Cowles Foundation Discussion Paper No. 1854

KANTIAN OPTIMIZATION, SOCIAL ETHOS, AND PARETO EFFICIENCY

John E. Roemer

March 2012

March 1, 2012

“Kantian optimization, social ethos, and Pareto efficiency”

by

John E. Roemer[1]

Abstract. Although evidence accrues in biology, anthropology and experimental economics that *homo sapiens* is a cooperative species, the reigning assumption in economic theory is that individuals optimize in an autarkic manner (as in Nash and Walrasian equilibrium). I here postulate an interdependent kind of optimizing behavior, called Kantian. It is shown that in simple economic models, when there are negative externalities (such as congestion effects from use of a commonly owned resource) or positive externalities (such as a social ethos reflected in individuals’ preferences), Kantian equilibria dominate Nash-Walras equilibria in terms of efficiency. While economists schooled in Nash equilibrium may view the Kantian behavior as utopian, there is some – perhaps much -- evidence that it exists. If cultures evolve through group selection, the hypothesis that Kantian behavior is more prevalent than we may think is supported by the efficiency results here demonstrated.

Key words: Kantian equilibrium, social ethos, implementation
JEL codes: D60, D62, D64, C70, H30

---

[1] Department of political science and Cowles Foundation, Yale University. My main debt is to Joaquim Silvestre, with whom I have discussed these ideas for many years. I also thank the following, for their suggestions and help: Roger Howe, François Maniquet, Juan Moreno-Ternero, Joseph Ostroy, Larry Samuelson, and Joel Sobel.

1. Introduction

Three strands of work in contemporary social science, evolutionary biology, and political philosophy unite in emphasizing this fact: that *homo sapiens* is a cooperative species. In evolutionary biology, this statement is accepted as a premise, and scientists are interested in explaining how cooperation and 'altruism' may have developed among humans. In economics, there is now a long series of experiments whose results are most easily explained by the hypothesis that individuals are to some degree altruistic. Altruism is to be distinguished from reciprocity: when behaving in a cooperative manner, a reciprocator expects cooperation in return, which will increase his/her net payoff (net, that is, of the original sacrifice entailed in cooperation), while an altruist cooperates without the expectation of a future reciprocating behavior. Many biologists, experimental economists, and anthropologists now accept the existence of altruistic as well as reciprocating behavior. A recent summary of the state-of-the-art in experimental economics, anthropology, and evolutionary biology is provided by Bowles and Gintis (2011). See Rabin (2006) for a summary of the evidence from experimental economics. An anthropological view is provided in Henrich and Henrich (2007). A recent paper which provides a good bibliography of work attempting to explain altruistic preferences as evolutionary equilibria is Alger and Weibull (2012).

In political philosophy, G.A. Cohen (2010) offers a definition of 'socialism' as a society in which earnings of individuals at first accord with a conception of equality of opportunity that has developed in the last thirty years in political philosophy (see Rawls (1971), Dworkin (1981), Arneson (1989), and Cohen(1989)), but in which inequality in those earnings is then reduced because of the necessity to maintain 'community,' an ethos in which '…people care about, and where necessary, care for one another, and, too, care that they care about one another.' Community, Cohen argues, may induce a society to reduce material inequalities (for example, through taxation) that would otherwise be acceptable according to 'socialist' equality of opportunity. But, Cohen writes:

> …the principal problem that faces the socialist ideal is that we do not know how to design the machinery that would make it run. Our problem is not, primarily, human selfishness, but our lack of a suitable organizational technology: our problem is a problem of design. It may be an insoluble design

> problem, and it is a design problem that is undoubtedly exacerbated by our selfish propensities, but a design problem, so I think, is what we've got.

An economist reading these words thinks of the first theorem of welfare economics. A Walrasian equilibrium is Pareto efficient in an economy with complete markets, private goods, and the absence of externalities. But under the communitarian ethos, people care about the welfare of others – which induces massive consumption externalities – and so the competitive equilibrium will not, in general, be efficient. What economic mechanism can deliver efficiency under these conditions[2]?

There is an important line of research, conducted by Ostrom (1990) and her collaborators, arguing that, in many small societies, people figure out how to avoid, or solve, the 'tragedy of the commons.' The 'tragedy' has in common with altruism the existence of an externality which conventional optimizing behavior does not properly address[3]. It may be summarized as follows. Imagine a lake which is owned in common by a group of fishers, who each possess preferences over fish and leisure, and perhaps differential skill (or sizes of boats) in (or for) fishing. The lake produces fish with decreasing returns with respect to the fishing labor expended upon it. In the game in which each fisher proposes as her strategy a fishing time, the Nash equilibrium is inefficient: there are congestion externalities, and all would be better off were they able to design a decrease, of a certain kind, in everyone's fishing. Ostrom has studied many such societies, and maintains that many or most of them learn to regulate 'fishing,' without privatizing the 'lake.' Somehow, the inefficient Nash equilibrium is avoided. This example is not one in which fishers care about other fishers (necessarily), but it is one in which cooperation is organized to deal with a negative externality of autarkic behavior.

Ostrom's observations pertain to small societies. In large economies, we observe the evolution of the welfare state, supported by considerable degrees of taxation

---

[2] In war-time Britain, many spoke of 'doing their bit' for the war effort – voluntary additional sacrifice for the sake of the common good. (See the wonderful BBC series 'Foyle's War' to understand the pervasiveness of this ethos.) But, if I want to contribute to the common struggle, how *much* extra should I do?

[3] In the case of altruism, 'conventional' behavior is market behavior, and in the case of the tragedy of the commons, it is autarkic optimizing behavior in using a resource which is owned in common.

of market earnings. It is not immediately evident that welfare states are due to a feeling of community (à la Cohen), or simply provide a more conventional public good or a good in which market failures abound (insurance), or reflect reciprocating behavior among citizens (welfare states expand after wars, perhaps as a reward to returning soldiers; see Scheve and Stasavage[in press]). Nevertheless, the large scope of welfare states, especially in Northern Europe, is perhaps most easily explained by a communitarian ethos. Redistributive taxation is, that is to say, at least some degree a reaction to the material deprivation of a section of society, which others view as undeserved, and desire to redress. Nevertheless, as is well-known, redistributive taxation induces, to some degree, allocative inefficiency. The solution is second-best.

Economic theorists are beginning to pay attention to the design problem – that is, how to achieve economic efficiency in a society where people care about other people. Perhaps to say they are ‘beginning’ to do so is uncharitable: implementation theory, largely initiated with Maskin’s (1999) work of thirty years ago, asks whether a social-choice rule can be implemented as the Nash equilibrium of a game. And before Maskin, Leonid Hurwicz pioneered the work on mechanism design, in which he studied the efficiency properties of different economic mechanisms at a highly abstract level. This work, however, did not focus upon the issue of externalities induced by the fact that people care about the welfare of other people.

A recent contribution which is relevant to this inquiry is that of Dufwenberg, Heidhues, Kirchsteiger, Riedel, and Sobel (2010), entitled “Other-regarding preferences in general equilibrium,” which studies, at an abstract level, the veracity of the first and second welfare theorems in the presence of other-regarding preferences. From the viewpoint of the evolution of economic thought, it is significant that their article is the result of combining three independent papers by subsets of the five authors: in other words, the problem of addressing seriously the efficiency consequences of the existence of other-regarding preferences is certainly in the air at present.

In this paper, I wish to offer a partial solution to two problems of economic allocation: how to achieve efficiency in environments where there are positive and negative externalities and individuals are conventionally self-interested, and secondly, how to achieve efficiency in the presence of a *social ethos* – I use the term, taken from

Bowles and Gintis (2011) -- although ‘other-regarding preferences’ is a synonym. (Perhaps social ethos includes the kind of second-order preference that G.A. Cohen refers to in defining community, that people care that they care about others, while ‘other-regarding preferences’ does not.) The ‘problem’ is that market equilibria are in general Pareto inefficient in the presence of a social ethos, and moreover, redistributive taxation is also inefficient.

I next describe the economic environment for this inquiry. There is a concave production function which produces a single output from a single input, called effort. Effort is supplied by individuals; it may differ in intensity or efficiency units, but effort can be aggregated across individuals when measured in the proper units. Individuals have conventional personal utility functions, representing their self-interested preferences over income and effort. In general, they care about the welfare of others as well. There are two aspects to this caring: how individuals choose to *aggregate* individual welfares into social welfare, and the *degree* to which social welfare counts in the individual’s preferences. We will assume here that individuals are homogeneous with respect to these two decisions.

An individual of type γ has preferences represented by an *all-encompassing utility function* which might be of the form:

$$U^{\gamma}(x(\cdot),E(\cdot)) = u^{\gamma}(x(\gamma),E(\gamma)) + \alpha \exp\int \log[u^{\tau}(x(\tau),E(\tau))]dF(\tau) \quad (1.1)$$

where $u^{\gamma}(\cdot,\cdot)$ is the *personal utility function* of type γ over consumption and effort, $E(\cdot):\mathbb{R}_+ \to \mathbb{R}_+$ is a function which describes the efforts of individuals of all types, $x(\cdot):\mathbb{R}_+ \to \mathbb{R}_+$ is a function which defines the amount of output (a single good) allocated to each type, α is a non-negative number measuring the degree of social ethos, *F* is the distribution of types in the society, and the social-welfare function (in this case) is given by a member of the CES family

$$W^{p}(u[i]) = \left(\int u[i]^{p}\, dF(i)\right)^{1/p}, \quad (1.2)$$

as $p \to 0$. (It is well-known that the function in (1.2) approaches the exponential of the average of the logarithms as $p \to 0$.) Think of an individual's type as signifying, inter alia, the degree to which effort is easy for him, or his natural talent.

A society in which people do not count the welfare of others is one with *individualistic ethos*: in such a society, $\alpha = 0$, A society in which they do is one with *social ethos*. Social ethos can be stronger or weaker, as represented by the parameter α. When $\alpha = \infty$, the economy is equivalent to the one in which for everyone, all-encompassing utility is equal to social welfare; this is the purely altruistic economy.

Production is described by a differentiable, concave production function *G*. In the continuum economy, the value $G(\bar{E})$ is per capita output of the good when the effort schedule is $E(\cdot)$ and $\bar{E} \equiv \int E(\gamma)\, dF(\gamma)$. When the number of agents is finite, I usually write the discrete effort vector as $E = (E^1, ..., E^n)$ and the sum of efforts as $E^S \equiv \sum E^j$. Total output is then $G(E^S)$.

Suppose that G is linear and there is a private-ownership economy with zero profits at competitive equilibrium. A typical allocation rule is the linear-tax rule:

$$x_t^{\gamma}(E(\cdot)) = (1-t)wE(\gamma) + t\int wE(\tau)\, dF(\tau) \;, \qquad (1.3)$$

where *w* is the wage paid by the firm and *t* is the tax rate. Under the competitive assumption, the firm pays a wage equal to the marginal product of effort, $w = G'(\bar{E})$. There are two important kinds of externality here – both positive: the tax system creates positive externalities to individual labor, because in general some of each worker's earnings is redistributed to others, and there are also positive consumption externalities due to social ethos. It is unfortunate that, under classical behavior, at least if the economy is large, individuals ignore the positive externalities induced by their labor. I call this classical behavior *autarkic*, and contrast it with behavior that I call *interdependent.* The equilibrium concept associated with autarkic behavior is *Nash* equilibrium; the concept associated with interdependent behavior is *Kantian* equilibrium. In Nash equilibrium, each person adjusts his action if and only if his situation would improve assuming that others do not adjust theirs. In Kantian equilibrium, a person adjusts his action if and only if his situation would improve if all others adjust their actions *in similar fashion* to the personal adjustment he is contemplating. Definitions will be provided in the next section. There is only one concept of Nash equilibrium, but there are many concepts of Kantian equilibrium, because the phrase 'in similar fashion' can be spelled out in various ways.

My main focus will be upon behavior: that is, upon how a change in optimizing behavior from autarkic to interdependent (Nash to Kantian) can (or cannot) resolve the inefficiencies due to positive and negative externalities, and in particular, those induced by the existence of a social ethos. Just as economists are often asked to accept the idea that the *formal* concept Nash equilibrium captures a common kind of *actual* stable point in human economic relations, so I will ask readers to accept, for the sake of argument, that Kantian equilibrium (in its various versions) can capture a kind of social equilibrium. In section 5, I will contemplate whether Kantian behavior is achievable in human societies, or is simply a utopian idea.

I contrast the approach here with that of almost all the literature on altruism, cited earlier. In that literature, the focus is upon explaining the emergence and stability of altruistic preferences. My focus here is upon a different protocols of optimization, and hence upon different conceptions of equilibrium in games[4].

