RATIONAL TAX SCHEME BEHIND A VEIL OF IGNORANCE

BIUNG-GHI JU AND JUAN D. MORENO-TERNERO

ABSTRACT. We explore the design of impartial tax schemes when agents’ incomes are completely determined by their inborn talents. Building on Harsanyi’s veil-of-ignorance approach, we conceptualize an impartial observer who chooses a tax scheme without knowing her own vNM utility function, and the distribution of talents, and whose vNM preferences behind the veil obey Harsanyi’s principle of acceptance and are independent of the distribution of talents. Our results in the resulting framework provide three main messages: (i) the veil of ignorance implies anonymity of tax schemes; (ii) the veil of ignorance generically rejects utilitarian tax schemes; (iii) the veil of ignorance endorses the (Rawlsian) leveling tax scheme.

JEL: D63, D71, H24.
Keywords: veil of ignorance, impartiality, anonymity, leveling tax, utilitarianism.

1. INTRODUCTION

Income tax schemes, when properly designed, are essential tools for the enhancement of distributive justice. Our aim in this paper is to design an “impartial” income tax scheme for a group of agents when their incomes are completely determined by their inborn talents, which are “morally arbitrary” (e.g., Rawls, 1971). We take Harsanyi’s veil-of-ignorance approach (e.g., Harsanyi, 1953; 1955; 1977) to conceptualize an impartial viewpoint. An impartial observer, briefly IO, contemplates becoming, with equal probability, one of the agents in the group, for whom the tax burden is going to be allocated, and obtaining one of the possible taxable incomes (or talents) under an income distribution with equal probability too. The decision by such an IO is considered as being impartial because it would not represent the interests of a particular person only, or a particular income group only. The veil of ignorance necessitates the IO to reflect the interests of all agents (and all income groups) equally well.

The “basic structure of society” (the subject of justice) here is how the society should “determine the division of advantages from social cooperation”
RATIONAL TAX SCHEME BEHIND A VEIL OF IGNORANCE 2

(Rawls, 1999, p.6) in the form of tax allocations. The principles of justice govern the way taxes are allocated across different income groups.¹ We investigate the principles, or tax schemes, over which agents agree when they make their decisions rationally through a “fair procedure” (a unanimous decision behind the veil of ignorance). In order to “nullify the effects of specific contingencies which put men at odds and temp them to exploit social and natural circumstances to their own advantage” (Rawls, 1999, p.118), the veil of ignorance hides each person’s “fortune in the distribution of natural assets and abilities... his aversion to risk” (Rawls, 1999, p.118). In our model, this takes the form of uncertainties with regard to both incomes (indicating talents) and risk preferences. The major difference with respect to the original approach by Harsanyi is the added fortune in the distribution of talents, and our focus on the basic structure of society that is not sensitive to risk preferences (tax schemes do not take into account risk preferences; they take into account ordinal preferences and income distributions).

Our main findings can be summarized in two parts. On the one hand, and contrary to the utilitarian connotation of Harsanyi’s approach, we show that the IO’s rational decision in our context rejects utilitarian tax schemes. This is due to a fundamental conflict with impartiality that utilitarian schemes exhibit. More precisely, we show that, under a mild condition, all utilitarian tax schemes violate a specific form of anonymity, which turns out to be a necessary condition for the optimality of IO’s decision.

On the other hand, we find that, within a broad domain of (not necessarily anonymous) tax schemes, the optimal scheme for the IO coincides with the most egalitarian tax scheme -the leveling tax- which is a natural formulation of the Rawlsian difference principle in our simple taxation model. This might be interpreted as rationale for the veil of ignorance as a tool for recommending what justice requires, with respect to the distribution of wealth. It has been recently argued against that view due to a systematic conflict between the veil of ignorance and a weak version of egalitarianism (e.g., Roemer, 2002; Moreno-Ternero and Roemer, 2008). We show in this paper that one can escape from that disturbing conflict, at least in our model, provided one is willing to thicken the veil in a natural way and to obey the fundamental constraints imposed in the basic structure of society.

These findings convey another important message from our work, which had not been highlighted in the sizable existing literature on the veil of ignorance; namely, that the veil of ignorance, when properly thickened, generically implies anonymity of allocation rules (tax schemes, in our context).

¹¹These principles primarily apply, as I have said, to the basic structure of society and govern the assignment of rights and duties and regulate the distribution of social and economic advantages.” (Rawls, 1999, p.53).
The rest of the paper is organized as follows. In Section 2, we present the model. The results are gathered in Section 3. Section 4 concludes with some remarks and further insights.

