u} \in \partial C(\phi)$ then immediately, by definition of h , $h(\mathbf{u}) = \rho(\phi)$ and we are done. If $\mathbf{u} \notin \partial C(\phi)$ then by assumption B there is a ψ such that:

$$\mathbf{u} \in \partial C(\psi).$$

Also as $\mathbf{u} \in C(\phi)$ and $\mathbf{u} \notin \partial C(\phi)$ there is a utility profile $\mathbf{v} \gg \mathbf{u}$ such that $\mathbf{v} \in C(\phi)$. Notice that $\mathbf{v} \in C(\phi) \setminus C(\psi)$, so:

$$C(\phi) \setminus C(\psi) \neq \emptyset.$$

by the NIP and Lemma A.1, $C(\psi) \sqsubset C(\phi)$ and therefore, by Part 1 of the Theorem:

$$h(\mathbf{u}) = \rho(\psi) < \rho(\phi),$$

as was to be shown.

Next for the other part of (26), let us show that:

$$h(\mathbf{u}) \leq \rho(\phi) \rightarrow \mathbf{u} \in C(\phi).$$

Let $h(\mathbf{u}) \leq \rho(\phi)$. Let ψ be such that $\mathbf{u} \in \partial C(\psi)$, which exists by Assumption B. By definition of h :

$$h(\mathbf{u}) = \rho(\psi).$$

Then:

$$\rho(\psi) = h(\mathbf{u}) \leq \rho(\phi),$$

so, by Part 1 of the theorem $C(\psi) \subseteq C(\phi)$ and therefore $\mathbf{u} \in C(\psi) \subseteq C(\phi)$ as we needed to show.

($\leftarrow$) This is easy, as from Part 1 of the theorem, we already know that (25) implies that $(C(\phi))_{\phi \in \Phi}$ satisfies the NIP.

Part 3

Assume that $(C(\phi))_{\phi \in \Phi}$ satisfies Assumptions A, B and C. Let ρ be defined by (18) and let h be defined by (27) above. We would like to show that both ρ and h are continuous.

As ρ is defined by (18), its continuity follows from Berge's maximization theorem (as $C(\cdot)$ is continuous, we have that ρ is also continuous).

Showing continuity of h takes a bit more work. Let $\mathbf{u} \in \mathcal{D}$. For a sequence $(\mathbf{x}^t)_{t \in \mathbb{N}}$ let us write $\mathbf{x}^t \rightarrow \mathbf{x}$ if the sequence converges to $\mathbf{x}$. In order to show that h is continuous in $\mathcal{D}$ it suffices to show that it is continuous at every point $\mathbf{u} \in \mathcal{D}$. In other words, we need to show that:

$$\begin{aligned} &\text{for all sequences } (\mathbf{u}^t)_{t \in \mathbb{N}} \text{ in } \mathcal{D} \text{ with } \mathbf{u}^t \rightarrow \mathbf{u}, \\ &\text{there is a subsequence, } (\mathbf{w}^t)_{t \in \mathbb{N}} \text{ such that } h(\mathbf{w}^t) \rightarrow h(\mathbf{u}) \end{aligned} \quad (28)$$

This is what we are going to demonstrate.

Let $(\mathbf{u}^t)_{t \in \mathbb{N}}$ be a sequence in $\mathcal{D}$ such that $\mathbf{u}^t \rightarrow \mathbf{u}$. By Assumption B and the definition of h above, there exists a corresponding sequence $(\phi^t)_{t \in \mathbb{N}}$ in Φ such that for all $t \in \mathbb{N}$:

$$h(\mathbf{u}^t) = \rho(\phi^t).$$

As Φ is a compact subset of $\mathbb{R}^k$ (by Assumption C), the Bolzano-Weierstrass theorem tells us that the sequence $(\phi^t)_{t \in \mathbb{N}}$ has a convergent subsequence, say, $(\psi^t)_{t \in \mathbb{N}}$. Assume:

$$\psi^t \rightarrow \psi \in \Phi. \quad (29)$$

Further, as ρ is continuous, we have that:

$$\rho(\psi^t) \rightarrow \rho(\psi). \quad (30)$$

Let $(\mathbf{w}^t)_{t \in \mathbb{N}}$ be the subsequence of $(\mathbf{u}^t)_{t \in \mathbb{N}}$ that corresponds to the subsequence $(\psi^t)_{t \in \mathbb{N}}$ of $(\phi^t)_{t \in \mathbb{N}}$. In other words:

$$h(\mathbf{w}^t) = \rho(\psi^t).$$

By (30) have that:

$$h(\mathbf{w}^t) = \rho(\psi^t) \rightarrow \rho(\psi).$$

Let us show that:

$$\rho(\psi) = h(\mathbf{u}), \quad (31)$$

thereby establishing (28), what we wanted to proof.