Section 2 defines Kantian equilibrium, and studies its efficiency properties in conventional economies where there is individualistic ethos. Section 3 looks at economies with a social ethos, and studies Kantian equilibrium there. Section 4 provides an existence theorem for Kantian equilibrium, and comments upon dynamic properties. Section 5 discusses the question whether Kantian optimization is a utopian idea, of only theoretical interest, or whether it might come to be characterize human societies.

I originally proposed the definition of (multiplicative) Kantian equilibrium in Roemer (1996), and showed its relationship to the 'proportional solution.' In Roemer (2010), I investigated multiplicative Kantian equilibrium more carefully. The present paper shows that there are many versions of Kantian optimization, and characterizes

[4] One can define a two-by-two taxonomy of models. Preferences can be either altruistic or individualistic, and optimization can either be Nash or Kantian. The classical world is modeled as (individualistic preferences, Nash optimization). The work I have cited focuses upon models of the form (altruistic, Nash). In this paper, I study (individualistic, Kantian) and (altruistic, Kantian).

when they deliver efficient outcomes in the presence of the various kinds of externality that I have reviewed in this section.

2. Kantian equilibrium in economies with an individualistic ethos

Immanuel Kant proposed the behavioral ethic known as the categorical imperative: take those actions and only those actions which you would have all others emulate[5]. This suggests the following formalization. Let $\{V^\gamma(E(\cdot))\}$ be a set of payoff functions for a game played by types γ, where the strategy of each player is a non-negative effort $E(\gamma)$. Thus the payoff of each depends upon the efforts of all. A *multiplicative Kantian equilibrium* is an effort schedule $E^*(\cdot)$ such that *nobody would prefer that everybody alter his effort by the same factor.* That is:

$$(\forall\gamma)(\forall r \geq 0)(V^\gamma(E^*(\cdot)) \geq V^\gamma(rE^*(\cdot))) . \tag{2.1}$$

In Roemer (1996, 2010), this concept was simply called ‘Kantian equilibrium.’

The remarkable feature of multiplicative Kantian equilibrium is that it resolves the tragedy of the commons. Consider the example given in section 1 of the community of fishers. At an effort allocation $E(\cdot)$, if each fisher of type γ keeps his catch, then his fish income will be :

$$x^f(E(\cdot),\gamma) = \frac{E(\gamma)}{\bar{E}} G(\bar{E}) . \tag{2.2}$$

.

Thus, the fishers’ game is defined by the payoff functions:

$$V^\gamma(E(\cdot)) = u^\gamma(x^f(E(\cdot),\gamma), E(\gamma)) . \tag{2.3}$$

It is proved in the two citations given above to Kantian equilibrium that *if a strictly positive effort allocation is a multiplicative Kantian equilibrium, then it is Pareto efficient in the economy* $\xi = (\mathbf{u}, G, F, 0)$, where $\mathbf{u}$ is a profile of concave utility functions, and the last co-ordinate in the description of the economy is the value of α. This is a

---

[5] The somewhat more general version of the categorical imperative is that one’s behavior should accord with 'universalizable maxims.’

stronger statement than saying the allocation is efficient in the game $\{V^{\gamma}\}$: for in the game, only certain types of allocation are permitted – ones in which fish are distributed in proportion to effort expended. But the economy $(\mathbf{u}, G, F, 0)$ defines any allocation as feasible, as long as $\int x(\gamma)\, dF(\gamma) \leq G(\bar{E})$. So Kantian behavior, if adopted by individuals, resolves the tragedy of the commons. The intuition is that the Kantian counterfactual (that *every* person will expand his labor by a factor r if I do so – or so I contemplate) forces each to internalize the externality associated with the congestion effect of his own fishing. It is not obvious that multiplicative Kantian equilibrium will internalize the externality in exactly the right way – to produce efficiency – but it does.

A *proportional solution* in the fisher economy is defined as an allocation $(x(\cdot), E(\cdot))$ with two properties:

(i) $x(\gamma) = x^f(E(\cdot), \gamma)$, and

(ii) $(x(\cdot), E(\cdot))$ is Pareto efficient.

The proportional solution was introduced in Roemer and Silvestre (1993), although the concept of (multiplicative) Kantian equilibrium came later. The proportional solutions of the fisher economy are exactly its positive multiplicative Kantian equilibria (see theorem 1 below). In the small societies which Ostrom has studied, which are (in the formal sense) usually ‘economies of fishers’ where each ‘keeps his catch,’ she argues that internal regulation assigns ‘fishing times’ that engender a Pareto efficient allocation. If this is so, these allocations are proportional solutions, and therefore (by the theorem just quoted) they are multiplicative Kantian equilibria in the game where participating fishers/hunters/miners propose labor times for accessing a commonly owned resource. This suggests that small societies discover their multiplicative Kantian equilibria. Ostrom (1990), however, does not provide any evidence for Kantian thinking among citizens of these socieities. Knowing the theory of multiplicative Kantian equilibrium, one is tempted to ask whether a ‘Kantian ethos’ exists in these small societies, which somehow leads to the discovery of the equilibrium.

I now introduce a second Kantian protocol which leads to *additive Kantian equilibrium*[6]. An effort allocation $E(\cdot)$ is an additive Kantian equilibrium if and only if no individual would have all individuals add (or subtract) the same amount of effort to everyone's present effort. That is:

$$(\forall\gamma)(\forall r \geq -\inf_{\tau} E(\tau))(V^{\gamma}(E(\cdot)) \geq V^{\gamma}(E(\cdot)+r)), \tag{2.4}$$

where $E(\cdot)+r$ is the allocation in which the effort of type γ individuals is $E(\gamma)+r$. The lower bound $(r \geq -\inf_{\tau} E(\tau))$ is necessary to avoid negative efforts, and to keep the optimization problem proposed in (2.4) a concave problem. (It is assumed that effort is unbounded above but bounded below by zero.) Additive Kantian equilibrium again postulates that each person 'internalizes' the effects of his contemplated change in effort, but now the variation is additive rather than multiplicative.

In the sequel, I will denote these two kinds of Kantian behavior as $K^{\times}$ and K^{+}.

We can moreover define a general 'Kantian variation' which includes as special cases additive and multiplicative Kantian equilibrium. We say a function $\varphi:\mathbb{R}_+^2 \to \mathbb{R}_+$ is a *Kantian variation* if :

$$\forall x \quad \varphi(x,1) = x,$$

and if, for any $x \neq 0$, the function $\varphi(x,\cdot)$ maps onto the non-negative real line.

Denote by $\varphi[E(\cdot),r]$ the effort schedule $\tilde{E}$ defined by $\tilde{E}(\gamma) = \varphi(E(\gamma),r)$.

Then an effort schedule $E(\cdot)$ is a $\varphi-$Kantian equilibrium if and only if:

$$(\forall\gamma)(V^{\gamma}(\varphi[E(\cdot),r]) \text{ is maximized at } r=1) \quad . \tag{2.5}$$

If we let $\varphi(x,r) = rx$, this definition reduces to multiplicative Kantian equilibrium; if we let $\varphi(x,r) = x+r-1$, it reduces to additive Kantian equilibrium.

Let $\varphi(x,r)$ be any Kantian variation that is concave in *r*, and let the payoff functions with respect to some sharing rule $\{V^{\gamma}\}$ be concave. Then a positive effort schedule $E(\cdot)$ is a positive $\varphi-$Kantian equilibrium if and only if:

[6] This variation of Kantian equilibrium was proposed to me by J. Silvestre in 2004.

$$\forall\gamma \quad \left.\frac{d}{dr}\right|_{r=1} V^{\gamma}(\varphi[E(\cdot),r]) = 0\,. \tag{2.6}$$

Eqn. (2.6) follows immediately from definition (2.5), since $V^{\gamma}(\varphi[E(\cdot),r])$ is a concave function of *r*, and hence its maximum, if it is interior, is achieved where its derivative with respect to *r* is zero. Note that both the additive and multiplicative Kantian variations are concave functions of *r*.

I find it convenient to describe allocation rules by *sharing rules.* Denote by **G** the set of all concave differentiable production functions, and by **E** the set of all effort vectors, that is, functions $E:\mathbb{R}_+ \to \mathbb{R}_+$. A *sharing rule* is a set of functions $\{\theta^{\gamma}\}$, one for each type, where $\theta^{\gamma}:\mathbf{E}\times\mathbf{G}\to[0,1]$ and for all (E,G):

$$\int \theta^{\gamma}(E,G)\,dF(\gamma) = 1. \tag{2.7}$$

The amount of output which type γ receives at $E(\cdot)$ when the production function is *G* is $\theta^{\gamma}(E,G)G(\bar{E})$, where $\bar{E}$ is interpreted as average effort in continuum economies, and as the sum of efforts in finite economies. Note that, although sharing rules (and hence allocation rules) can depend on *G*, they do not depend on the utility functions of agents.

Examples.

1. The *proportional sharing rule* is given by $\theta^{\gamma,P}(E(\cdot),G) = \dfrac{E(\gamma)}{\bar{E}}$
2. The *equal division sharing rule* is given by

$$\theta^{\gamma,ED}(E(\cdot),G) = \begin{cases} 1, \text{ in continuum economies} \\ \dfrac{1}{n}, \text{ in economies with } n \text{ agents} \end{cases}$$

3. The *Walrasian sharing rules* are given by:

$$\theta^{\gamma,W}(E(\cdot),G) = \frac{G'(\bar{E})E(\gamma)}{G(\bar{E})} + \sigma(\gamma)\left(1 - \frac{\bar{E}G'(\bar{E})}{G(\bar{E})}\right),$$

where $\sigma(\gamma)$ is the share of the firm that operates *G* owned by each agent of type γ. Note that although the proportional and equal-division sharing rules do not, in fact, depend upon *G*, the Walrasian sharing does (except when *G* is linear). This is one reason it is important to allow sharing rules to depend on *G*.

Once we propose a sharing rule, then we can define, for any economy $(\mathbf{u}, G, F, 0)$, its payoff functions $\{V^\gamma\}$, and hence its $K^\times$ and K^+ equilibria. Define the domain of concave economies $\mathfrak{G}$ as all economies $(\mathbf{u}, G, F, \alpha)$ where $\mathbf{u}$ is a profile of concave personal utility functions $u: \mathbb{R}^2_+ \to \mathbb{R}$, $G \in \mathbf{G}$, *F* is a distribution function of types, and $\alpha \geq 0$ is any degree of social ethos. (We fix the social-welfare function – for instance, the one displayed in (1.1).) Denote by $\mathfrak{G}^0$ the class of economies with $\alpha = 0$, by $\mathfrak{G}^{fin}$ the class of economies with a finite number of agents, and by $\mathfrak{L}$ the class of economies where *G* is linear, and so. (E.g., $\mathfrak{L}^{0,fin}$ is the class of finite economies with $\alpha = 0$.) Although proofs of theorems will generally appear in the appendix, it is important to demonstrate the most important idea in this paper by proving the first proposition in the text.

Proposition 1 *Any strictly positive* $K^\times$ *equilibrium with respect to the proportional sharing rule is Pareto efficient on the domain* $\mathfrak{G}^0$. *Any strictly positive* K^+ *equilibrium with respect to the equal-division sharing rule is Pareto efficient on the domain* $\mathfrak{G}^0$.

Proof:

1. Let $E(\cdot)$ be a strictly positive $K^\times$ equilibrium w.r.t. the proportional sharing rule θ^P. The first-order condition stating this fact is:

$$(\forall\gamma) \quad \left.\frac{d}{dr}\right|_{r=1} u^\gamma\left(\frac{rE(\gamma)}{r\bar{E}} G(r\bar{E}), rE(\gamma)\right) = 0, \tag{2.8}$$

which means:

$$(\forall\gamma) \quad u_1^\gamma \cdot \left(\frac{E(\gamma)}{\bar{E}} G'(\bar{E})\bar{E}\right) + u_2^\gamma E(\gamma) = 0. \tag{2.9}$$

Since $E(\gamma) > 0$, divide through (2.9) by $E(\gamma)$, giving:

$$(\forall\gamma) \quad -\frac{u_2^\gamma}{u_1^\gamma} = G'(\bar{E}). \tag{2.10}$$

Eqn. (2.10) states that the marginal rate of substitution between income and effort is, for every agent, equal to the marginal rate of transformation, which is exactly the condition for Pareto efficiency at an interior solution. This proves the first claim.