2. Model

There is a group $N$ of $n$ agents. For ease of notation, let $N \equiv \{1, \ldots, n\}$. Each agent $i \in N$ is born with a certain level of talent, which is the sole determinant of his pre-tax income. Thus, agents have no moral responsibility for their incomes. Throughout the paper, we use talent and pre-tax income interchangeably. For each agent $i \in N$, let $y_i$ denote $i$’s talent or pre-tax income. A fixed amount of revenue $R > 0$ needs to be collected through a tax scheme. The tax revenue is used to supply a pure public good that provides uniform benefits to all agents, and so the only distributional concern here is with regard to post-tax incomes or tax allocations. Assume that the total income is not short of collecting the revenue $R$, i.e., $\sum_{i \in N} y_i \geq R$. Seminal investigation on such tax problems can be found, for instance, in Young (1987, 1988). The profile of incomes $y = (y_1, \ldots, y_n)$ and the revenue $R$ constitute a tax problem. Let $\mathcal{D}$ be the collection of all tax problems we are interested in solving and call it the domain of tax problems. Throughout the paper, assume that for all tax problems $(y, R) \in \mathcal{D}$, there is an agent with zero income.

A tax allocation for a problem $(y, R) \in \mathcal{D}$ is a profile of tax payments $t = (t_1, \ldots, t_n)$ such that $\sum_{i=1}^{n} t_i = R$ and $0 \leq t_i \leq y_i$ for each $i \in N$ (thus, we rule out subsidies as a means of redistribution). We say that $t = (t_1, \ldots, t_n)$ is an interior tax profile if $0 < t_i < y_i$ for each $i \in N$. A tax scheme is a function $f$ associating with each problem in $\mathcal{D}$ a tax allocation. Well-known tax schemes are the head tax; which distributes the tax burden equally subject to no agent paying more than his income; the leveling tax, which equalizes post-tax income across agents subject to no agent being subsidized; and the flat tax, which equalizes tax rates across agents. Formally, under the head tax, $t_i = \min\{-1/\lambda, y_i\}$ for some $\lambda \in \mathbb{R}_-$ with $\sum_{i \in N} \min\{-1/\lambda, y_i\} = R$; under the leveling tax, $t_i = \max\{y_i - 1/\lambda, 0\}$ for some $\lambda \in \mathbb{R}_+$ with $\sum_{i \in N} \max\{y_i - 1/\lambda, 0\} = R$; under the flat tax, $t_i = \lambda y_i$ for some $\lambda \in [0, 1]$ with $\sum_{i \in N} \lambda y_i = R$. The flat tax always yields interior tax profiles whenever $\sum_{i \in N} y_i > R$. This is, however, not the case with the other two schemes.

---

2In other words, we assume that labor is perfectly inelastically supplied.

3O’Neill (1982) used earlier the same mathematical framework to analyze the problem of adjudicating conflicting claims. Readers are referred to Moulin (2002) or Thomson (2003) for extensive treatments of diverse problems (such as taxation, conflicting claims, bankruptcy, cost sharing, and surplus sharing) fitting this framework.
Note that, under the leveling tax, exempt agents end up with lower post-tax incomes than the other agents: for each $i, j, k$ with $t_i, t_j > 0$ and $t_k = 0$, $y_k \leq y_i - t_i = y_i - t_j$. In our simple model, income can be considered as an index of primary goods. Then, the leveling tax allows us to achieve the maximum possible index (post-tax income) of worst-off agents. Hence, the leveling tax here formalizes the Rawlsian difference principle.

Assume that each agent possesses Neumann-Morgenstern (vNM) preferences over wealth lotteries (probability distributions over wealth levels). Denote agent $i$’s vNM utility function on wealth by $u_i: \mathbb{R}_+ \to \mathbb{R}$. Throughout the paper, we assume that all agents are strictly risk averse, i.e., $u_i(\cdot)$ is strictly concave for each $i \in N$. A (weighted) utilitarian tax scheme selects tax profiles maximizing the weighted average of (post-tax income) utilities, i.e., $(t_i)_{i \in N}$ maximizes

$$\sum_{i=1}^{n} \alpha_i u_i(y_i - t_i),$$

where $(\alpha_i)_{i \in N} \in \mathbb{R}_+^N$ is the vector of weights on individual utilities.

**Anonymity.** In the literature of fair allocation, impartiality is formalized as various axioms of (social choice) rules. One of the most basic ones is anonymity, the requirement that the tax payment of each agent should not depend on who the tax payer is but solely on her income level and the income distribution. This will require that switching the positions of agents $i$ and $j$ should switch the tax payments of the two and there should be no change in the tax payments by everyone else. Formally, let $\Pi$ denote the set of bijections from the set of agents $N$ into itself. For each $\pi \in \Pi$, let $y_{\pi}$ denote the resulting income profile after permuting $N$ via $\pi$. A tax scheme $f(\cdot)$ is anonymous if $f_{\pi(i)}(y, R) = f_i(y_{\pi}, R)$, for each $(y, R) \in \mathcal{D}$, $\pi \in \Pi$ and $i \in N$. Thus, if $f$ is anonymous, we write, for ease of notation, $t_\pi \equiv f(y_{\pi}, R)$, when $t = f(y, R)$. If $R$ remains fixed, and no confusion is possible, we shall skip it while referring to a scheme.