First, observe that from the definitions of ρ and h , for all $t \in \mathbb{N}$:

$$\mathbf{w}^t \in \partial C(\psi^t) \subseteq C(\psi^t). \quad (32)$$

Given (30) and upper-hemicontinuity of $C(\cdot)$, there should be a convergent subsequence of $(\mathbf{w}^t)_{t \in \mathbb{N}}$ whose limit is in $C(\psi)$. This convergent subsequence of $(\mathbf{w}^t)_{t \in \mathbb{N}}$ is also a subsequence of $(\mathbf{u}^t)_{t \in \mathbb{N}}$, so its limit has to be equal to $\mathbf{u}$. Therefore:

$$\mathbf{u} \in C(\psi).$$

and consequentially:

$$h(\mathbf{u}) \leq \rho(\psi).$$

This gives one part of (31). We still need to show that $h(\mathbf{u}) \geq \rho(\psi)$. The proof is by contradiction. Assume that $h(\mathbf{u}) < \rho(\psi)$. Then by the definition of h , i.e. (27), we have that:

$$\mathbf{u} \notin \partial C(\psi).$$

Given that $\mathbf{u} \in C(\psi)$, there must be a profile $\mathbf{z}$ such that:

$$\mathbf{z} \gg \mathbf{u} \text{ and } \mathbf{z} \in C(\psi). \quad (33)$$

Take a number $\varepsilon > 0$ small enough such that:

$$\mathbf{z} \gg \mathbf{u} + \varepsilon \mathbf{1}, \quad (34)$$

where $\mathbf{1}$ is the vector of ones.

By (29) and (33) we have that $\psi^t \rightarrow \psi$ and $\mathbf{z} \in C(\psi)$, so by lower hemicontinuity of $C(\cdot)$ there is a number T , and a sequence $(\mathbf{z}^t)_{t \geq T}$ such that:

$$\mathbf{z}^t \rightarrow \mathbf{z} \text{ and for all } t \geq T, \mathbf{z}^t \in C(\psi^t). \quad (35)$$

Let $T \in \mathbb{N}$ be large enough such that for all agents $j \leq N$ and all $t \geq T$:

$$|z_j^t - z_j| < \varepsilon/2 \quad (36)$$

which can be found given that $\mathbf{z}^t \rightarrow \mathbf{z}$. Also, let T be large enough such that for all agents $j \leq N$ and all $t \geq T$:

$$|w_j^t - u_j| < \varepsilon/2, \quad (37)$$

which can be found given that $\mathbf{w}^t \rightarrow \mathbf{u}$. The latter is due to the fact that $(\mathbf{w}^t)_{t \in \mathbb{N}}$ is a subsequence of $(\mathbf{u}^t)_{t \in T}$ and that $\mathbf{u}^t \rightarrow \mathbf{u}$. Then, for all agents $j \leq N$ and all $t \geq T$:

$$\begin{aligned} z_j^t &\geq z_j - \varepsilon/2, \\ &\geq u_j - \varepsilon/2 + \varepsilon, \\ &> u_j + \varepsilon/2, \\ &\geq w_j^t. \end{aligned}$$

The first inequality is due to (36), the second is due to (34) and the last is due to (37). This shows that for t large enough:

$$\mathbf{z}^t \gg \mathbf{w}^t.$$

However, this contradicts with the fact that for all $t \in \mathbb{N}$: $\mathbf{z}^t \in C(\psi^t)$ (by (35)) and $\mathbf{w}^t \in \partial C(\psi^t)$ (by (32)). Conclude that $h(\mathbf{u}) = \rho(\psi)$.

Part 4

Assume that there is a continuous function $h : \mathcal{D} \rightarrow \mathbb{R}$ and a function $\rho : \Phi \rightarrow \mathbb{R}$ such that for all $u \in \mathcal{D}$ and $\phi \in \Phi$:

$$h(\mathbf{u}) \leq \rho(\phi) \leftrightarrow \mathbf{u} \in C(\phi). \quad (38)$$

We need to show that $(C(\phi))_{\phi \in \Phi}$ satisfies the NIP. Let us first show that:

$$\text{if } \mathbf{u} \in \partial C(\phi), \text{ then } h(\mathbf{u}) = \rho(\phi). \quad (39)$$

If $\mathbf{u} \in \partial C(\phi)$ then by (38), we know that $h(\mathbf{u}) \leq \rho(\phi)$. If, towards a contradiction $h(\mathbf{u}) < \rho(\phi)$ then, by continuity of h , there exists a $\varepsilon > 0$ such that:

$$h(\mathbf{u} + \varepsilon \mathbf{1}) < \rho(\phi).$$

But then, again by (38):

$$\mathbf{u} \ll \mathbf{u} + \varepsilon \mathbf{e} \in C(\phi),$$

which contradicts $\mathbf{u} \in \partial C(\phi)$.