2. For the second claim, let $E(\cdot)$ be a K^{+} equilibrium w.r.t. the equal-division sharing rule θ^{ED} for any economy in $\mathfrak{G}^{0}$. Then:

$$(\forall\gamma)\quad \left.\frac{d}{dr}\right|_{r=0} u(G(\bar{E}+r),E(\gamma)+r)=0\,, \tag{2.11}$$

which means:

$$(\forall\gamma)\quad u_1^{\gamma}\cdot G'(\bar{E})+u_2^{\gamma}=0\,. \tag{2.12}$$

(Strict positivity of E is here used so that the range of r includes a small neighborhood of zero.) Clearly (2.12) implies (2.10), and again the allocation is Pareto efficient. ■

Examine the proof of the first part of this proposition, and compare the reasoning that agents who are Kantian employ to Nash reasoning. When a fisher contemplates increasing his effort on the lake by 10%, she asks herself, “How would I like it if everyone increased his effort by 10%?” She is thereby forced to internalize the externality that her increased labor would impose on others, when G is strictly concave.

It is important to note that, in Kantian optimization, agents evaluate deviations from their own viewpoints, as in Nash optimization. They do not put themselves in the shoes of others, as they do in Rawls’s original position, or in Harsanyi’s (1977) thought experiment in which agents employ *empathy*. In this sense, Kantian behavior requires *less of a displacement of the self* than ‘veil-of-ignorance’ thought experiments require. Agents require no empathy to conduct Kantian optimization: what changes from Nash behavior is the supposition about the counterfactual.

Indeed, the next theorem states that there is a unidimensional continuum of sharing rules, with the proportional and equal-division rules as its two endpoints, each of which can be efficiently implemented on $\mathfrak{G}^{0}$ using a particular Kantian variation. Define the allocation rule:

$$\theta_{\beta}^{\gamma}(E(\cdot))=\frac{E(\gamma)+\beta}{\bar{E}+\beta},\quad 0\le\beta\le\infty \tag{2.13}$$

and the Kantian variations:

$$\varphi_\beta(x,r) = rx + (r-1)\beta, \quad 0 \le \beta \le \infty. \tag{2.14}$$

(For finite economies, we write (2.13) as $\theta_\beta^\gamma(E(\cdot)) = \dfrac{E(\gamma)+\beta}{E^S + n\beta}, \quad E^S = \sum E(\tau)$.) Note that for $\beta = 0$, θ_β is the proportional rule and φ_β is the multiplicative Kantian variation, and for $\beta = \infty$, θ_β is the equal-division rule and φ_β is the additive Kantian variation (this last fact is perhaps not quite obvious). We will call a Kantian equilibrium associated with the variation φ_β, a K^β equilibrium. (So $K^0 \equiv K^\times$, etc.)

Before stating the next theorem we must define the following. Fix β and an effort vector $E \in \mathbb{R}^n_{++}$. Define $r_i^j = \dfrac{E^i + \beta}{E^j + \beta}$. Now consider the set of vectors in $\mathbb{R}^n_+$ of the form $(\varphi_\beta(x, r_1^j), \varphi_\beta(x, r_2^j), ..., \varphi_\beta(x, r_j^n))$ where x varies over the positive real numbers, but restricted to an interval that keeps the defined vector non-negative. This is a unidimensional manifold in $\mathbb{R}^n_+$ which I denote as $\mathfrak{M}_E^j$. We have:

<u>Theorem 1</u>[7] *For* $0 \le \beta \le \infty$:

A. *If* $E(\cdot)$ *is a strictly positive* K^β *equilibrium w.r.t. the sharing rule* θ_β *at any economy in* $\mathfrak{G}^0$*, then the induced allocation is Pareto efficient.*

B. θ_0 *is the only sharing rule for which the* $K^\times$ *equilibrium is Pareto efficient on the domain* $\mathfrak{G}^{0,fin}$.

C. *For* $\beta > 0$, *the only sharing rules that are efficiently implementable on* $\mathfrak{G}^{0,fin}$ *are of the form* $\theta^j(E,G) = \theta_\beta^j(E) + \dfrac{k^j(E)}{G(E)}$ *where* $\{k^j\}$ *are any functions satisfying:*

(i) $\sum_j k^j(E) \equiv 0$

[7] Theorem 3 of Roemer (2010) stated something similar to part B of the present theorem, but the proof offered there is incorrect.

(ii) $(\forall j,E)(\theta^j(E,G)\in[0,1])$, and

(iii) $(\forall j,E)(k^j$ is constant on the manifold $\mathfrak{M}_E^j)$. *That is, on* $\mathfrak{M}_E^j$,

$$\nabla k^j \cdot r^j \equiv 0 .$$

Proof: See appendix[8].

The theorem states first that for all $\beta \geq 0$, the pair $(\varphi_\beta,\theta_\beta)$ is an *efficient Kantian pair*: i.e., that the sharing rule θ_β is efficiently implementable in K^β equilibrium on the convex domain $\mathfrak{E}^{0,fin}$. Part C states that the only other sharing rules that are K^β implementable are ones which add numbers to the θ_β shares that are constant on certain sets of lines in $\mathbb{R}_+^n$. Part B states that (in the unique case when $\beta = 0$) these constants must be zero.

Unfortunately, part C makes theorem 1 difficult to state. One may ask, is it necessary? That is, do there in fact exist sharing rules satisfying conditions C (i)-C(iii) of the theorem where the functions k^j are not identically zero? The following example shows that there are.

Example 4.

We consider K^+ equilibrium (i.e., $\beta = \infty$) where $n = 2$. In this case

$$\theta_\infty^j(E^1,E^2) = \frac{1}{2} ,$$

that is, the equal-division sharing rule. Now consider:

$$\tilde{\theta}^1(E) = \begin{cases} \dfrac{1}{2} + \dfrac{G(E^1-E^2)}{2G(E^1+E^2)}, \text{ if } E^1 \geq E^2 \\ \dfrac{1}{2} - \dfrac{G(E^1-E^2)}{2G(E^1+E^2)}, \text{ if } E^1 < E^2 \end{cases} \quad . \qquad (2.15)$$

$$\tilde{\theta}^2(E) = 1 - \tilde{\theta}^1(E)$$

[8] I believe that parts B and C are also true on the space of continuum economies, but proving that would require more sophisticated mathematical techniques.

The $\tilde{\theta}$ rule satisfies conditions C(i)-C(iii). To explain in words why rules like this work, think of what happens when we apply the appropriate Kantian variation to a vector E under this rule. In this case, we add a constant r to all effort levels. Notice that the second term in $\tilde{\theta}^j(E)$ is unaffected, because $E^1 + r - (E^2 + r) = E^1 - E^2$. Therefore, when looking at Kantian deviations, and setting the derivative of utility equal to zero at $r = 0$ (in the additive case), these k^j terms vanish. Sharing rules like $\tilde{\theta}$ can be constructed for any n and any $\beta > 0$.

From the history-of-thought vantage point, the case $\beta = 0$ is the classical socialist economy: that is, it's an economy where output is distributed in proportion to labor expended *and efficiently so*. The rule θ_β in case $\beta = \infty$ is the classical 'communist' economy: output is distributed 'according to need' (here, needs are identical across persons), *and efficiently so*. Obviously, the sharing rules θ_β associated with $\beta \in (0,\infty)$ are (in a sense) averages between these two classical concepts. If the 'interactive' optimization reflected in the Kantian way of thinking is akin to a kind of cooperation, it is perhaps not surprising to note that these classical concepts of cooperative societies are efficiently implemented by different versions of Kantianism.

I believe that history displays examples of both the proportional and equal-division solutions. The former has been discussed in relation to Ostrom's work on fisher economies. And anthropologists conjecture that many hunter-gatherer societies employed equal-division. Many Israeli kibbutzim employed the equal division rule, at least in the early days. (Whether they found Pareto efficient equal-division allocations is another question.) Theorem 1 suggests that we look for societies that implemented some of the other allocation rules in the β continuum, although the Kantian variation involved for $\beta \notin \{0,\infty\}$ may be too arcane for human societies.

It remains to ask, when we discover an example of a society which appears to implement one of these sharing rules, whether Kantian thinking among its members plays a role in maintaining its stability. Just as a Nash equilibrium is stable, so a Kantian allocation will be stable if the players in the game employ Kantian optimization.

The analogous result to theorem 1 for *Nash* equilibrium is:

Theorem 2

A. There is no sharing rule that is efficiently implementable in Nash equilibrium on the domain $\mathfrak{E}^{0,fin}$.

B. On continuum economies, Walrasian rules are efficiently Nash implementable[9].

Proof: Appendix.

The reason that the Walrasian sharing rule, as defined in the previous footnote, is not efficiently implementable in Nash equilibrium on *finite* economies is that an individual's Nash behavior at the Walrasian sharing rule takes account of her affect on $G'(E^S)$ and on her share of profits as she deviates her effort (i.e., agents are not price takers). It is only in the continuum economy that the agent rationally ignores such affects, and hence, Nash behavior induces efficiency. Of course, this is the point that Makowski and Ostroy (2001) have focused upon in their work on the distinction between perfect competition and Walrasian equilibrium.

[9] A Walrasian rule allocates output to an individual of type γ equal to his value marginal product $E(\gamma)G'(\bar{E})$ plus a fixed share of the firm's profits.

March 1, 2012

3. Economies with a social ethos

A. Efficiency results

We begin by characterizing interior Pareto efficient allocations in continuum economies where individuals have all-encompassing utility functions like those in (1.1), except we use the more general CES social-welfare function. That is, we assume that:

$$U^{\gamma}(x(\cdot),E(\cdot)) = u^{\gamma}(x(\gamma),E(\gamma)) + \alpha\left(\int_0^{\infty} u^{\tau}(x(\tau),E(\tau))]^p \, dF(\tau)\right)^{1/p} , \quad (3.1)$$

where $1 \geq p > -\infty$. As noted, the case $p = 0$ generates the formulation in (1.1).

At an allocation $(x^*(\cdot),E^*(\cdot))$, we write $u^{\gamma}(x^*(\gamma),E^*(\gamma)) \equiv u[*,\gamma]$, and for the two partial derivatives of $u, u_j^{\gamma}(x^*(\gamma),E^*(\gamma)) \equiv u_j[*,\gamma]$.

Theorem 3 *A strictly positive allocation is Pareto efficient in the economy* $(\mathbf{u},G,F,\alpha)$ *if and only if:*

(a) $\quad \forall\gamma \quad \dfrac{u_2[*,\gamma]}{u_1[*,\gamma]} = -G'(\bar{E})$, *and*

(b) $\quad \forall\gamma \quad \dfrac{1}{u_1[*,\gamma]} \geq \dfrac{\alpha(Q^*)^{(1-p)/p} u[*,\gamma]^{p-1} \int u_1[*,\tau]^{-1} \, dF(\tau)}{1+\alpha(Q^*)^{(1-p)/p} \int u[*,\tau]^{p-1} \, dF(\tau)}$,

where $Q^* \equiv \int u[*,\gamma]^p \, dF(\gamma)$.

I offer some remarks about and corollaries to theorem 3.

First, we introduce a *quasi-linear economy* for which the results take a particularly simple and intuitive form. In the quasi-linear economy, we take

$$u^{\gamma}(x,E) = x - \frac{E^2}{\gamma}. \quad (3.2)$$

1. Note the separate roles played by the conditions (a) and (b) of theorem 3. Condition (a) assures allocative efficiency in the economy with $\alpha = 0$. Condition (b) is entirely responsible for the efficiency requirement induced by social ethos. Note that the function G does not appear in (b).

Indeed, it is obvious that any allocation which is Pareto efficient in the α-economy (for any α) must be efficient in the economy with $\alpha = 0$. For suppose not.

Then the allocation in question is Pareto-dominated by some allocation in the 0-economy. But immediately, that allocation must dominate the original one in the α-economy, as it causes the social-welfare function to increase (as well as the private part *u* of all-encompassing utility). It is therefore not surprising that the characterization of theorem 2 says that 'the allocation is efficient in the 0-economy (part (a)) and satisfies a condition which becomes increasingly restrictive as α becomes larger (part (b)).'

2. Define $PE(\alpha)$ as the set of interior Pareto efficient allocations for the α-economy. It follows from condition (b) of theorem 3 that the Pareto sets are nested, that is:

$$\alpha > \alpha' \Rightarrow PE(\alpha) \subset PE(\alpha').$$

Hence, denoting the fully altruistic economy by $\alpha = \infty$, we have:

$$PE(\infty) = \cap_{\alpha \geq 0} PE(\alpha).$$

$PE(\infty)$ will generally be a unique allocation – the allocation that maximizes social welfare.