Anonymity requires first that when there occurs a reshuffling of agents’ income positions, all those agents whose positions remain fixed should not experience any change in their tax burdens; that is, their tax burdens should be invariant after the reshuffling. Anonymity also requires that reshuffled agents’ tax burdens should also be reshuffled in the same order; tax burdens should be covariant with the reshuffling. Formally,

**Invariance-Anonymity:** For each $\pi, \pi' \in \Pi$ and $i \in N$ with $\pi(i) = \pi'(i)$, $f_i(y_{\pi}) = f_i(y_{\pi'}).$

\[\text{It is straightforward to show that, in our model, the leveling tax profile is the best one under the lexicographic extension of the Rawlsian maximin principle.}\]
Covariance-Anonymity: For each \( \pi, \pi' \in \Pi \) and \( h, k \in N \) with \( h \neq k \) and \( \pi(h) = \pi'(k) \), \( f_h(y_\pi) = f_k(y_{\pi'}) \).

It is evident that anonymity is equivalent to the combination of these two partial formulations of anonymity. The equivalence will remain intact even if covariance-anonymity is replaced with weak covariance-anonymity: for each \( \pi, \pi' \in \Pi \), if for each \( i \in N \) with \( \pi(i) = \pi'(i) \), \( f_i(y_\pi) = f_i(y_{\pi'}) \), then for each \( h, k \in N \) with \( h \neq k \) and \( \pi(h) = \pi'(k) \), \( f_h(y_\pi) = f_k(y_{\pi'}) \).

Most utilitarian tax schemes are not anonymous when agents have different utility functions. They violate both invariance-anonymity and covariance-anonymity. As we shall see later, violation of invariance-anonymity is actually quite widespread among utilitarian schemes.

In the contractarian tradition, impartiality of moral norms (tax schemes, here) comes out from the impartial nature of the original position of the social contract. The most rigorous and effective development in the 20th century of this issue is characterized by the “veil of ignorance” (Harsanyi 1953, 1955, 1977; Rawls 1963, 1971).

The Veil of Ignorance. The veil of ignorance conceptualizes an impartial viewpoint where tax profiles can be decided without any influence of morally arbitrary characteristics such as risk preferences or talent. A decision maker behind the veil of ignorance aims to select “optimal” tax schemes (i.e., schemes benefitting her most) within the domain \( \mathcal{D} \). As a necessary condition to deliver impartiality, the domain should be closed with respect to any name permutation, that is, if \( y \in \mathcal{D} \), then for each \( \pi \in \Pi \), \( y_\pi \in \mathcal{D} \). In addition, if two problems differ only on the names of the agents then they should occur with equal probability. Furthermore, the decision maker could be in the positions of agent \( i \) (endowed with \( u_i(\cdot) \)) and agent \( j \) (endowed with \( u_j(\cdot) \)) with equal probability too. Formally, we assume that the state space \( \mathcal{S} \) is \( N \times \mathcal{D} \) and that there is a probability distribution \( P: N \times \mathcal{D} \to [0,1] \) such that for each \( i, j \in N \), \( y \in \mathcal{D} \), and \( \pi \in \Pi \), \( P(i,y) = P(j,y_\pi) \). The decision maker behind the veil of ignorance, so defined, will be referred to as the impartial observer, briefly, \( IO \).

The Impartial Observer. In what follows, for the sake of simplicity, we fix a distribution of talent \( y \in \mathbb{R}_+^n \) and a revenue \( R \) and restrict the domain of tax problems so that it only comprises those problems that have the same talent distribution and revenue as the problem \( (y,R) \). That is, we assume \( \mathcal{D} \equiv \{(y_\pi,R) : \pi \in \Pi\} \). Thus, our state space can be simply written as \( \mathcal{S} \equiv N \times \Pi \), and each \((i,\pi) \in N \times \Pi \) will be representing the state in which

\(^5\)A somewhat related, albeit different, alternative is the so-called notion of probabilistic egalitarianism (e.g., Lerner, 1944, Sen, 1971) that we shall not treat here.
the IO is in the position of agent \( i \) and faces the problem \((y_\pi, R)\).\(^6\) Under this simplification, for each \((i, \pi) \in \mathcal{S}\),

\[
P(i, y_\pi) = \frac{1}{n} \times \frac{1}{n!}
\]

where \( n! \equiv n \times (n - 1) \times \cdots \times 2 \times 1 \). Thus, \( \sum_{\pi \in \Pi} P(i, y_\pi) = 1/n \) for each \( i \in N \).

An extended prospect for the IO specifies the realized state \((i, \pi) \in \mathcal{S}\) and the level of post-tax income, or wealth \( W \in \mathbb{R}_+ \). A lottery is a probability distribution over the set of all extended prospects, \( \mathcal{S} \times \mathbb{R}_+ \). A simple lottery is a profile of probability-income pairs at all states; that is, \( L \equiv (p_i, \pi, W_i)_{(i, \pi) \in \mathcal{S}} \), (a sure income at each state). With a slight abuse, we call a probability distribution over wealth a wealth lottery.