Now, take a utility profile $\mathbf{u}$ such that:

$$\mathbf{u} \in \partial C(\phi) \cap \partial C(\psi).$$

By (39) above:

$$\rho(\phi) = h(\mathbf{u}) = \rho(\psi), \quad (40)$$

so for any $\mathbf{v} \in C(\phi)$ we have, by (38), that $h(\mathbf{v}) \leq \rho(\phi)$ and by (40) $\rho(\phi) = \rho(\psi)$ which means that $h(\mathbf{v}) \leq \rho(\psi)$. Invoking (38) once more shows that $\mathbf{v} \in C(\psi)$. As $\mathbf{v}$ was an arbitrary element of $C(\phi)$, we can conclude that $C(\phi) \subseteq C(\psi)$. A similar argument shows that $C(\psi) \subseteq C(\phi)$, so $C(\psi) = C(\phi)$, which means that the NIP condition is satisfied.

Proof of Theorem 4.3

Let $(C(\phi))_{\phi \in \Phi}$ satisfy the NIP and Assumptions A and B.

($\rightarrow$) Assume that TU is satisfied. We need to show that the Pareto function h is additively separable. Let us first show that:

$$\rho(\phi) \geq \rho(\psi) \leftrightarrow \kappa(\phi) \geq \kappa(\psi). \quad (41)$$

If $\rho(\phi) \geq \rho(\psi)$ then by Theorem 3.2 $C(\psi) \subseteq C(\phi)$. Take $\mathbf{u} \in \partial C(\psi)$. Then by continuity of the functions $g_i(\cdot)$ it follows that $\sum_{i=1}^N g_i(u_i) = \kappa(\psi)$. As $\mathbf{u} \in C(\phi)$ it follows by the definition of TU that:

$$\kappa(\psi) = \sum_{i=1}^N g_i(u_i) \leq \kappa(\phi).$$

For the reverse, if $\kappa(\phi) \geq \kappa(\psi)$ and $\mathbf{u} \in \partial C(\psi)$ then again by continuity of g_i , $\sum_{i=1}^N g_i(u_i) = \kappa(\psi)$. Then:

$$\sum_{i=1}^N g_i(u_i) = \kappa(\psi) \leq \kappa(\phi),$$

which means, by definition of TU, that $\mathbf{u} \in C(\phi)$. This means that $C(\psi) \subseteq C(\phi)$. By the NIP and Lemma A.1, we have that $C(\psi) \subseteq C(\phi)$. By Theorem 3.2, we have that:

$$\rho(\psi) \leq \rho(\phi).$$

Now, for the rest of the proof, if (41) holds then $\rho(\phi) = g(\kappa(\phi))$ for some increasing function $g(\cdot)$. Then, for $\mathbf{u} \in \partial C(\phi)$ we have that:

$$h(\mathbf{u}) = \rho(\phi) = g(\kappa(\phi)) = g\left(\sum_i^N g_i(u_i)\right),$$

so h is additively separable.

($\leftarrow$) Let h be additively separable. We need to show that TU is satisfied. By Theorem 3.2. $\mathbf{u} \in C(\phi)$ if and only if:

$$h(\mathbf{u}) = g\left(\sum_{i=1}^N g_i(u_i)\right) \leq \rho(\phi).$$

This shows that $(C(\phi))_{\phi \in \Phi}$ satisfies the definition of TU where we define $\kappa(\phi) = g^{-1}(\rho(\phi))$.

Proof of Theorem 5.8

We begin by proving the following Lemma.

Lemma A.3. *If $\mathbf{y} \in Y$ and $(\mathbf{x}, \mathbf{y}) \in \mathcal{P}|_Y$ then $\mathbf{u}(\mathbf{x}, \mathbf{y}) \in \partial C(\mathbf{y})$. On the other hand if $\mathbf{v} \in \partial C(\mathbf{y})$ then there is an allocation $(\mathbf{x}, \mathbf{y}) \in \mathcal{P}|_{\{\mathbf{y}\}}$ such that $\mathbf{v} = \mathbf{u}(\mathbf{x}, \mathbf{y})$.*

Proof. The first part is obvious from the definition of $\mathcal{P}|_Y$. For the second part, let:

$$\mathbf{v} \in \partial C(\mathbf{y}).$$

Then, by definition of $C(\mathbf{y})$, there is an allocation $(\mathbf{x}, \mathbf{y})$ such that:

$$f(\mathbf{x}, \mathbf{y}) \leq 0 \text{ and } \mathbf{v} \leq \mathbf{u}(\mathbf{x}, \mathbf{y})$$

Let us show that $\mathbf{v} = \mathbf{u}(\mathbf{x}, \mathbf{y})$ thereby finishing the proof.