3. Let $\alpha \to \infty$; then condition (b) of theorem 3 reduces to:

$$\forall \gamma \quad \frac{u_1[*,\gamma]^{-1}}{\int u_1[*,\tau]^{-1}\, dF(\tau)} \geq \frac{u[*,\gamma]^{p-1}}{\int u[*,\tau]^{p-1}\, dF(\tau)}. \tag{3.3}$$

We have:

<u>Corollary 1</u> *An interior allocation is efficient in the fully altruistic economy (i.e., maximizes social welfare) if and only if*:

$$\text{(a)} \ \forall \gamma \quad \frac{u_2[*,\gamma]}{u_1[*,\gamma]} = -G'(\bar{E}),$$

and (c) *for some* $\lambda > 0, \quad \forall \gamma \quad u_1[*,\gamma] = \lambda u[*,\gamma]^{1-p}$.

<u>Proof:</u>

We need only show that (3.3) implies (c). (The converse is obviously true.)

Denote $\lambda = \dfrac{\int u_1[*,\tau]^{-1}\, dF(\tau)}{\int u[*,\tau]^{p-1}\, dF(\tau)}$. Then (3.3) can be written:

$$\forall \gamma \quad u_1[*,\gamma]^{-1} \geq \lambda u[*,\gamma]^{p-1}. \tag{3.4}$$

Suppose there is a set of types of positive measure for which the inequality in (3.4) is slack. Then integrating (3.4) gives us:

$$\int u_1[*,\gamma]^{-1}\,dF(\gamma) > \lambda\int u[*,\gamma]^{p-1}\,dF(\gamma)\,,$$

which says $\lambda > \lambda$, a contradiction. Therefore (3.4) holds with equality for almost all γ, and the corollary follows. ■

4. Consider the quasi-linear economy. Then $u_1 \equiv 1$. Now corollary 1 implies that *in the quasi-linear economy, the only Pareto efficient interior allocation as* $\alpha \to \infty$ *is the equal-utility allocation for which condition* (a) *holds.*

Let us compute this allocation in the quasi – linear economy in which production is linear: $G(x) = x$. Then these conditions reduce to:

(i) $\dfrac{2E(\gamma)}{\gamma} = 1$, and

(ii) $k = x(\gamma) - \dfrac{E(\gamma)^2}{\gamma}$, and

(iii) $\int x(\gamma)\,dF(\gamma) = \int E(\gamma)\,dF(\gamma)$.

It is not hard to show that (i), (ii), and (iii) characterize the equal utility allocation:

$$E(\gamma) = \frac{\gamma}{2}, \quad x(\gamma) = \frac{\gamma+\bar{\gamma}}{4}\text{, where } \bar{\gamma} = \int \gamma\,dF(\gamma).$$

5. Consider the preferences when $p = 0$. In this case, the altruistic part of U is $\exp[\int \log(u[*,\gamma])dF(\gamma)$, and $Q^* = 1$. Therefore condition (b) of theorem 2 becomes simpler:

$$(\forall\gamma)\quad \frac{u[*,\gamma]}{u_1[*,\gamma]} \geq \frac{\alpha\int u_1^{-1}[*,\tau]\,dF(\tau)}{1+\alpha\int u^{-1}[*,\tau]\,dF(\tau)}\,.$$

We next prove:

<u>Theorem</u> <u>4.</u> *Let a sharing rule* θ *be given, and denote the set of* $\beta-$*Kantian equilibria for the economy* $(\mathbf{u},G,F,\alpha)$ by $\mathbf{K}^\beta(\theta,\alpha)$. *Then* $\mathbf{K}^\beta(\theta,\alpha) = \mathbf{K}^\beta(\theta,0)$.

Indeed, the theorem is more general than stated: different agents can have different values of the altruistic parameter α. The argument shows that the Kantian equilibria of these economies are identical to the Kantian equilibria of the associated economy where all $\alpha's$ are zero. This is apparently a disturbing result: for it says that Kantian optimization cannot deal, at least explicitly, with the externalities induced by altruism!

We do, however, have one instrument – namely, β -- which may help achieve Pareto efficient allocations when $\alpha > 0$. Indeed, consider the family of quasi-linear economies, where, for some fixed $\rho > 1$:

$$u^{\gamma}(x,E) = x - \frac{E^{\rho}}{\rho\gamma}. \tag{3.5}$$

For these economies we can always choose a value β so that the K^{β} equilibrium w.r.t. the allocation rule θ_{β} is efficient for economies with any value of α: that is to say, the $(K^{\beta},\theta_{\beta})$ allocation maximizes social welfare (and so is in $PE(\infty)$).

<u>Theorem</u> <u>5</u> *Let* $u^{\gamma}(x,E) = x - \dfrac{E^{\rho}}{\rho\gamma}$, *some* $\rho > 1$. *Let G be any concave production function. Define* $\bar{E}$ *by the equation* $\bar{E} = \bar{\gamma}_{\rho} G'(\bar{E})^{1/(\rho-1)}$ *where* $\bar{\gamma}_{\rho} \equiv \int \gamma^{1/(\rho-1)}\, dF(\gamma)$. *Then for this economy :*

(a) *An allocation is* PE(0) iff $E(\gamma) = \gamma^{1/(\rho-1)} G'(\bar{E})^{1/(\rho-1)}$.

(b) *Define* $\beta(\rho) = \rho \dfrac{G(\bar{E})}{G'(\bar{E})} - \bar{E}$. *The* K^{β} *allocation w.r.t. the sharing rule* θ_{β} *is in* $PE(\infty)$.

(c) *As* $\beta \to \beta(\rho)$ *from below, the maximum value of* α *for which the* $(K^{\beta},\theta_{\beta})$ *allocation is in* $PE(\alpha)$ *approaches infinity.*

The reader is entitled to ask: What happens for $\beta > \beta(\rho)$? The answer is that, in the $(K^{\beta},\theta_{\beta})$ allocation, some utilities become negative, and so social welfare for the CES family of functions is undefined, and so all-encompassing utility *U* is undefined.

B. Taxation in private-ownership economies

The K^{β} equilibria for the sharing rules θ_{β} are not implementable with markets in any obvious way. This is most easily seen by noting that the proportional rule is not so implementable[10]. Of course, according the second theorem of welfare economics, there is some division of shares in the firm which operates *G* which would implement these rules in Walrasian equilibrium in continuum economies, but to compute those shares, one would have to know the preferences of the agents. The advantage of the Kantian approach is that the Kantian allocations are decentralizable in the sense that agents need only know the production function G, average effort $\bar{E}$, and their own preferences, to compute the deviation they would like (everybody) to make.

Nevertheless, one would like Kantian optimization to be useful in market economies as well. For the linear economies, we have a hopeful result. Before stating it, let us define the sharing rules associated with linear taxation. Define the linear sharing rule for *linear* economies with production function $G(x) = ax$ by:

$$\theta^j_{[t]}(E) = (1-t)\frac{E^j}{E^S} + \frac{t}{n}. \qquad (3.6)$$

That is, each agent receives $(1-t)$ times the marginal product of his labor plus an equal share of tax revenues.

<u>Theorem</u> <u>6</u>

A. *For any* $t \in [0,1]$, *the* K^+ *equilibria for the linear tax rule* $\theta_{[t]}$ *is Pareto efficient on* $\mathfrak{L}^{0,fin}$.

B. *The only allocation rules which are efficiently implementable in* K^+ *on* $\mathfrak{L}^{0,fin}$ *are of the form* $\theta^j(E) = \theta^j_{[t]}(E) + \frac{k^j(E)}{G(E)}$ *for some* $t \in [0,1]$ *where:*

(i) *for all E,* $\sum k^j(E) = 0$

[10] However, the equal-division sharing rule is market-implementable. Impose linear taxation in a Walrasian economy and set the tax rate equal to unity. This is equivalent to the equal-division sharing rule.

(ii) *for all* (*j*,*E*) $\theta^j(E) \in [0,1]$, and

(iii) *for all* (*j*,*E*), $\nabla k^j(E) \cdot E = 0$.

Proof: See appendix.

Part A of the theorem states that for finite economies with linear production, linear taxation provides a redistributive mechanism which is consistent with efficiency – for any tax rate in the [0,1] interval. Therefore, in such an economy with a social ethos, citizens could choose a high tax rate to redistribute income substantially, without sacrificing allocative efficiency. Part B of the theorem is analogous to part C of theorem 1.

As in theorem 1, one is entitled to ask whether there are examples of sharing rules where the functions k^j are not identically zero. There are, as the next example shows.

Example 5.

Let *n* = 2, and consider the sharing rule:

$$\theta^1(E) = \begin{cases} (1-t)\dfrac{E^1}{E^S} + \dfrac{t}{2} + \dfrac{t^2(E^1 - E^2)}{2E^S}, \text{ if } E^1 \geq E^2 \\ (1-t)\dfrac{E^1}{E^S} + \dfrac{t}{2} - \dfrac{t^2(E^2 - E^1)}{2E^S}, \text{ if } E^1 \geq E^2 \end{cases}, \qquad (3.7)$$

$$\theta^2(E) = 1 - \theta^1(E)$$

for $t \in (0,1)$. It is easy to verify that these rules satisfy conditions B(i)-(iii), and these rules are clearly not linear tax rules.

We are not interested in linear economies as such, because they are so special. Theorem 6 is presented because it motivates us to ask how linear taxation performs in concave economies with a continuum of agents. Let us postulate that a linear-taxation sharing rule is applied to a person's income, which is equal to his effort times the Walrasian wage plus an equal-per-capita share of the firm's profits. The effort allocation $E(\cdot)$ is a K^+ equilibrium for the *t*-linear tax rule if:

$$(\forall\gamma)\left.\frac{d}{dr}\right|_{r=0} u^{\gamma}((1-t)(E(\gamma)+r)G'(\bar{E}+r)+(1-t)(G(\bar{E}+r)-(\bar{E}+r)G'(\bar{E}+r))+tG(\bar{E}+r),E(\gamma)+r)=0$$

or: $$u_1^{\gamma}\cdot\Big((1-t)(E(\gamma)-\bar{E})G''(\bar{E})+G'(\bar{E})\Big)+u_2^{\gamma}=0\,, \qquad (3.8)$$

and so the marginal rate of substitution of type γ is:

$$-\frac{u_2^{\gamma}}{u_1^{\gamma}}=G'(\bar{E})+(1-t)(E(\gamma)-\bar{E})G''(\bar{E})\,. \qquad (3.9)$$

What is noteworthy is that the wedge between the MRS and the MRT, which is $(1-t)(E(\gamma)-\bar{E})G''(\bar{E})$, goes to zero as t approaches one. Of course, this must be the case, since the allocation at $t=1$ is the kibbutz allocation, which we know is 0-efficient on concave economies. (Of course, (3.9) gives the proof that the linear share rules are Pareto efficient on linear economies.)

Compare (3.9) with Nash-Walras equilibrium in the same private-ownership economy, which is given by:

$$-\frac{u_2^{\gamma}}{u_1^{\gamma}}=(1-t)G'(\bar{E})\,. \qquad (3.10)$$

Here, the wedge between the MRS and the MRT is $tG'(\bar{E})$ which becomes equal to the whole MRT as *t* goes to one. If there is a social ethos, citizens might well wish to redistribute market incomes via taxation. Under Nash optimization, it becomes increasingly costly to do so (as taxes increase), while with K^{+} optimization, equation (3.9) suggests it may become decreasingly costly to do so.

We study this issue with some simulations. I choose $G(x,r)=\frac{x^r}{r}$, for several values of $r\in(0,1)$, and use the quasi-linear utility $u^{\gamma}(x,y)=x-\frac{y^2}{\gamma}$. The distribution *F* is lognormal with a mean of 50 and a median of 40. Let the distribution of profit shares be egalitarian: $\sigma(\gamma)\equiv 1$. (If we desire an *anonymous* Walrasian rule, we must choose this distribution.)

I describe the computational procedure by which the K^+ equilibrium is computed for various tax rates. The characterization of the effort schedule in K^+ equilibrium for the quasi-linear utility profile is given by:

$$(1-t)G''(\bar{E})(E(\gamma)-\sigma(\gamma)\bar{E})+G'(\bar{E})=\frac{2}{\gamma}E(\gamma) \qquad (3.11)$$

For the specified production function, this equation may be solved to yield:

$$E(\gamma,t)=\frac{\bar{E}(t)^{r-1}\gamma(1+(1-r)(1-t))}{2+\gamma(1-r)(1-t)\bar{E}(t)^{r-2}}\,, \qquad (3.12)$$

where $\bar{E}(t)$ is the integral of $E(\gamma,t)\ dF$. Integrating (3.12) and manipulating the result gives an equation in the single unknown $\bar{E}(t)$:

$$1=\int\frac{(1+(1-r)(1-t))\gamma}{2\bar{E}(t)^{2-r}+(1-r)(1-t)\gamma}dF(\gamma)\,. \qquad (3.13)$$

Fixing *r*, we solve (3.13) for $\bar{E}(t)$ numerically, for various values of *t*, and then compute the Kantian equilibrium effort schedule from (3.12). Then we compute social welfare at the various values of *t*.