The IO’s preferences over lotteries follow the vNM axioms. Thus, there is a utility index function \( U : \mathcal{S} \times \mathbb{R}_+ \to \mathbb{R} \) for which the IO’s preferences over lotteries are represented by the expected value of \( U(\cdot) \). The decision problem of the IO behind the veil of ignorance is then to find the tax scheme that maximizes the expected value of the utility index function \( U \). In order to determine, at least partially, what the function \( U \) is, we follow Harsanyi’s first step and assume what he calls:

**The Principle of Acceptance:** For each \( i \in N \) and \( \pi \in \Pi \), \( U(i, \pi, \cdot) \) represents the same vNM preferences over wealth lotteries as \( u_i(\cdot) \) represents.

In other words, once the IO takes the position of person \( i \), the IO’s preferences over wealth lotteries coincide with \( i \)'s preferences. This axiom implies that the IO’s vNM preferences over wealth lotteries do not depend on the distribution of talents.

In addition, we require that the IO’s talent, as well as the distribution of talents, do not make any difference in the IO’s preferences over wealth either. That is, consumptions of any level of wealth with \( y_\pi \) and with \( y_{\pi'} \), for each \( \pi, \pi' \in \Pi \), are indifferent for the IO who is in the position of person \( i \). We shall refer to this axiom as talent irrelevance.

**Talent Irrelevance:** For each \( i \in N \) and all \( \pi, \pi' \in \Pi \), \( U(i, \pi, \cdot) = U(i, \pi', \cdot) \).

By the principle of acceptance, and the vNM theorem, it follows that, for each \((i, \pi) \in \mathcal{S}\), there exist \( a_{i\pi} > 0 \) and \( b_{i\pi} \) such that, for each \( W \in \mathbb{R}_+ \),

\[
U(i, \pi, W) = a_{i\pi} \cdot u_i(W) + b_{i\pi}.
\]

\(^6\)Note that any domain that is closed with respect to name permutations can be expressed as the union of our simple domains. Thus, our simplification, which is only intended for ease of exposition, will not imply any loss of generality. Each of our results would apply to the general framework.
RATIONAL TAX SCHEME BEHIND A VEIL OF IGNORANCE

Then, by talent irrelevance, for each \( i \), there exist \( a_i > 0 \) and \( b_i \) such that \( a_i \equiv a_{i\pi} \) and \( b_i \equiv b_{i\pi} \) for each \( \pi \in \Pi \). Hence, for each \( (i, \pi) \in \mathcal{S} \) and all \( W \in \mathbb{R}_+ \),

\[
U(i, \pi, W) = a_i \cdot u_i(W) + b_i.
\]

If a tax scheme \( f: \mathcal{D} \rightarrow \mathbb{R}^N \) is chosen, then the IO faces a simple lottery \( L_f \equiv (P(i, \pi), y_{\pi(i)} - f_i(y_{\pi}))_{(i, \pi) \in \mathcal{S}} \). Thus, \( f \) yields the following expected utility for the IO:

\[
U(L_f) = \sum_{i \in N} \sum_{\pi \in \Pi} \frac{1}{n \times n!} \left[ a_i u_i(y_{\pi(i)}) - f_i(y_{\pi}) \right].
\]

In summary, we have the following statement regarding the choice of the impartial observer in our setting.

**Lemma 1.** Under the principle of acceptance and talent irrelevance, the IO’s preferences over tax schemes \( f \) are determined by the expected utility of \( f \),

\[
\sum_{i \in N} \sum_{\pi \in \Pi} \frac{1}{n \times n!} [a_i u_i(y_{\pi(i)}) - f_i(y_{\pi})] + b_i]
\]

for some \( a_1, \ldots, a_n \geq 0 \) and \( b_1, \ldots, b_n \in \mathbb{R} \).

Rational Tax Scheme. The IO chooses a tax scheme that is optimal both behind the veil of ignorance and behind the veil with interim information of her own identity and talent. A tax scheme \( f \) is interim optimal if for any realization of the IO’s identity and talent, the IO, after being informed of her own identity and talent, has no incentive to revise the tax scheme. It is rational if it is interim optimal and is preferred by the IO to any other interim optimal tax schemes.

3. Results

3.1. The veil of ignorance implies anonymity of tax schemes. We start this section with a result that provides a necessary condition for optimal tax schemes.