The proof is by contradiction. Given that $\mathbf{v} \in \partial C(\mathbf{y})$ it is impossible that $\mathbf{v} \ll \mathbf{u}(\mathbf{x}, \mathbf{y})$. So if $\mathbf{v} \neq \mathbf{u}(\mathbf{x}, \mathbf{y})$ then $\mathbf{v} < \mathbf{u}(\mathbf{x}, \mathbf{y})$ and there should be an agent j such that:

$$v_j < u_j(x_j, \mathbf{y}).$$

From $u_j(x_j, \mathbf{y}) > 0$, we have that by Assumption F, $x_j > 0$. As such, by continuity of f there should exist a number $\varepsilon > 0$ and $\delta > 0$ such that:

$$v_j < u_j(x_j - \varepsilon, \mathbf{y}), \text{ and } f\left(\mathbf{x} - \varepsilon \mathbf{e}_j + \sum_{i \neq j} \delta \mathbf{e}_i, \mathbf{y}\right) \leq 0.$$

Here $\mathbf{e}_i$ is the vector that has a one at position i and zeros otherwise. Define:

$$\mathbf{x}' = \mathbf{x} - \varepsilon \mathbf{e}_j + \sum_{i \neq j} \delta \mathbf{e}_i.$$

Now, we have that $v_j < u_j(x'_j, \mathbf{y})$ and for all $i \neq j$:

$$u_i(x'_i, \mathbf{y}) = u_i(x_i + \delta, \mathbf{y}) > u_i(x_i, \mathbf{y}) = v_i.$$

Conclude that:

$$f(\mathbf{x}', \mathbf{y}) \leq 0 \text{ and } \mathbf{u}(\mathbf{x}', \mathbf{y}) \gg \mathbf{v} \geq \mathbf{0}.$$

This implies that $\mathbf{u}(\mathbf{x}', \mathbf{y}) \in C(\mathbf{y})$ and $\mathbf{u}(\mathbf{x}, \mathbf{y}) \gg \mathbf{v}$, which contradicts the assumption that $\mathbf{v} \in \partial C(\mathbf{y})$. $\square$

Lemma A.4. *The collection $(C(\mathbf{y}))_{\mathbf{y} \in \mathcal{Y}}$ satisfies Assumption A.*

Proof. Let $\mathbf{y} \in \mathcal{Y}$. Let us first show that $C(\mathbf{y})$ is non-empty. By definition of $\mathcal{Y}$, there is a $(\mathbf{x}, \mathbf{y})$ such that $f(\mathbf{x}, \mathbf{y}) \leq 0$ and $\mathbf{u}(\mathbf{x}, \mathbf{y}) \geq \mathbf{0}$, so $\mathbf{u}(\mathbf{x}, \mathbf{y}) \in C(\mathbf{y})$ which means that $C(\mathbf{y})$ is non-empty.

Next, let us show that $C(\mathbf{y})$ is bounded. Let $\mathbf{v} \in C(\mathbf{y})$. Then there is a vector $\mathbf{x}$, such that $f(\mathbf{x}, \mathbf{y}) \leq 0$ and for all agents $\mathbf{v} \leq \mathbf{u}(\mathbf{x}, \mathbf{y})$. Also, by Assumption F, we have that for all agents $i \leq N$, there is a $\bar{x}_i$ such that $f(\bar{x}_i \mathbf{e}_i, 0) > 0$. As f is increasing in all components, this implies that $x_i \leq \bar{x}_i$ and consequentially, $\mathbf{v} \leq \mathbf{u}(\bar{\mathbf{x}}, \mathbf{y})$. Then:

$$\|\mathbf{v}\| \leq \sqrt{\sum_{i=1}^N (u_i(\bar{x}_i, \mathbf{y}))^2},$$

which shows that $C(\mathbf{y})$ is bounded.