It is a standard exercise to compute the effort schedule for Walrasian equilibrium. Individual effort is given by $\tilde{E}(\gamma,t)=\frac{(1-t)w\gamma}{2}$, and average effort is given by $\tilde{E}(t)=\frac{(1-t)w\bar{\gamma}}{2}$, where *w* is the Walrasian wage, which solves to be:

$$w=G'(\tilde{E})=\left(\frac{(1-t)\bar{\gamma}}{2}\right)^{(r-1)/(2-r)}\,.$$

We will perform a political-economy simulation. For each voter, we may define an indirect (all-encompassing) utility function which gives her utility at the K^+ equilibrium as a function of the tax rate, and another indirect utility function which gives her (all-encompassing) utility at the Nash-Walras equilibrium as a function of the tax rate. These indirect utility functions are single-peaked in *t*, and so we will assume that the politically chosen tax rate is the ideal tax rate of the median-type voter. (This will be the median ideal tax rate.) We compute these tax rates for various values of the social-ethos parameter α, for both K^+ and Nash-Walras equilibrium. We compare social welfare in these two equilibria, using the social-welfare function that citizens use in their all-encompassing utility functions.

Tables 1a and 1b report results for $r = 0.90$ and $r = 0.50$. In the first case, the maximum admissible tax rate is about 0.70, because for higher rates, some utilities become negative, and the social-welfare function is undefined. For each value of α, I compute the ideal tax rate of the median type at the Kantian and Walrasian equilibrium, and report the values of social welfare at those political equilibria. For $r = 0.5$, the maximum admissible tax rate is about 0.8. In both cases, it turns out that the ideal tax rate of the median type, in the Kantian regime, is the maximum admissible rate. We see from the tables that the ideal tax rate of the median type, in the Walrasian regime, is much smaller, and decreases slightly as α increases.

Out[44]//TableForm=

alpha	t-Kant	t-Walras	Soc Wel @ Kant	Soc Wel @ Walras
0.	0.7	0.166667	8.47076	7.72644
0.1	0.7	0.167536	8.47076	7.72656
0.2	0.7	0.168222	8.47076	7.72665
0.3	0.7	0.168778	8.47076	7.72672
0.4	0.7	0.169238	8.47076	7.72676
0.5	0.7	0.169624	8.47076	7.7268
0.6	0.7	0.169953	8.47076	7.72683
0.7	0.7	0.170237	8.47076	7.72686
0.8	0.7	0.170484	8.47076	7.72687
0.9	0.7	0.170701	8.47076	7.72689
1.	0.7	0.170894	8.47076	7.7269

<u>Table 1a</u> Political-equilibrium tax rates and social welfare in Kantian and Walrasian regimes, for the quasi-linear economy with $G(x) = x^{0.9} / 0.9$ and $\sigma(\gamma) \equiv 1$

Out[98]//TableForm=

alpha	t-Kant	t-Walras	Soc Wel @ Kant	Soc Wel @ Walras
0.	0.8	0.166667	4.32975	4.28841
0.1	0.8	0.160879	4.32975	4.28933
0.2	0.8	0.156081	4.32975	4.29003
0.3	0.8	0.152041	4.32975	4.29057
0.4	0.8	0.148592	4.32975	4.29099
0.5	0.8	0.145615	4.32975	4.29133
0.6	0.8	0.143019	4.32975	4.2916
0.7	0.8	0.140736	4.32975	4.29183
0.8	0.8	0.138712	4.32975	4.29202
0.9	0.8	0.136906	4.32975	4.29218
1.	0.8	0.135285	4.32975	4.29231

Table 1b Political-equilibrium tax rates and social welfare in Kantian and Walrasian regimes, for the quasi-linear economy with $G(x) = x^{0.5} / 0.5$ and $\sigma(\gamma) \equiv 1$

This is a consequence of the deadweight loss experienced with taxation in the Walrasian regime. We see that, even with substantial concavity, the political equilibrium in the Kantian regime dominates that of the Walrasian regime in terms of social welfare, at least for values of α in $[0,1]$.

4. Existence and dynamics

The existence of *proportional solutions*, which are the $K^{\times}$ equilibria of convex economies $(\mathbf{u}, G, F, \alpha)$ was proved in Roemer and Silvestre (1993). Here, we provide conditions under which $\beta-$Kantian equilibria exist, with respect to the sharing rules described in Theorem 1.

Theorem 7. . *Let* $\xi \in \mathfrak{E}^{fin}$. Let the component functions of $\mathbf{u}$ be strictly concave.

A. *If for all* $u \in \mathbf{u}$, $\frac{\partial^2 u}{\partial x \partial y} \leq 0$, *then a strictly positive* K^{+} *equilibrium w.r.t. the equal-division sharing rule* θ^{ED} *exists on* ξ.

B. *Let* $0 \leq \beta < \infty$. *If for all* $u \in \mathbf{u}$, *u is quasi-linear, then a strictly positive* $\beta-$*Kantian equilibrium w.r.t. the sharing rule* θ_{β} *exists.*

Proof: Appendix.

The premises of this theorem can surely be weakened.

We turn briefly to dynamics. There will not be robust dynamics for Kantian equilibrium, as there are not for Nash equilibrium. There is, however, a simple dynamic mechanism that will, in well-behaved cases, converge to a Kantian equilibrium from any initial effort vector. The mechanism is based on the mapping Θ defined in the proof of theorem 7. We illustrate it here for the case of a profile of quasi-linear utility functions and the equal-division sharing rule. Thus, let $u^j(x, y) = x - c^j(y)$, for $j = 1, ..., n$, where

c^j is a strictly convex function. For any vector $E_0 \in \mathbb{R}^n_{++}$, define $r^j(E_0)$ as the unique solution of:

$$\max_r \left(\frac{G(E_0 + nr)}{n} - c^j(E_0^j + r) \right). \tag{4.1}$$

Define $\Theta^j(E_0) = E_0^j + r^j(E_0)$. The mapping $\Theta = (\Theta^1,...,\Theta^n)$ maps $\mathbb{R}^n_+ \to \mathbb{R}^n_+$ and is analogous to the best-reply correspondence in Nash equilibrium. A fixed point of Θ is a K^+ equilibrium for the equal-division sharing rule, since at a fixed point E^*, $r^j(E^*) = 0$ for all *j*. Since the example is special, the next result is proved only for the case *n* = 2, although it is true for finite *n*. The next proposition shows that if we iterate the mapping Θ indefinitely from any initial starting vector E_0 it converges to (the unique) K^+ equilibrium for the equal-division sharing rule.

<u>Proposition</u> <u>2</u> *For n = 2, there exists a unique fixed point of the mapping Θ, which is a K^+ equilibrium for the equal-division sharing rule with quasi-linear preferences. The dynamic process defined by iterating the application of Θ from any initial effort vector converges to the K^+ equilibrium.*
<u>Proof</u>: Appendix.

5. <u>Discussion</u>

My analysis has been positive rather than normative. I have argued that if agents optimize in the Kantian way, then certain allocation rules will produce Pareto efficient allocations, while Nash optimization will not. While the *analysis* is positive, Kantian optimization, if people follow it, is motivated by a moral point of view: each must think that he should take an action if and only if he would advocate that all others take a similar action. Is it plausible to think that there are (or could be) societies where individuals do (or would) optimize in the Kantian manner?

Certainly parents try to teach Kantian behavior to their children, at least in some contexts. "Don't throw that candy wrapper on the ground: How would you feel if everyone did so?" The golden rule ("Do unto others as you would have them do unto you") is a special case of Kantian ethics. (And wishful thinking ["if I do *X*, then all those who are similarly situated to me will do *X*"], although a predictive claim, rather than an ethical one, will also induce Kantian equilibrium – if all think that way.) This may explain why people vote in large elections, and make charitable contributions. So there is some reason to believe that Kantian equilibria are accessible to human societies.

Consider the relationship between the theoretical concept of Nash equilibrium and the empirical evidence that agents play the Nash equilibrium in certain social situations that can be modeled as games. We do not claim that agents are consciously computing the Nash equilibrium of the game: rather, we believe there is some process by which players *discover* the Nash equilibrium, and once it is discovered, it is stable, given autarkic reasoning. We now know there are many experimental situations in which players in a game do not play (what we think is) the Nash equilibrium. Conventionally, this 'deviant' behavior has been rationalized by proposing that players have different payoff functions from the ones that the experimenter is trying to induce in them, or that they are adopting behavior that is Nash in repeated games generated by the one-shot game under consideration. Another possibility, however, is that players in these games are playing some kind of Kantian equilibrium. In Roemer (2010), I showed that if, in the prisoners' dilemma game, agents play mixed strategies on the two pure strategies of {cooperate, defect}, then all multiplicative Kantian equilibria entail both players' cooperating with probability at least one-half (i.e., no matter how great the payoff to defecting is). It can also be shown that, in a stochastic dictator game, where the dictator is chosen randomly at stage 1 and allocates the pie between herself and the other player in stage 2, the unique $K^{\times}$ equilibrium is that each player gives one-half the pie to the other player, if he is chosen.

The non-experimental (i.e., real-world) counterpart, as I have said in the introduction, may be the games that the societies that Elinor Ostrom has studied are playing. If these games can be modeled as 'fisher' economies, with common ownership of a resource whose use displays negative congestion externalities, and if, as Ostrom contends, these societies figure out how to engender efficient allocations of labor applied to the common resource, then they are discovering the multiplicative Kantian equilibrium of the game. Perhaps Kantian reasoning helps to maintain the equilibrium, if optimizing behavior is 'interdependent' and not 'autarkic.' Ostrom explains the maintenance of the efficient labor allocation by the use of sanctions and punishments, but that may not be the entire story: it may be that most fishers are thinking in the Kantian manner, and that punishments and monitoring are needed only to control a minority who are Nash optimizers. What I am proposing is that an ethic may have evolved, in these societies, in which the fisher says to himself, "I would like to increase my fishing time by 5 hours a week, but I have a right to do so only if all others could similarly increase their fishing times, and that I would not like." Armed only with the theory of Nash equilibrium, one naturally thinks that these Pareto efficient solutions to the tragedy of the commons require punishments to keep *everyone* in line. But this may not be the case.

As I noted earlier, Kantian ethics, and therefore the behavior they induce, require *less* selflessness than another kind of ethic: putting oneself in the shoes of others. Consider charity. "I should give to the unfortunate, because I could have been that unfortunate soul – indeed, there but for the grace of God go I." The Kantian ethic says, "I will give to the unfortunate an amount which I would like all others who are similarly situated to me to give." Assuming that there is a social ethos (that is, $\alpha > 0$) this kind of reasoning may induce substantial charity – or, in the political case, fiscal redistribution. The Kantian ethic does not require the individual to place herself in the shoes of another. In this sense, it requires a less radical departure from self than the 'grace of God' reasoning does.

My analysis has studied the consequences of assuming that the optimizing behavior of individuals might not be autarkic, as in Nash equilibrium, but interdependent, as in the various kinds of Kantian optimization. To the extent that human societies have prospered by invoking the ability of individuals of members of our species to cooperate with each other, it is perhaps likely that Kantian reasoning is a cultural adaptation, selected by evolution (the classic reference is Boyd and Richerson [1985]). Because we have shown that Kantian behavior can resolve, in many cases, the inefficiency of autarkic behavior, cultures which discover it, and attempt to induce that behavior in their members, will thrive relative to others. Group selection may produce Kantian optimization as a meme.

One can rightfully ask whether it is utopian to suppose that the allocation rules studied here can be used in large economies. I am skeptical, because the market mechanism is so important in large economies, and most of the allocation rules described in theorem 1 cannot be decentralized using markets. (I noted that the equal-division rule can be.) This motivated my simulations of the linear-tax sharing rules where the market allocation is Walrasian. We do not get full Pareto efficiency, but the results are much better when agents are Kantian than when they are Nash optimizers.

One of the main motivations I gave for studying Kantian optimization was in order to resolve the inefficiencies in economies with a social ethos, due to the consumption externalities that they entail. It seems that, if a society is solidaristic in the sense of possessing a social ethos, then its members should *behave* in a cooperative fashion. The behavior upon which I have focused in this paper is optimizing behavior. I leave the reader with a question. Is there reason to think that *if* a society is characterized by having a high degree of social ethos, it is *more likely* that its members can learn to optimize in the Kantian manner? My intuition indicates this is probably so, but I cannot yet provide an argument to show it.