**Proposition 1.** All rational tax schemes satisfy invariance-anonymity.

**Proof.** Let \( f: \mathcal{D} \rightarrow \mathbb{R}^N \) be an optimal scheme for the IO. Let \( i^*, j^* \in N \) and \( \tilde{\pi}, \tilde{\pi}' \in \Pi \) be such that \( \tilde{\pi}(i^*) = \tilde{\pi}'(i^*) = j^* \). Suppose, by contradiction, that \( f_i^*(y_{\pi}) \neq f_{j^*}(y_{\tilde{\pi}'}) \). For each \( h \in N \), let \( \bar{t}_h \equiv \sum_{\pi \in \Pi: \pi(i^*) = h' \frac{f_{\pi(i^*)}(y_{\pi})}{(n-1)!} \) be the mean of all tax payments by the person with \( h \)’s talent at all problems \( y_{\pi} \) with \( \pi(i^*) = j^* \). In particular, \( \bar{t}_{j^*} \equiv \sum_{\pi \in \Pi: \pi(i^*) = j^* \frac{f_{\pi(i^*)}(y_{\pi})}{(n-1)!} \) Note that for each
π ∈ Π with π(i∗) = j∗, (i∗(h))h∈N is a feasible tax profile at tax problem yπ. By Lemma 1, the IO’s expected utility from tax scheme f can be written as

\[ U(L_f) = \sum_{i \in N} a_i \sum_{j \in N} \frac{1}{n^2} \sum_{\pi \in \Pi : \pi(i) = j} \frac{1}{(n-1)!} u_i(y_j - f_i(y_\pi)) + \frac{1}{n} \sum_{i \in N} b_i. \]

Then, by strict risk aversion of agent i∗, i.e., strict concavity of u∗(·),

\[ u_i(y_j - \bar{t}_j) > \sum_{\pi \in \Pi : \pi(i) = j} \frac{1}{(n-1)!} u_i(y_j - f_i(y_\pi)). \]

Thus, by replacing f(yπ) with \(\bar{t}_\pi\) for each \(\pi \in \Pi\) with \(\pi(i) = j\∗\), and keeping the rest of the values of \(f(·)\), we can define another tax scheme that gives the IO with the position of \(i^\ast\) and income level \(y_j^\ast\), a higher expected utility than \(f(·)\) does, contradicting interim optimality.

The following lemma, whose straightforward proof we omit, provides a useful matrix representation of tax schemes satisfying invariance-anonymity.

**Lemma 2.** A tax scheme f satisfies invariance-anonymity if and only if it is represented by a tax matrix \(T \equiv (T_{ij})_{(i,j) \in N \times N}\) such that, for each \(\pi \in \Pi\), and \(i \in N\),

- \(0 \leq T_{i\pi(i)} \leq y_{\pi(i)}\);
- \(\sum_{i \in N} T_{i\pi(i)} = R\);
- \(T_{i\pi(i)} = y_{\pi(i)} - f_i(y_\pi)\).

In words, the third statement of the previous lemma says that the \((i, j)\)'s entry of the tax matrix yields i's tax payment when i is endowed with j's talent. It follows from there that if the tax scheme also satisfies (weak) covariance-anonymity, and hence it is fully anonymous, all row vectors of \(T\) are identical.

A straightforward consequence of Proposition 1 is that, if an optimal tax scheme satisfies covariance-anonymity then it is fully anonymous. Moreover, we show that when there is an agent with zero income, if a rule satisfies invariance-anonymity, it also satisfies covariance-anonymity and so it is fully anonymous.

**Proposition 2.** Assume that y is such that \(y_h = 0\), for some \(h \in N\). Then, invariance-anonymity implies covariance-anonymity.

**Proof.** Let \((y, R) \in \mathcal{D}\) and \(x \equiv f(y, R)\). Let \(h \in N\) be such that \(y_h = 0\). Let \(i, j \in N\), \(i \neq j\), and \(R' \equiv x_i + x_j\). Let \(\pi^ij: N \to N\) be the transposition of \(i\) and \(j\).
By invariance-anonymity, for all \( k \in N \backslash \{ i, h \} \), \( f_k(y_{\pi^{jh}}, R) = x_k \). By claims-boundedness, \( f_i(y_{\pi^{jh}}, R) = 0 = f_h(y, R) \). Hence by efficiency, \( f_h(y_{\pi^{jh}}, R) = x_i \). If \( j = h \), we are done.

Suppose \( j \neq h \). Now let \( y' \equiv y_{\pi^{jh}} \). Again by invariance-anonymity, for each \( k \in N \backslash \{ i, j \} \), \( f_k(y'_{\pi^{ij}}, R) = f_k(y', R) \), which means \( f_h(y'_{\pi^{ij}}, R) = x_i \) and for each \( k \in N \backslash \{ i, j, h \} \), \( f_k(y'_{\pi^{ij}}, R) = x_k \). By claims-boundedness, \( f_j(y'_{\pi^{ij}}, R) = 0 \). Hence by efficiency, \( f_j(y'_{\pi^{ij}}, R) = x_j \).