To finalize assumption A we have to show that $C(\mathbf{y})$ is closed, let $(\mathbf{v}^t)_{t \in \mathbb{N}}$ be a sequence in $C(\mathbf{y})$ with $\mathbf{v}^t \rightarrow \mathbf{v}$. We need to show that $\mathbf{v} \in C(\mathbf{y})$. As $\mathbf{v}^t \in C(\mathbf{y})$ for all t , there is a sequence $(\mathbf{x}^t)_{t \in \mathbb{N}}$ such that for all $t \in \mathbb{N}$:

$$f(\mathbf{x}^t, \mathbf{y}) \leq 0 \text{ and } \mathbf{v}^t \leq \mathbf{u}(\mathbf{x}^t, \mathbf{y}).$$

The sequence $(\mathbf{x}^t)_{t \in \mathbb{N}}$ takes values in a bounded set (we previously showed that each $\mathbf{x}^t$ is bounded from above by $\bar{\mathbf{x}}$), so it has a convergent subsequence:

$$\mathbf{x}^{t_j} \rightarrow \mathbf{x}.$$

Now, along this subsequence:

$$\mathbf{v}^{t_j} \leq \mathbf{u}(\mathbf{x}^{t_j}, \mathbf{y}) \text{ and } \mathbf{v}^{t_j} \rightarrow \mathbf{v}.$$

Also, the functions $u_i(x_i, \mathbf{y})$ are continuous, so taking limits along the subsequence gives:

$$\mathbf{v} \leq \mathbf{u}(\mathbf{x}, \mathbf{y}).$$

Next, $f(\mathbf{x}^{t_j}, \mathbf{y}) \leq 0$ for all elements in the subsequence, so by continuity of f :

$$f(\mathbf{x}, \mathbf{y}) \leq 0.$$

This shows that $\mathbf{v} \in C(\mathbf{y})$. □

Let us now prove Theorem 5.8.

($\rightarrow$) The proof is by contradiction. Assume that the Coase Independence Property is satisfied but the NIP is violated. This means that there is a $\mathbf{v}$ such that:

$$\mathbf{v} \in \partial C(\mathbf{y}) \cap \partial C(\mathbf{y}') \text{ and } C(\mathbf{y}) \neq C(\mathbf{y}').$$

Without loss of generality, assume that there is a $\mathbf{w}$ such that:

$$\mathbf{w} \in C(\mathbf{y}) \setminus C(\mathbf{y}'). \tag{42}$$

As $\mathbf{v} \in \partial C(\mathbf{y}') \cap \partial C(\mathbf{y})$, by Lemma A.3, there should be an $\mathbf{x}^1$ and $\mathbf{x}^2$ such that:

$$\mathbf{v} = \mathbf{u}(\mathbf{x}^1, \mathbf{y}) = \mathbf{u}(\mathbf{x}^2, \mathbf{y}'), \tag{43}$$

$$(\mathbf{x}^1, \mathbf{y}) \in \mathcal{P}|_{\{\mathbf{y}\}} \text{ and } (\mathbf{x}^2, \mathbf{y}') \in \mathcal{P}|_{\{\mathbf{y}'\}}. \tag{44}$$

The proof takes several steps.

Step 1: Proof that $(\mathbf{x}^2, \mathbf{y}') \in \mathcal{P}|_{\{\mathbf{y}, \mathbf{y}'\}}$

If not, then there is an allocation $(\mathbf{x}^3, \mathbf{y}'')$ such that:

$$\begin{aligned} \mathbf{y}'' &\in \{\mathbf{y}, \mathbf{y}'\}, \\ f(\mathbf{x}^3, \mathbf{y}'') &\leq 0, \\ \mathbf{u}(\mathbf{x}^3, \mathbf{y}'') &\gg \mathbf{u}(\mathbf{x}^2, \mathbf{y}') = \mathbf{u}(\mathbf{x}^1, \mathbf{y}) = \mathbf{v}. \end{aligned}$$

If $\mathbf{y}'' = \mathbf{y}'$, this means that $\mathbf{v} \notin \partial C(\mathbf{y}')$ and if $\mathbf{y}'' = \mathbf{y}$ this means that $\mathbf{v} \notin \partial C(\mathbf{y})$. Both cases give a contradiction.