References

Alger, I. and J.W. Weibull, 2012. "Homo Moralis: Preference evolution under incomplete information and assortative matching," http://pzp.hhs.se/media/5245/homomoralis120206_wp.pdf

Arneson, R. 1989. "Equality and equality of opportunity for welfare," *Phil. Stud.* 56, 77-93

Bowles, S. and H. Gintis, 2011. *A cooperative species: Human reciprocity and its evolution,* Princeton: Princeton University Press

Boyd, R. and P. Richerson, 1985. *Culture and the evolutionary process,*Chicago: University of Chicago Press

Cohen, G.A. 1989. "On the currency of egalitarian justice," *Ethics* 99, 906-944

Cohen, G.A. 2010. *Why not socialism?* Princeton: Princeton University Press

Dufwenberg, M., P. Heidhues, G. Kirchsteiger, F. Riedel, and J. Sobel, 2011. "Other regarding preferences and general equilibrium," *Rev. Econ. Stud.* 78, 613-639

Dworkin, R. 1981. "What is equality? Part 2: Equality of resources," *Phil. & Pub. Affairs* 10, 283-345

Harsanyi, J. 1977. *Rational behavior and bargaining equilibrium in games and social situations,* Cambridge University Press

Henrich, N. and J. Henrich, 2007. *Why humans cooperate: A cultural and evolutionary explanation*, Oxford: Oxford University Press

Maskin, E. 1999. "Nash equilibrium and welfare optimality," *Rev. Econ. Stud.* 66, 23-38

Makowski, L. and J. Ostroy, 2001. "Perfect competition and the creativity of the market,"*J.Econ. Lit.* 39, 479-535

Ostrom, E. 1990. *Governing the commons: The evolution of institutions for collective action,* Cambridge: Cambridge University Press

Rabin, M. 2006. "The experimental study of social preferences,"*Social Research* 73, 405-428

Rawls, J. 1971. *A theory of justice,* Cambridge, MA: Harvard University Press

Roemer, J. 1996. *Theories of distributive justice*, Cambridge, MA: Harvard University Press

Roemer, J. 2010. "Kantian equilibrium," *Scandinavian J. Econ.* 112, 1-24

Roemer, J. and J. Silvestre, 1993. "The proportional solution for economies with both private and public ownership," *J. Econ. Theory* 59, 426-444

Scheve, K. and D. Stasavage, in press. "The conscription of wealth: Mass warfare and the demand for progressive taxation," *International Organization*

(for “Kantian optimization, social ethos, and Pareto efficiency” by J.E. Roemer)

“**Appendix: Proofs of theorems**”

Proof of Theorem 1.

The proof of part A simply mimics the proof of Proposition 1. We prove part B.

1. Consider the Kantian variation $\varphi^{\beta}(x,r) = rE + (r-1)\beta$, and any share rule $\{\theta^{j}, j = 1,...,n\}$, defined for a finite economy with *n* agents.

The condition that must hold for a rule θ to be efficiently implemented on $\mathfrak{E}$ in K^{β} equilibrium is the FOC:

$$(\forall j) \quad \frac{\nabla\theta^{j}(E)\cdot(E+\beta)G(E^{S}) + \theta^{j}(E)G'(E^{S})(E^{S}+n\beta)}{E^{j}+\beta} = G'(E^{S}), \qquad \text{(A.1)}$$

which is the statement that that at a K^{β} equilibrium $E = (E^{1},...,E^{n})$, the marginal rate of substitution between effort and income for each agent is equal to the marginal rate of transformation. Recall that $E^{S} \equiv \sum E^{j}$, $\nabla\theta^{j}$ is the gradient of the function θ^{j} with respect to its *n* arguments, $E+\beta$ is the vector whose *j*th component is $E^{j}+\beta$, and $\nabla\theta^{j}(E)\cdot(E+\beta)$ is the scalar product of two *n* vectors. (A.1) can be written as:

$$(\nabla\theta^{j}(E)\cdot\frac{(E+\beta)}{E^{j}+\beta})\frac{G(E^{S})}{G'(E^{S})} + \theta^{j}(E)\frac{(E^{S}+n\beta)}{E^{j}+\beta} = 1. \qquad \text{(A.2)}$$

2. We now argue that (A.2) must hold as a set of partial differential equations on $\mathbb{R}^{n}_{++}$. For let $E \in \mathbb{R}^{n}_{++}$ be any vector. Fix a production function *G*. We can always construct *n* utility functions whose marginal rates of substitution at the points $(\theta^{j}(E), E^{j})$ are exactly given by the value of the left-hand side of equation (A.1). For the economy thus defined, *E* is indeed a K^{β} equilibrium. This demonstrates the claim.

3. Continue to fix a vector $E \in \mathbb{R}^{n}_{++}$. Define $r_{i}^{j} = \dfrac{E_{i}+\beta}{E_{j}+\beta}$ for $i = 1,...,n$ and notice that $\varphi_{\beta}(E_{j}, r_{i}^{j}) = E_{i}$. Consider the one dimensional manifold gotten by varying *x*: $\mathfrak{M}_{E}^{j} = (\varphi_{\beta}(x,r_{1}),\varphi_{\beta}(x,r_{2}),...,\varphi_{\beta}(x,r_{n}))$. Note that when $x = E_{j}$, this picks out the vector *E*.

We will reduce the system (A.2) of PDEs to ordinary differential equations on $\mathfrak{M}_{E}^{j}$.

Define $\psi^j(x) = \theta^j(\varphi_\beta(x, r_1^j), \ldots, \varphi_\beta(x, r_n^j))$. Note that :

$$(\psi^j)'(x) = \nabla\theta^j(\varphi_\beta(x, r^j)) \cdot r^j \tag{A.3}$$

where $\varphi_\beta(x, r)$ is the generic vector in the manifold, and $r^j = (r_1^j, \ldots, r_n^j)$.

Define $\mu^j(x) = G(\sum \varphi(x, r_i^j))$ and note that:

$$(\mu^j)'(x) = G'(\sum \varphi(x, r_i^j)) \sum r_i^j . \tag{A.4}$$

It follows that we may write (A.2) restricted to the manifold $\mathfrak{M}_E^j$ as:

$$(\psi^j)'(x) r^{S,j} \frac{\mu^j(x)}{(\mu^j)'(x)} + \mu^j(x) r^{S,j} = 1, \tag{A.5}$$

where $r^{S,j} \equiv \sum_i r_i^j$.

4. (A.5) is a first-order ODE. A particular solution is given by the constant function:

$$\psi^j(x) = \frac{1}{r^{S,j}}, \tag{A.6}$$

and the general solution to its homogeneous variant is:

$$\hat{\psi}^j(x) = \frac{k^j(r^j)}{\mu^j(x)}, \tag{A.7}$$

where k^j a constant that depends on r^j (i.e., on the manifold $\mathfrak{M}_E^j$). Therefore the general solution of (A.5) is

$$\psi^j(x) = \frac{1}{r^{S,j}} + \frac{k^j(r^j)}{\mu^j(x)}. \tag{A.8}$$

Now, evaluating this equation at $x = E^j$ gives:

$$\psi^j(E^j) = \theta^j(E) = \frac{1}{r^{S,j}} + \frac{k^j(r^j)}{G(E)} = \frac{E^j + \beta}{\sum(E^i + \beta)} + \frac{k^j(r^j)}{G(E)}. \tag{A.9}$$

Since the n shares in (A.8) sum to one, (A.8) tells us that we must have $\sum k^j(r^j) = 0$.

This proves part C.

5. To prove part B, return to equation (A.8) which holds on the manifold $\mathfrak{M}_E^j$. For $\beta = 0$ (i.e., $K^\times$ equilibrium), the manifold $\mathfrak{M}_E^j = \{(r_1^j x, \ldots, r_n^j x) \mid x \geq 0\}$. Hence, as x approaches zero $\mu^j(x)$ approaches zero. If, for some j, $k^j(r^j) \neq 0$, then for sufficiently small x, $\psi^j(x)$ would violate the constraint that it lie in $[0,1]$. Hence, for the case when

$\beta = 0$ (and only for that case) we may conclude that the constants k^j are identically zero, and the claim of part B follows. ■

Proof of Theorem 2:

1. An interior allocation E is Nash implementable on the class of finite convex economies for the sharing rule θ if and only if

$$\forall j \quad u_1^j \cdot (\frac{\partial \theta^j(E)}{\partial E_j} G(E^S) + \theta^j(E) G'(E^S)) + u_2^j = 0 \tag{A.10}$$

Therefore θ is efficiently implementable iff:

$$\forall j \quad 1 = \theta^j(E) + \frac{G(E^S)}{G'(E^S)} \frac{\partial \theta^j(E)}{\partial E_j}. \tag{A.11}$$

2. Indeed, (A.11) must hold for the entire positive orthant $\mathbb{R}^n_{++}$, for given any positive vector E, we can construct n concave utility functions such that (A.10) holds at E.

3. For fixed E, define $\psi^j(x) = \theta^j(E_1, E_2, ..., E_{j-1}, x, E_{j+1}, ..., E_n)$ and $\mu^j(x) = G(x + E^S - E_j)$. Then (A.11) gives us the differential equation:

$$1 = \psi^j(x) + \frac{\mu^j(x)}{(\mu^j)'(x)} (\psi^j)'(x), \tag{A.12}$$

which must hold on $\mathbb{R}_{++}$.

4. But (A.12) implies that

$$\frac{(\psi^j)'(x)}{1 - \psi^j(x)} = \frac{(\mu^j)'(x)}{\mu^j(x)} \tag{A.13}$$

which implies that $\mu^j(x)(1 - \psi^j(x)) = k^j$ and therefore $\psi^j(x) = 1 - \dfrac{k^j(E^{-j})}{\mu^j(x)}$ where the constant k^j may depend on the manifold $(E^1, .., E^{j-1}, x, E^{j+1}, .., E^n)$ on which ψ^j is defined.

.

5. In turn, this last equation says that on the unidimensional space $(E_1, ..., E_{j-1}, x, E_{j+1}, ..., E_n)$ we have:

$$\theta^j(E_1, ..., E_{j-1}, x, E_{j+1}, ..., E_n) G(x + E^S - E_j) = G(x + E^S - E_j) - k^j(E^{-j}), \tag{A.14}$$

which says that ‘every agent receives his entire marginal product’ on this space. To be precise:

$$(\forall x, y > 0)$$
$$(\theta^j(E_1,...,E_{j-1},x,E_{j+1},...,E_n)G(x+E^S-E_j)-\theta^j(E_1,...,E_{j-1},y,E_{j+1},...,E_n)G(y+E^S-E_j)=$$
$$G(x+E^S-E_j)-G(y+E^S-E_j)) \tag{A.15}$$

Now let $y=0$ and $x=E_j$ and let $z_j = G(E^S-E_j)$, Then (A.15) says that:

$$(\forall j)(\theta^j(E)G(\bar{E})-z_j = G(\bar{E})-G(\bar{E}-E_j)). \tag{A.16}$$

6. Adding up the equations in (A.16) over *j*, and using the fact that $z_j \geq 0$, we have:

$$G(\bar{E}) \geq nG(\bar{E}) - \sum G(\bar{E}-E_j) \tag{A.17}$$

or:

$$G(\bar{E}) \leq \frac{1}{n-1}\sum G(\bar{E}-E_j). \tag{A.18}$$

7. Now note that $\frac{1}{n-1}\sum(E^S-E_j)=E^S$. Therefore (A.18) can be written:

$$G(\frac{1}{n-1}\sum(E^S-E_j)) \leq \frac{1}{n-1}\sum G(E^S-E_j), \tag{A.19}$$

which is impossible for any strictly concave *G*. This proves part A of the theorem.

8. The proof of part B is familiar: for part B just says that Nash behavior, taking prices as given, at the Walrasian sharing rule, induces Pareto efficiency.

■

<u>Proof of Theorem 3:</u>

Consider the program:

$$\max_{K,h(),q()} \int_{\tau\in D} u^\tau(x^*(\tau)+h(\tau),E^*(\tau)+q(\tau))dF(\tau)+\alpha F(D)K$$

subject to

$$\forall\gamma \quad u^\gamma(x^*(\gamma)+h(\gamma),E^*(\gamma)+q(\gamma))+\alpha K \geq u^\gamma(x^*(\gamma),E^*(\gamma))+\alpha K^*$$
$$\forall\gamma \quad x^*(\gamma)+h(\gamma)\geq 0$$
$$\forall\gamma \quad E^*(\gamma)+q(\gamma)\geq 0$$
$$K \leq \left(\int u^\gamma(x^*(\gamma)+h(\gamma),E^*(\gamma)+q(\gamma))^p\, dF(\gamma)\right)^{1/p}$$
$$G(\int(E^*(\gamma)+q(\gamma))dF(\gamma)) \geq \int(x^*(\gamma)+h(\gamma))dF(\gamma)$$

where D is any set of types of positive measure. Suppose the solution to this program is $h^* \equiv 0, q^* \equiv 0, K = K^*$. (K* is the value of the social-welfare function – given in the K constraint in the program -- when $h = q = 0$.) Then $(x^*(\cdot), E^*(\cdot))$ is a Pareto efficient allocation. Since we are studying strictly positive allocations, the second and third sets of constraints at the proposed optimal solution will be slack.