Finally let \( y^* \equiv y'_{\pi^{ij}} \). Clearly, \( y^*_{\pi^{jh}} = y_{\pi^{ij}} \). Again by invariance-anonymity, for each \( k \in N \backslash \{ j, h \} \), \( f_k(y^*_{\pi^{jh}}, R) = f_k(y^*, R) \), which means \( f_i(y^*_{\pi^{jh}}, R) = x_j \) and for each \( k \in N \backslash \{ i, j, h \} \), \( f_k(y^*_{\pi^{jh}}, R) = x_k \). By claims-boundedness, \( f_h(y^*_{\pi^{jh}}, R) = 0 \). Hence by efficiency, \( f_j(y^*_{\pi^{jh}}, R) = x_i \).

Therefore, \( f(y_{\pi^{ij}}, R) = x_{\pi^{ij}} \).

Now the result follows from the fact that all permutations \( \pi : N \to N \) can be represented by a composition of a finite sequence of transpositions. \( \square \)

The following result, which is a straightforward consequence of Propositions 1 and 2, conveys our first message in this paper, as stated in the title of this subsection; namely, that the veil of ignorance implies anonymity of tax schemes.

**Theorem 1.** If a tax scheme is optimal for the IO, it is anonymous.

3.2. **The veil of ignorance generically rejects utilitarian tax schemes.**

Proposition 1 shows that a rational decision behind the veil of ignorance implies invariance-anonymity. As mentioned earlier, utilitarian tax schemes violate anonymity quite often. In fact, violation of invariance-anonymity occurs under the following mild condition for tax schemes, which says that an agent’s income is lower than another’s agent post-tax income for some interior tax profile. Formally,

**Minimal Difference:** There exist \( \pi \in \Pi \) and \( y \in \mathcal{D} \) such that for each \( i \in N \), \( 0 < f_i(y_{\pi}) < y_{\pi(i)} \) and, for some \( j, k \in N \), \( y_{\pi(j)} < y_{\pi(k)} - f_k(y_{\pi}) \).

**Proposition 3.** Assume \( n \geq 3 \). If a utilitarian tax scheme satisfies minimal difference, then it violates invariance-anonymity.

**Proof.** Let \( f \) be a utilitarian tax scheme satisfying minimal difference. Let \( \pi \in \Pi \) and \( y \in \mathcal{D} \) be such that for each \( i \in N \), \( 0 < f_i(y_{\pi}) < y_{\pi(i)} \) and, for some \( j, k \in N \), \( y_{\pi(j)} < y_{\pi(k)} - f_k(y_{\pi}) \). Let \( \pi' \) be such that \( \pi'(k) = \pi(j) \), \( \pi'(j) = \pi(k) \), and for each \( h \neq k, j \), \( \pi'(h) = \pi(h) \). By the interior assumption, marginal utilities at \( \pi \)-post-tax incomes must be equal, i.e., \( MU_1(y_{\pi(1)} - f_1(y_{\pi})) = \cdots = MU_n(y_{\pi(n)} - f_n(y_{\pi})) \equiv \lambda \). As \( y_{\pi'(k)} = y_{\pi(j)} < y_{\pi(k)} - f_k(y_{\pi}) \).

**Claim:** For each \( h \neq k \), \( MU_h(y_{\pi'(h)} - f_h(y_{\pi'})) < \lambda \).
Proof. We first show that $MU_j(y_{\pi'}(j) - f_j(y_{\pi'})) < \lambda$. Suppose $MU_j(y_{\pi'}(j) - f_j(y_{\pi'})) \geq \lambda = MU_j(y_{\pi}(j) - f_j(y_{\pi}))$, which gives $y_{\pi'}(j) - f_j(y_{\pi'}) \leq y_{\pi}(j) - f_j(y_{\pi})$. As both $j$ and $k$ have lower post-tax wealth, there must be another agent $h \neq j, k$ whose post-tax wealth is higher, i.e., $y_{\pi'}(h) - f_h(y_{\pi'}) > y_{\pi}(h) - f_h(y_{\pi})$. Then, by the strict concavity of $u_h$, $MU_h(y_{\pi'}(h) - f_h(y_{\pi'})) < MU_h(y_{\pi}(h) - f_h(y_{\pi})) = \lambda$. Therefore $MU_h(y_{\pi'}(h) - f_h(y_{\pi'})) < MU_j(y_{\pi'}(j) - f_j(y_{\pi'}))$. Note that $y_{\pi'}(h) - f_h(y_{\pi'}) > 0$ and $y_{\pi'}(j) - f_j(y_{\pi'}) < y_{\pi'}(j) > 0$. Thus, it is possible to reduce $h$'s post-tax wealth and increase $j$'s post-tax wealth simultaneously, which leads to a higher aggregate utility level, contradicting that $f(y_{\pi'})$ maximizes the aggregate utility.