Step 2: Proof that for $\alpha^* = \max\{\alpha \geq 0 : \alpha \mathbf{w} \in C(\mathbf{y}')\}$, there is an allocation $(\mathbf{x}^4, \mathbf{y}')$ such that $\mathbf{u}(\mathbf{x}^4, \mathbf{y}') = \alpha^* \mathbf{w}$ and $(\mathbf{x}^4, \mathbf{y}') \in \mathcal{P}|_{\{\mathbf{y}, \mathbf{y}'\}}$

Observe that α^* is well defined by Lemma A.4. Let:

$$\mathbf{z} = \alpha^* \mathbf{w}.$$

By definition of α^* , we have that $\mathbf{z} \in \partial C(\mathbf{y}')$ and by (42):

$$\mathbf{z} < \mathbf{w}, \tag{45}$$

i.e. $\alpha^* < 1$. Using Lemma A.3 and $\mathbf{z} \in \partial C(\mathbf{y}')$, we know that there is an allocation $(\mathbf{x}^4, \mathbf{y}')$ such that:

$$(\mathbf{x}^4, \mathbf{y}') \in \mathcal{P}|_{\{\mathbf{y}'\}}, \text{ and } \mathbf{u}(\mathbf{x}^4, \mathbf{y}') = \mathbf{z}.$$

Also, by Step 1 above:

$$(\mathbf{x}^2, \mathbf{y}') \in \mathcal{P}|_{\{\mathbf{y}, \mathbf{y}'\}}$$

Applying the Coase Independence Property, it follows that $(\mathbf{x}^4, \mathbf{y}') \in \mathcal{P}|_{\{\mathbf{y}, \mathbf{y}'\}}$.

Step 3: Proof that there is an allocation $(\mathbf{x}^5, \mathbf{y})$ such that $\mathbf{u}(\mathbf{x}^5, \mathbf{y}) \gg \mathbf{u}(\mathbf{x}^4, \mathbf{y}')$ and $f(\mathbf{x}^5, \mathbf{y}) \leq 0$.

From (45), we know that:

$$\alpha^* \mathbf{w} = \mathbf{z} = \mathbf{u}(\mathbf{x}^4, \mathbf{y}') < \mathbf{w}$$

Also, $\mathbf{w} \in C(\mathbf{y})$ so there is some allocation $(\mathbf{x}^6, \mathbf{y})$ such that:

$$\mathbf{u}(\mathbf{x}^4, \mathbf{y}') \leq \mathbf{w} \leq \mathbf{u}(\mathbf{x}^6, \mathbf{y}) \text{ with } f(\mathbf{x}^6, \mathbf{y}) \leq 0.$$

Also, there is at least one agent j such that:

$$u_j(x_j^4, \mathbf{y}') < u_j(x_j^6, \mathbf{y}).$$

As $x_j^6 > 0$ (as $u_j(x_j^6, \mathbf{y}) > 0$, and by continuity of the utility functions and f , there are numbers $\varepsilon > 0$ and a $\delta > 0$ small enough such that:

$$\begin{aligned} u_j(x_j^6, \mathbf{y}') &< u_j(x_j^6 - \varepsilon, \mathbf{y}), \\ f\left(\mathbf{x}^6 - \varepsilon \mathbf{e}_j + \sum_{i \neq j} \delta \mathbf{e}_i, \mathbf{y}\right) &\leq 0. \end{aligned}$$

Let:

$$\mathbf{x}^5 = \mathbf{x}^6 - \varepsilon \mathbf{e}_j + \sum_{i \neq j} \delta \mathbf{e}_i.$$

Then:

$$\mathbf{u}(\mathbf{x}^5, \mathbf{y}) \gg \mathbf{u}(\mathbf{x}^4, \mathbf{y}') \text{ and } f(\mathbf{x}^5, \mathbf{y}) \leq 0,$$

as was to be shown.

From Step 3, it follows that $(\mathbf{x}^4, \mathbf{y}') \notin \mathcal{P}|_{\{\mathbf{y}, \mathbf{y}'\}}$ which contradicts the finding of Step 2.

($\leftarrow$) Assume that the NIP is satisfied. We need to show that the Coase independence property holds. Let Y be a subset of $\mathcal{Y}$. As $(C(\mathbf{y}))_{\mathbf{y} \in \mathcal{Y}}$ satisfies Assumptions A, we know by Lemma A.1 that the relation $\sqsubseteq$ is a complete ordering over $\mathcal{Y}$ with asymmetric part $\sqsubset$.