We will show that conditions (a) and (b) of the proposition characterize the * allocations for which this statement is true. Let (h,q,K) be any feasible triple in the above program, for a fixed positive allocation (x^*, E^*). Let $\Delta K = K - K^*$. Then define the Lagrange function:

$$
\begin{aligned}
\Delta(\varepsilon) = &\int_{\tau\in D} u^{\tau}(x^*(\tau)+\varepsilon h(\tau), E^*(\tau)+\varepsilon q(\tau))dF(\tau) + \alpha F(D)(K^* + \varepsilon\Delta K) + \\
&\rho\Big(G(\int (E^*(\tau)+\varepsilon q(t))\,dF(\tau) - \int (x^*(\tau)+\varepsilon h(\tau))\,dF(\tau)\Big) + \lambda\left(\int u^{\tau}(x^*(\tau)+\varepsilon h(\tau), E^*(\tau)+\varepsilon q(\tau))^p\,dF(\tau)\right)^{1/p} - \\
&\lambda\Big(K^* + \varepsilon\Delta K\Big) + \int B(\gamma)(u(x^*(\tau)+\varepsilon h(\tau), E^*(\tau)+\varepsilon q(\tau),\tau) + \alpha\varepsilon\Delta K - u(x^*(\tau), E^*(\tau),\tau))dF(\tau).
\end{aligned}
$$

Suppose there is non-negative function $B(\cdot)$ and non-negative numbers (λ,ρ) for which the function Δ is maximized at zero. Note $\Delta(0)$ is the value of the objective of the above program, when $h^* \equiv 0 \equiv q^*$ and $K = K^*$, and $\Delta(1)$ equals the value of the objective at (h,q,K) plus some non-negative terms. The claim will then follow. Since Δ is a concave function, it suffices to produce an allocation (x^*, E^*) for which non-negative (B,λ,ρ) exist such that $\Delta'(0) = 0$.

Compute the derivative of Δ at zero:

$$\Delta'(0) = \int_D \Big(u_1[*,\gamma]h(\gamma) + u_2[*,\gamma]q(\gamma)dF(\gamma)\Big) + \alpha F(D)\Delta K +$$
$$\rho\Big(G'(\int E^*(\tau)\,dF(\tau))\int q(\tau)\,dF(\tau) - \int h(\tau)\,dF(\tau)\Big) +$$
$$\frac{\lambda}{p}(Q^*)^{(1-p)/p}\, p\int u[*,\gamma]^{p-1}\Big(u_1[*,\gamma]h(\gamma) + u_2[*,\gamma]q(\gamma)\Big)dF(\gamma) -$$
$$\lambda\Delta K + \int B(\gamma)\Big(u_1[*,\gamma]h(\gamma) + u_2[*,\gamma]q(\gamma) + \alpha\Delta K\Big)dF(\gamma).$$

We now gather together the coefficients of $\Delta K, h,$ and q in the above expression and set them equal to zero:

Coefficient of ΔK: $\quad \alpha F(D) + \alpha\int B(\gamma)\,dF(\gamma) - \lambda = 0 \qquad$ (A.9)

Coefficient of $h(\gamma)$: $\quad u_1[*,\gamma]\mathbf{1}_D - \rho + \lambda(Q^*)^{(1-p)/p}u[*,\gamma]^{p-1}u_1[*,\gamma] + B(\gamma)u_1[*,\gamma] = 0$,

(A.10)

Coefficient of $q(\gamma)$: $\quad u_2[*,\gamma]\mathbf{1}_D + \rho G'(\bar{E}) + \lambda(Q^*)^{(1-p)/p}u[*,\gamma]^{p-1}u_2[*,\gamma] + B(\gamma)u_2[*,\gamma] = 0$,

(A.11)

where $\mathbf{1}_D(\gamma) = \begin{cases} 1, \text{ if } \gamma \in D \\ 0, \text{ if } \gamma \notin D \end{cases}$ and $\bar{E} = \int E^*(\gamma)\,dF(\gamma)$.

By setting all these coefficients equal to zero, and solving for the Lagrange multipliers, we will discover the characterization of the allocation $(x^*(), E^*())$. Note that, at an interior Pareto efficient solution, we must have:

$$\frac{u_2[*,\gamma]}{u_1[*,\gamma]} = -G'(\bar{E}) ,$$

for this is the statement that the marginal rate of substitution for each type between labor and output is equal to the marginal rate of transformation between labor and output. Therefore write:

$$u_1[*,\gamma] + u_2[*,\gamma] = u_1[*,\gamma]\left(1 + \frac{u_2[*,\gamma]}{u_1[*,\gamma]}\right) = u_1[*,\gamma]\Big(1 - G'(\bar{E})\Big). \quad \text{(A.12)}$$

Now add together the equations for the coefficients of $q(\gamma)$ and $h(\gamma)$, divide this new equation by $1-G'(\bar{E})$, use equation (A.12), and the result is exactly the equation (A.11). Therefore, eqn. (A.12) has enabled us to eliminate equation (A.11): if we can produce non-negative values $(B(\cdot),\lambda,\rho)$ satisfying (A.9) and (A.10), we are done.

Solve eqn. (A.10) for $B(\gamma)$:

$$B(\gamma)=\frac{\rho-u_1[*,\gamma]\mathbf{1}_D-u_1[*,\gamma]\lambda(Q^*)^{(1-p)/p}u[*,\gamma]^{p-1}}{u_1[*,\gamma]}\ . \quad \text{(A.13)}$$

From eqn. (A.9), we have $\lambda=\alpha F(D)+\alpha\int B(\gamma)\,dF(\gamma)$, and substituting the expression for $B(\gamma)$ into this equation, we integrate and solve for λ:

$$\lambda=\frac{\alpha\rho\int u_1[*,\gamma]^{-1}\,dF(\gamma)}{1+\alpha(Q^*)^{(1-p)/p}\int u[*,\gamma]^{p-1}\,dF(\gamma)} \quad \text{(A.14).}$$

Eqn. (A.13) says that $B(\gamma)$ is non-negative if and only if

$$\rho\geq u_1[*,\gamma](\mathbf{1}_D+\lambda(Q^*)^{(1-p)/p}u[*,\gamma]^{p-1})\ ; \qquad \text{(A.15)}$$

substituting the expression for λ from (A.14) into (A.15) yields an inequality in ρ which, by rearranging terms, can be written as:

$$\rho\left(1-u_1[*,\gamma]\frac{\alpha(Q*)^{(1-p)/p}u[*,\gamma]^{p-1}\int u_1[*,\tau]^{-1}\,dF(\tau)}{1+\alpha(Q*)^{(1-p)/p}\int u[*,\tau]^{[-1}\,dF(\tau)}\right)\geq u_1[*,\gamma]\,. \qquad \text{(A.16)}$$

In sum, we can find non-negative Lagrange multipliers iff we can produce a non-negative number ρ such that (A.16) is true for all γ. This can be done iff:

$$\forall\gamma \quad \frac{1}{u_1[*,\gamma]}\geq\frac{\alpha(Q^*)^{(1-p)/p}u[*,\gamma]^{p-1}\int u_1[*,\tau]^{-1}\,dF(\tau)}{1+\alpha(Q^*)^{(1-p)/p}\int u[*,\tau]^{p-1}\,dF(\tau)},$$

proving the theorem. ■

Proof of Theorem 4.

We prove the generalization of the theorem stated in the text. We prove the result for $K^{\times}$ equilibrium for simplicity's sake, although the proof for K^{β} equilibrium is the same. Also for simplicity's sake, we use the social-welfare function of (1.1).

1. For the sharing rule θ, an allocation E is a $K^{\times}$ equilibrium iff:

$$\frac{d}{dr}\Big|_{r=1}\left(u^{\gamma}(\theta^{\gamma}(rE)G(r\bar{E}),rE(\gamma))+\alpha^{\gamma}\exp\int\log(u^{\tau}(\theta^{\tau}(rE)G(r\bar{E}),rE(\tau))\,dF(\tau)\right)=0\,, \quad (A.17)$$

where we assume that the altruism parameters $\{\alpha^{\gamma}\}$ are non-negative. Expand this derivative, writing it as:

$$(\forall\gamma)\; D^{\gamma}(E)+\alpha^{\gamma}\exp\int\log(u^{\tau}(\theta^{\tau}(E)G(\bar{E}),E(\tau))dF(\tau)\left(\int\frac{D^{\tau}(E)}{u^{\tau}}dF(\tau)\right)=0, \quad (A.18)$$

where $D^{\tau}(E)=\frac{d}{dr}\Big|_{r=1} u^{\tau}(\theta^{\tau}(rE)G(r\bar{E}),rE(\tau))$.

2. Now (A.18) says that :

$$(\forall\gamma)(D^{\gamma}(E)=-\alpha^{\gamma}k)$$

where k is a constant (independent of γ). Therefore we can substitute $-\alpha^{\tau}k$ for $D^{\tau}(E)$ on the r.h.s. of eqn. (A.18), and re-write that equation as:

$$-\alpha^{\gamma}k-\alpha^{\gamma}km=0\,, \quad (A.\ 19)$$

where m is a positive constant. If $\alpha^{\gamma}=0$, we have from (A.18) that $D^{\gamma}(E)=0$. Id $\alpha^{\gamma}\neq 0$, it follows from (A.19) that $k=0$. But this means that for all γ, $D^{\gamma}(E)=0$, which is exactly the condition that E is a Kantian equilibrium for the economy with $\alpha=0$. ■

<u>Proof of Theorem 5:</u>

1. The effort allocation in part (a) maximizes the surplus, which is the condition for efficiency in the quasi-linear economy with $\alpha=0$.

2. Integrating the expression for $E(\gamma)$, we have that the equation $\bar{E}=\bar{\gamma}_{\rho}G'(\bar{E})^{1/(\rho-1)}$,characterizing $\bar{E}$.

3. To prove claim (b), we show that the $\beta(\rho)$ -Kantian equilibrium produces equal utilities across γ. From Remark 4 stated after Theorem 2, this suffices to show that the allocation will be in $PE(\infty)$. We have:

$$u[\gamma,\beta] = \frac{\gamma^{1/(\rho-1)} G'(\bar{E})^{1/(\rho-1)} + \beta}{\bar{\gamma}_\rho G'(\bar{E})^{1/(\rho-1)} + \beta} G(\bar{E}) - \frac{\gamma^{\rho/(\rho-1)} G'(\bar{E})^{\rho/(\rho-1)}}{\rho\gamma} = \\ \gamma^{1/(\rho-1)} \left(\frac{G'(\bar{E})^{1/(\rho-1)} G(\bar{E})}{\bar{\gamma}_\rho G'(\bar{E})^{1/(\rho-1)} + \beta} - \frac{G'(\bar{E})^{\rho/(\rho-1)}}{\rho} \right) + k \tag{A.17}$$

where k is a constant independent of γ. Calculation shows that the value of β that causes the coefficient of $\gamma^{1/(\rho-1)}$ in (A.17) to vanish is $\beta(\rho)$ as defined in claim (b). It is easy to observe that $\beta(\rho) > 0$ by the concavity of G, and because $\rho > 1$. This proves claim (b).

4. Claim (c) follows from analyzing the condition (b) of theorem 2, which for quasi-linear economies is:

$$(\forall\gamma) \quad 1 + \alpha \int u[*,\tau]^{-1}\, dF(\tau) \geq \alpha u[*,\gamma]^{-1},$$

as β approaches $\beta(\rho)$ from below. ■

<u>Proof</u> <u>of</u> <u>Theorem</u> <u>6:</u>

1. A simple calculation shows that if E is a K^+ equilibrium for an economy with a linear production function $G(x) = ax$ w.r.t. *any* linear tax sharing rule $\theta_{[t]}$, for $t \in [0,1]$, then the allocation is 0-Pareto efficient.