Now suppose that, for some $h \neq j, k$, $MU_h(y_{\pi'}(h) - f_h(y_{\pi'})) \geq \lambda$. Then $y_{\pi'}(h) - f_h(y_{\pi'}) \leq y_{\pi}(h) - f_h(y_{\pi})$. Note that $y_{\pi'}(j) - f_j(y_{\pi'}) > y_{\pi}(j) - f_j(y_{\pi}) > 0$ and that, by assumption, $y_{\pi'}(h) - f_h(y_{\pi'}) < y_{\pi}(h)$ and so $y_{\pi'}(h) - f_h(y_{\pi'}) < y_{\pi}(h)$. Thus it is possible to simultaneously increase the post-tax wealth of $h$ and reduce the post-tax wealth of $j$, which increases the aggregate utility (for $h$ has a higher marginal utility than $j$). $\diamond$

Hence, for each $h \neq j, k$, $MU_h(y_{\pi'}(h) - f_h(y_{\pi'})) < \lambda = MU_h(y_{\pi}(h) - f_h(y_{\pi}))$, which implies $y_{\pi'}(h) - f_h(y_{\pi'}) > y_{\pi}(h) - f_h(y_{\pi})$. Therefore, as $\pi'(h) = \pi(h)$, we get $f_h(y_{\pi'}) < f_h(y_{\pi})$, contrary to what invariance-anonymity requires. $\square$

Combining Propositions 1 and 3 we reach the conclusion that utilitarian tax schemes are typically rejected by the IO (the second message we wanted to convey from this work).

**Theorem 2.** If a utilitarian tax scheme satisfies minimal difference, then it is not optimal for the IO.

3.3. The veil of ignorance endorses the (Rawlsian) leveling tax scheme.

In this final subsection, we show that the most egalitarian tax scheme—the leveling tax—which is a natural formulation of the Rawlsian difference principle in our simple taxation model, is the optimal scheme for the IO. Formally,

**Theorem 3.** The leveling tax is the only optimal tax scheme for the IO.

**Proof.** By Theorem 1, an optimal tax scheme for the IO is anonymous. Then by Lemma 2, it is represented by a tax matrix and by anonymity, the matrix is composed of a constant row vector $t \equiv (t_i)_{i \in N}$. Consequently, the IO’s problem can be expressed as follows:

$$\max_t \left\{ \sum_{i=1}^{n} \sum_{j=1}^{n} a_i \cdot u_i(y_j - t_j) : 0 \leq t_j \leq y_j, \sum_{j \in N} t_j = R \right\}.
$$

(3.1)
In the remaining part of the proof, we assume, for simplicity, that individual utility index functions \( u_i(\cdot) \) are differentiable. The Lagrangian associated with (3.1) is given by

\[
\mathcal{L}(\cdot) = \sum_{j=1}^{n} \sum_{i=1}^{n} a_i \cdot u_i(y_j - t_j) + \lambda \left( R - \sum_{j=1}^{n} t_j \right) + \sum_{j=1}^{n} \mu_j(y_j - t_j) + \sum_{j=1}^{n} \gamma_j t_j.
\]

As all agents are strictly risk averse, \( \sum_{i,j=1}^{n} a_i \cdot u_i(\cdot) \) is strictly concave, and the above program can be solved by the Kuhn-Tucker Theorem (e.g., Mas-Collell et al., 1995). We consider two types of solutions to (3.1) for which the first constraint binds, i.e., solutions \( t \) such that \( \sum_{j=1}^{n} t_j = R \). The first type refers to interior solutions (i.e., those for which \( 0 \leq t_j \leq y_j \) for each \( j \in N \)) and the second one to corner solutions (i.e., those for which either \( t_j = 0 \) or \( t_j = y_j \) for some \( j \in N \)).

Formally,

**Case 1.** \( \mu_j = \gamma_j = 0 \) for each \( j \in N \) (the interior solutions).

In this case, we would have to solve the following system of equations:

\[
- \sum_{i=1}^{n} a_i \cdot u_i'(y_j - t_j) = \lambda \quad \text{for all} \quad j \in N.
\]

Let \( j, k \in N \). Then,

\[
\sum_{i=1}^{n} a_i \cdot u_i'(y_j - t_j) = \sum_{i=1}^{n} a_i \cdot u_i'(y_k - t_k).
\]

As \( \sum_{i,j} a_i u_i(\cdot) \) is strictly concave, it follows that

\[
y_j - t_j = y_k - t_k.
\]

In other words, post-tax incomes are equalized across agents.