Let:

$$(\mathbf{x}^1, \mathbf{y}) \in \mathcal{P}|_Y, \quad (46)$$

and assume that:

$$(\mathbf{x}^2, \mathbf{y}) \in \mathcal{P}|_{\{\mathbf{y}\}}. \quad (47)$$

We need to show that:

$$(\mathbf{x}^2, \mathbf{y}) \in \mathcal{P}|_Y.$$

From Lemma A.3, (47) implies that:

$$\mathbf{u}(\mathbf{x}^2, \mathbf{y}) \in \partial C(\mathbf{y}). \quad (48)$$

We proof that $(\mathbf{x}^2, \mathbf{y}) \in \mathcal{P}|_Y$ by contradiction. Towards a contradiction, assume that:

$$(\mathbf{x}^2, \mathbf{y}) \notin \mathcal{P}|_Y.$$

Then there is a vector of public goods $\mathbf{y}' \in Y$ and an allocation $(\mathbf{x}^3, \mathbf{y}')$ such that:

$$\mathbf{u}(\mathbf{x}^3, \mathbf{y}') \gg \mathbf{u}(\mathbf{x}^2, \mathbf{y}) \text{ and } f(\mathbf{x}^3, \mathbf{y}') \leq 0.$$

But then $\mathbf{u}(\mathbf{x}^3, \mathbf{y}') \gg \mathbf{u}(\mathbf{x}^2, \mathbf{y})$ and, by Lemma A.3 $\mathbf{u}(\mathbf{x}^3, \mathbf{y}') \in C(\mathbf{y}')$. Also by (48): $\mathbf{u}(\mathbf{x}^2, \mathbf{y}) \in \partial C(\mathbf{y})$ which implies that:

$$\mathbf{u}(\mathbf{x}^3, \mathbf{y}') \in C(\mathbf{y}') \setminus C(\mathbf{y}).$$

From Lemma A.1 it follows that:

$$C(\mathbf{y}) \sqsubset C(\mathbf{y}').$$

As $\mathbf{u}(\mathbf{x}^1, \mathbf{y}) \in C(\mathbf{y})$, it follows that there is a profile $\mathbf{z}$ such that:

$$\mathbf{z} \gg \mathbf{u}(\mathbf{x}^1, \mathbf{y}) \text{ and } \mathbf{z} \in C(\mathbf{y}').$$

This contradicts (46) and $\mathbf{y}' \in Y$.

B Construction and properties of $\Gamma(\mathbf{u}, \phi)$

Assume that $(C(\phi))_{\phi \in \Phi}$ satisfies Assumption A. Define the correspondence $\Gamma : \mathbb{R}_{+,0}^N \times \Phi \rightrightarrows \mathbb{R}_{++}$, such that:

$$\Gamma(\mathbf{u}, \phi) = \{\gamma \geq 0 : \gamma \cdot \mathbf{u} \in \partial C(\phi)\}.$$

Fact 1: $\Gamma(\mathbf{u}, \phi)$ is non-empty, convex valued.

Let:

$$\gamma^* = \sup\{\gamma \geq 0 : \gamma \cdot \mathbf{u} \in \partial C(\phi)\}.$$

The aim is to show that $\gamma^* \in \Gamma(\mathbf{u}, \phi)$. Let us first show that it is well defined. By Assumption A, $\partial C(\phi)$ is non-empty, which means that $\mathbf{0} \in C(\phi)$. As such, $\gamma = 0$ is a feasible solution to this problem. Next, from Assumption A, we know that $C(\phi)$ is bounded so for all $\mathbf{u} > \mathbf{0}$, there is a $\gamma > 1$ such that $\gamma \mathbf{u} \notin C(\phi)$. This value is an upper-bound for $\{\gamma \geq 0 : \gamma \cdot \mathbf{u} \in C(\phi)\}$. As such, we know that γ^* is well defined.

Also, as γ^* is the supremum, we know that for all $t \in \mathbb{N}$ there is a γ^t such that:

$$\begin{aligned} \gamma^t &\in \{\gamma \geq 0 : \gamma \cdot \mathbf{u} \in C(\phi)\}, \\ \text{and } \gamma_t + \frac{1}{t} &\geq \gamma^* \geq \gamma^t. \end{aligned}$$

Observe that $(\gamma^t \cdot \mathbf{u}) \rightarrow (\gamma^* \cdot \mathbf{u})$ and that $(\gamma^t \cdot \mathbf{u}) \in C(\phi)$ for all t . By Assumption A, we have that $\gamma^* \cdot \mathbf{u} \in C(\phi)$. Define:

$$\mathbf{u}^* = \gamma^* \cdot \mathbf{u}.$$

If $\mathbf{u}^* \notin \partial C(\phi)$, then there is a $\mathbf{v} \gg \mathbf{u}^*$ such that $\mathbf{v} \in C(\phi)$. Let:

$$\gamma' = \min_i \{v_i / u_i^*\} > 1.$$

Then $\mathbf{u}^* < \gamma' \mathbf{u}^* \leq \mathbf{v}$ so by comprehensiveness $\gamma' \mathbf{u}^* \in C(\phi)$. But then:

$$\gamma' \cdot \gamma^* > \gamma^*,$$

a contradiction with the optimality of γ^* . This shows that $\gamma^* \mathbf{u} \in \partial C(\phi)$ so $\Gamma(\mathbf{u}, \phi)$ is non-empty.