2. Now let E be a K^+ equilibrium w.r.t. any sharing rule θ on $(\mathbf{u}, G, F, 0)$ which is Pareto efficient on that economy. E is a K^+ equilibrium means:

$$u_1^j \left((\nabla\theta^j(E) \cdot \mathbf{1}) a E^S + \theta^j(E) a n \right) + u_2^j = 0,$$

and so Pareto efficiency means that:

$$\left((\nabla\theta^j(E) \cdot \mathbf{1}) a E^S + \theta^j(E) a n \right) = a,$$

or:

$$(\nabla\theta^j(E) \cdot \mathbf{1}) E^S + n\theta^j(E) = 1. \tag{A.18}$$

As has been argued in previous proofs, (A.18) must hold as a system of partial differential equations on $\mathbb{R}^n_{++}$.

3. Define $r_i^j = E^i - E^j$. Define $\psi^j(x) = \theta^j(x + r_1^j, ..., x + r_n^j)$. Note that $(\psi^j)'(E^j) = (\nabla\theta^j(E) \cdot \mathbf{1})$. Hence, on the manifold $\mathfrak{M}_E^j = \{(x + r_1^j, ..., x + r_n^j)\}$, we may write the differential equation (A.18) as:

$$(\psi^j)'(x)(nx + r^{j,S}) + n\psi^j(x) = 1, \qquad \text{(A.19)}$$

where $r^{j,S} = \sum_i r_i^j$. Since the linear tax rules satisfy (A.18) by step 1, it follows that a particular solution of (A.19) is $\psi^j(x) = (1-t)\dfrac{x}{nx + r^{j,S}} + \dfrac{t}{n}$, for any $t \in [0,1]$. The general solution to the homogeneous variant of (A.19) is $\psi^j(x) = \dfrac{k^j}{nx + r^{j,S}}$, where k^j is a constant that may depend upon the manifold $\mathfrak{M}_E^j$. Therefore the general solution to (A.19) is:

$$\psi^j(x) = (1-t)\frac{x}{nx + r^{j,S}} + \frac{t}{n} + \frac{k^j}{nx + r^{j,S}},$$

where t may be chosen freely, and k^j is as described. Translating back, this means that

$$\theta^j(E) = \theta^j_{[t]}(E) + \frac{k^j(E)}{E^S}$$

where we must have:

(i) for all E, $\sum k^j(E) = 0$

(ii) $\theta^j(E) \in [0,1]$

(iii) for all j and E, $\nabla k^j(E) \cdot \mathbf{1} = 0$.

Statements (i) and (ii) are obvious requirements, while statement (iii) says that the functions k^j are constant on the manifolds $\mathfrak{M}_E^j$. ■

Proof of Theorem 7:

Part A

1. Define the functions:

$$r^j(K, y) = \max_r u^j\left(\frac{G(K + y + nr)}{n}, y + r\right) \text{ for } (K, y) \in \mathbb{R}_+^2.$$

These are single-valued functions, by strict concavity of *u*.

The first-order condition defining r^j is:

$$u_1^j(\cdot)G'(K+y+nr)+u_2^j(\cdot)=0.$$

2. Using the implicit function theorem, compute that the derivatives of r^j w.r.t. its arguments are :

$$\frac{dr}{dK}=-\frac{u_1^jG''+u_{11}^jG'^2+u_{12}^jG'}{n(u_1^jG''+u_{11}^j(G')^2+2G'u_{12}^j+u_{22}^j)}<0 .$$

The denominator of this fraction is negative by concavity of *u* and *G*, the the numerator is negative since $u_{12}^j\le 0$, and hence $\frac{dr}{dK}<0$. And:

$$\frac{dr}{dy}=-\frac{u_{22}^j+(G')^2u_{11}^j+(n+1)u_{12}^j+u_1^jG''}{n(u_{22}^j+(G')^2u_{11}^j+2G'u_{12}^j+u_1^jG'')}<0.$$

Likewise, $\frac{dr}{dy}<0$.

3. Define y^j by $r^j(0,y^j)=0$. If all agents other than *j* are putting in zero effort, then y^j is the amount of effort for *j* at which he would not like to increase all efforts by any number. Now define $K^{-j}=\sum_{i\neq j}y^j$. Next define z^j by $r^j(K^{-j},z^j)=0$. z^j is the amount of effort for *j* such that, if all other agents *i* are expending y^i and he is expending z^j, he would not like to add or subtract any amount from all efforts.

4. We argue that $z^j<y^j$ for all *j*. Just note that $r^j(K^{-j},z^j)=0=r^j(0,y^j)$. Since $K^{-j}>0$, it follows that $z^j<y^j$, because the r^j are decreasing functions.

5. Hence we may define the non-degenerate rectangle $\Delta=\{E\in\mathbb{R}^n_{++}\mid z\le E\le y\}$.

6. By applying the definition of $r^j(K,y)$, note that we have the identity:

$$r^j(K+(n-1)b,a+b)=r^j(K,a)-b .$$

7. We now define a function $\Theta:\mathbb{R}^n_+ \to \mathbb{R}^n_+$:

$$\Theta(E^1,...,E^n) = (E^1 + r^1(\hat{E}^{-1}, E^1),..., E^n + r^n(\hat{E}^{-n}, E^n))$$

where $\hat{E}^{-j} \equiv \sum_{i \neq j} E^i$. Θ is like the best-reply correspondence in Nash equilibrium.

Θ is single-valued and continuous, by the Berge maximum theorem.

We next show that $\Theta(\Delta) \subseteq \Delta$. Let $E = (E^1,...,E^n) \in \Delta$. We must show:

$$(\forall j)(z^j \leq E^j + r^j(\hat{E}^{-j}, E^j) \leq y^j. \qquad \text{(A.20)}$$

By step 6, we have

$$r^j(\hat{E}^{-j}, E^j) - (y^j - E^j) = r^j(\hat{E}^{-j} + (n-1)(y^j - E^j), y^j) \leq 0,$$

where the inequality follows because r^j is decreasing and $r^j(0, y^j) = 0$ and $\hat{E}^{-j} + (n-1)(y^j - E^j) \geq 0$. This proves the second inequality in (A. 20).

Again by step 6, we have:

$$r^j(\hat{E}^{-j}, E^j) - (z^j - E^j) = r^j(\hat{E}^{-j} + (n-1)(z^j - E^j), z^j) \geq 0$$

where the inequality follows because r^j is decreasing and $\hat{E}^{-j} + (n-1)(z^j - E^j) \leq K^{-j}$ (note that $(n-1)(z^j - E^j) \leq 0$). This proves the first inequality in (A.20).

8. Hence, the function Θ satisfies all the premises of Brouwer's Fixed Point Theorem, and hence possesses a fixed point. But a fixed point of Θ is a vector E such that for all *j*, $r^j(\hat{E}^{-j}, E^j) = 0$, which is precisely a K^+ equilibrium. (Note that the rectangle is in the strictly positive orthant, which implies that the equilibrium is strictly positive.)

Part B

9. The proof proceeds in the same fashion as above, except we now define the functions:

$$r^j_\beta(K, y) = \arg\max_r u^j\left(\frac{ry + \beta(r-1) + \beta}{r(K+y) + n\beta(r-1) + n\beta} G(r(K+y) + n(r-1)\beta, ry + \beta(r-1)\right).$$

Recall that *y* will be evaluated at E^j and K at $\hat{E}^{-j}$ for a vector *E*.

The first-order condition defining the functions r^j_β is:

$$u^j_1 \cdot G' + u^j_2 = 0,$$

where u is evaluated at the point $(\frac{y+\beta}{K+y+n\beta}G(r(K+y)+(r-1)n\beta), ry+(r-1)\beta)$. We compute, using the implicit function theorem, that:

$$\frac{dr_\beta^j}{dK}=-\frac{(G'u_{11}^j+u_{12}^j)\frac{y+\beta}{K+y+n\beta}\left(G'r_\beta^j-\frac{G}{K+y+n\beta}\right)+r_\beta^j G''u_1^j}{(y+\beta)(G'^2u_{11}^j+2G'u_{12}^j+u_{22}^j)+u_1^jG''(K+y+n\beta)}.$$

The denominator is negative by the concavity of u and G. Quasi-linearity implies that $G'u_{11}^j+u_{12}^j=0$ and so the numerator is negative if $r_\beta^j>0$. But note that we must have $ry+(r-1)\beta\geq 0$, since efforts cannot be negative, and so r is restricted to the interval with lower bound $r\geq\frac{\beta}{y+\beta}>0$, and so $r_\beta^j>0$. Hence $\frac{dr_\beta^j}{dK}<0$.

Compute that:

$$\frac{dr_\beta^j}{dy}=-\frac{u_{11}^j\left(r_\beta^jG'^2\frac{(y+\beta)}{K+y+n\beta}+G'G\frac{(K+(n-1)\beta)}{(K+y+n\beta)^2}\right)+u_{12}^j\left(r_\beta^jG'\left(\frac{K+2y+(n+1)\beta}{K+y+n\beta}\right)+\frac{K+y+(n-1)\beta}{(K+y+n\beta)^2}\right)+r_\beta^ju_{22}^j}{(y+\beta)(G'^2u_{11}^j+2G'u_{12}^j+u_{22}^j+u_1G''(K+y+n\beta)}$$

The denominator is negative by concavity, and the numerator is negative since $u_{12}^j=0$, and so $\frac{dr_\beta^j}{dy}<0$.

10. Hence the functions r_β^j are decreasing, and the proof proceeds as before, from steps 3 through 8. ■

<u>Proof of Proposition 2:</u>

The proof proceeds by showing that the mapping Θ is a contraction mapping. It uses the following well-known result:

<u>Lemma</u> Let $\|\ \|$ be a norm on $\mathbb{R}^n$ and let $[\![A]\!]$ be the associated sup norm on mappings $A:\mathbb{R}^n\to\mathbb{R}^n$, defined by $[\![A]\!]=\sup_{\|x\|=1}\|A(x)\|$. Let $J(A)$ be the Jacobian matrix of A. If $[\![J(A)]\!]<1$, then A is a contraction mapping.

If we can show that Θ is a contraction mapping, then it possesses a unique fixed point, and the dynamic process induced by iterating the application of Θ from any initial effort vector will converge to the fixed point.

1. For $n=2$, the Jacobian of the map Θ is $\begin{pmatrix} 1+r_1^1 & r_2^1 \\ r_1^2 & 1+r_2^2 \end{pmatrix}$, where

$r_i^j(E^1,E^2)=\dfrac{\partial r^j}{\partial E^i}(E^1,E^2)$, assuming that these derivatives exist. Thus, the lemma requires that we show the norm of this matrix is less than unity. We take $\|\ \|$ to be the Euclidean norm on $\mathbb{R}^2$. We must show that:

$$\|E\|=1 \Rightarrow \left\| \begin{pmatrix} 1+r_1^1(E) & r_2^1(E) \\ r_1^2(E) & 1+r_2^2(E) \end{pmatrix} \begin{pmatrix} E^1 \\ E^2 \end{pmatrix} \right\| < 1 . \qquad \text{(A.21)}$$

2. Assuming differentiability of c^j, the function $r^j(E)$ is defined by the following first-order condition:

$$G'(E^S + 2r^j(E)) = (c^j)'(E^j + r^j(E)) , \qquad \text{(A.22)}$$

which has a unique solution under standard assumptions. By the implicit function theorem, the derivatives of $r^j(\cdot)$ are given by:

$$G''(y^j)(1+2r_i^j(E)) = (c^j)''(x^j)(\delta_i^j + r_i^j(E)) ,$$

where $y^j = G(E^S + nr^j(E)),\ x^j = E^j + r^j(E)$ and $\delta_i^j = \begin{cases} 1, \text{ if } i=j \\ 0, \text{ if } i \neq j \end{cases}$; or

$$r_i^j(E) = \frac{\delta_i^j (c^j)''(x^j) - G''(y^j)}{2G''(y^j) - (c^j)''(x^j)} . \qquad \text{(A.23)}$$

3. It follows from step 1 that the Jacobian of Θ is given by:

$$\begin{pmatrix} \dfrac{G''(y^1)}{2G''(y^1)-(c^1)''(x^1)} & \dfrac{-G''(y^1)}{2G''(y^1)-(c^1)''(x^1)} \\ \dfrac{-G''(y^2)}{2G''(y^2)-(c^2)''(x^2)} & \dfrac{G''(y^2)}{2G''(y^2)-(c^2)''(x^2)} \end{pmatrix}$$

and so, from step 1, we need only show that:

$$(Q^1(E^1-E^2))^2+(Q^2(E^1-E^2))^2<1 \qquad \text{(A. 24)}$$

where $\left\|(E^1,E^2)\right\|=1$ and $Q^j=\dfrac{G''(y^j)}{2G''(y^j)-(c^j)''(x^j)}$. Note that $\left|Q^j\right|<\dfrac{1}{2}$. Therefore (A.24) reduces to showing that $\dfrac{1}{2}(1-E^1E^2)<1$, which is obviously true, proving the proposition. ■