**Case 2.** \( \mu_j \neq 0 \) or \( \gamma_j \neq 0 \) for some \( j \in N \) (the corner solutions).

In this case, we would have to solve the following system of equations:

\[
- \sum_{i=1}^{n} a_i \cdot u_i'(y_j - t_j) - \lambda - \mu_j + \gamma_j = 0 \quad \text{for all} \quad j \in N.
\]

Thus, it is clear that for each those agents \( j \in N \) with interior solutions, i.e., \( \mu_j = \gamma_j = 0 \), then we have \( - \sum_{i=1}^{n} a_i \cdot u_i'(y_j - t_j) = \lambda \). Then Case 1 concludes that these agents will have equal post-tax incomes. Let now \( j \in N \) be such that \( \mu_j > 0 \). Then, \( t_j = y_j \), which implies that \( 0 \geq - \sum_{i=1}^{n} a_i \cdot u_i'(0) = \lambda + \mu_j > 0 \), a contradiction. Finally, let \( j \in N \) be such that \( \gamma_j \neq 0 \). Then, \( t_j = 0 \) and, therefore, \( - \sum_{i=1}^{n} a_i \cdot u_i'(y_j) = \lambda - \gamma_j \). As \( - \sum_{i=1}^{n} a_i \cdot u_i'(y_k - t_k) = \lambda \), for each \( k \) such that \( t_k > 0 \), it follows that \( y_j = y_j - t_j \leq y_k - t_k \) for each \( j, k \) such that \( t_k > 0 = t_j \). □
A close examination of the above proof tells us that it is not necessary to assume that all agents are "strictly" risk averse. The result would also hold assuming that only one agent is "strictly" risk averse (as this would be enough to guarantee the strict concavity of \( \sum_{i \in N} a_i u_i(\cdot) \)). "Strict" concavities of all individual utility functions, however, is needed to obtain invariance-anonymity in the proof of Proposition 1.

4. CONCLUDING REMARKS

The veil of ignorance has been an influential concept in political philosophy during the last half century. Many prominent authors have employed it in different forms as a tool to guarantee impartiality in resource allocation, albeit with no clear consensus about how thick the veil should be (e.g., Harsanyi, 1953; 1955; 1977; Rawls, 1971; Dworkin, 1981a; 1981b). One of the early contributions on the veil of ignorance is due to John Harsanyi, which we have partially endorsed here. In Harsanyi’s original framework, an impartial observer engages in a thought experiment in which he imagines having an equal chance of being any individual, complete with that individual’s tastes and objective circumstances. Harsanyi obtains that lotteries on the set of social alternatives should be ranked according to the average of the individual utilities, for suitably chosen vNM utility representations of the individual preferences.\(^7\) There is a broad agreement about adopting Harsanyi’s original framework and his principle of acceptance. Nevertheless, beyond such agreement on his first step, there are different options in the literature to complete it.\(^8\) Some of these options (e.g., Karni, 1998) arise from challenging Harsanyi’s formalization of impartiality, or the interpretation of his result as a justification for average utilitarianism.\(^9\)

In this paper, we have explored the selection of optimal and impartial tax schemes in a slight modification of Harsanyi’s original framework to impose a thicker veil of ignorance. In our model, agents’ incomes are completely determined by their talents. The IO contemplates becoming, with equal probability, one of the agents in the group for whom the tax burden is going to be allocated, and also obtaining one of the possible talents with equal probability. Following Harsanyi’s first step, we impose the principle of acceptance but, due to the thickness of our veil of ignorance, we also impose an additional axiom (talent irrelevance) with an inherent flavor of

\(^7\)Karni and Weymark (1993) provide an extension of this theorem to an informationally parsimonious context in which the IO is only assumed to have preferences on the extended lotteries in which there is an equal chance of being any person in society.

\(^8\)See, for instance, Moreno-Ternero and Roemer (2008) and the literature cited therein.

\(^9\)See, for instance, Weymark (1991) for an elaborated critique against this interpretation.
impartiality, according to which the distribution of talents does not make any difference in the IO’s preferences. We then scrutinize the role of impartiality by means of several anonymity axioms and obtain three main results. The first result amounts to obtain anonymity of tax schemes as a byproduct of the veil-of-ignorance approach. The second result amounts to the IO’s rejection of utilitarian tax schemes, which could be seen as an additional rationale against the interpretation of Harsanyi’s IO theorem as a justification for average utilitarianism. The third result amounts to the IO’s endorsement of the leveling tax. Thus, the IO is in line with giving priority to the worst-off.\textsuperscript{10} Actually, the IO selects the most extreme version of priority, as it advocates for the most preferred taxation scheme for the worst-off individual. In other words, we provide a veil-of-ignorance argument to endorse a Rawlsian recommendation in our context. This is in line with Rawls (1971) himself, who put forth an argument for the difference principle invoking the veil of ignorance and the original position.

REFERENCES


\textsuperscript{10}See Moreno-Ternero and Roemer (2006) for an instance of the powerful implications of this notion in resource allocation.

DEPARTMENT OF ECONOMICS, SEOUL NATIONAL UNIVERSITY, GWANAK-RO 1, GWANAK-GU, SEOUL 151-746, KOREA
E-mail address: bgju@snu.ac.kr

DEPARTMENT OF ECONOMICS, UNIVERSIDAD PABLO DE OLAVIDE, AND CORE, UNIVERSITÉ CATHOLIQUE DE LOUVAIN.
E-mail address: jdmoreno@upo.es