Now, let us show that $\Gamma(\mathbf{u}, \phi)$ is convex-valuedness, let $\gamma, \gamma' \in \Gamma(\mathbf{u}, \phi)$ and assume without loss of generality that $\gamma' > \gamma$. Pick $\theta \in [0, 1]$, and set $\gamma_\theta = \theta\gamma + (1 - \theta)\gamma'$. Then:

$$\gamma \mathbf{u} \leq \gamma_\theta \mathbf{u} \leq \gamma' \mathbf{u}.$$

As $\gamma' \mathbf{u} \in C(\phi)$ it must be, by comprehensiveness that $\gamma_\theta \mathbf{u} \in C(\phi)$. The proof that $\gamma_\theta \mathbf{u} \in \partial C(\phi)$ is by contradiction. If:

$$\gamma_\theta \mathbf{u} \notin \partial C(\phi),$$

it follows from $\gamma_\theta \mathbf{u} \in C(\phi)$ that there is a $\mathbf{v} \gg \gamma_\theta \mathbf{u}$ such that $\mathbf{v} \in C(\phi)$. But then, also $\mathbf{v} \gg \gamma \mathbf{u}$ which contradicts the assumption that $\gamma \mathbf{u} \in \partial C(\phi)$.

Fact 2: $\Gamma(\mathbf{u}, \phi)$ is upper hemicontinuous in $\mathbf{u}$ on $\mathcal{D} \setminus \{\mathbf{0}\}$.

Non-emptiness and compactness of $\Gamma(\mathbf{u}, \phi)$ was shown above. Now, take any sequence $\mathbf{u}^t \rightarrow \mathbf{u}$ and any sequence $\gamma^t \in \Gamma(\mathbf{u}^t, \phi)$ we need to show that γ^t has a convergent subsequence whose limit is in $\Gamma(\mathbf{u}, \phi)$.

Step 1: The sequence γ^t is bounded

As $\mathbf{u}^t \rightarrow \mathbf{u}$ we know that for any $\varepsilon > 0$, there is a T such that for all $t \geq T$:

$$(1 - \varepsilon)\mathbf{u} < \mathbf{u}^t.$$

Let $\gamma^* = \sup\{\gamma \in \Gamma(\mathbf{u}, \phi)\}$. Next, choose $\gamma > \gamma^*$ such that:

$$\gamma(1 - \varepsilon) > \gamma^*.$$

Then, $\gamma(1 - \varepsilon)\mathbf{u} > \gamma^*\mathbf{u}$, so $\gamma(1 - \varepsilon)\mathbf{u} \notin C(\phi)$.

Now, $\gamma(1 - \varepsilon)\mathbf{u} < \gamma\mathbf{u}^t$, which means that $\gamma\mathbf{u}^t \notin C(\phi)$. This means that $\gamma > \gamma^t$ for all $\gamma^t \in \Gamma(\mathbf{u}^t, \phi)$ and all $t \geq T$. Taking the maximum of γ and $\cup_{t \leq T} \Gamma(\mathbf{u}^t, \phi)$ gives the desired upper bound.

Step 2: γ^t has a convergent subsequence that converges to $\gamma \in \Gamma(\mathbf{u}, \phi)$.

Given that $(\gamma^t)_{t \in \mathbb{N}}$ is bounded, it has a convergent subsequence $\gamma^{t_i} \rightarrow \gamma$. Consequently:

$$\gamma^{t_i}\mathbf{u}^{t_i} \rightarrow \gamma\mathbf{u}.$$

If we can show that $\gamma\mathbf{u} \in \partial C(\phi)$, then we are done as this means that $\gamma \in \Gamma(\mathbf{u}, \phi)$.

Now, for all t_i , $\gamma^{t_i}\mathbf{u}^{t_i} \in \partial C(\phi) \subseteq C(\phi)$. Given that $C(\phi)$ is closed, we have that:

$$\gamma\mathbf{u} \in C(\phi).$$

If, towards a contradiction $\gamma\mathbf{u} \in C(\phi) \setminus \partial C(\phi)$ then there is a $\mathbf{v} \gg \gamma\mathbf{u}$ such that $\mathbf{v} \in C(\phi)$. Then, for t_i large enough, $\gamma^{t_i}\mathbf{u}^{t_i} \ll \mathbf{v}$ which contradicts the assumption that $\gamma^{t_i}\mathbf{u}^{t_i} \in \partial C(\phi)$ for all t_i .