Probabilistic Lump-Sum Taxation

Maciej Konrad Dudek*

November 23, 2016

Abstract

In this paper, we describe an elementary, fully implementable, and revenue neutral randomizing mechanism that leads to a Pareto improvement over allocations induced by labor income tax functions. We illustrate, by providing a theoretical example, that our mechanism can be applied to optimal tax functions identified by Saez [17] and Diamond [4]. Furthermore, we provide an explicit numerical example, which confirms that our procedure can lead to a Pareto improvement over allocations induced by optimal tax functions of Mirrlees [12]. Moreover, we show that our randomizing procedure is not only implementable, as it requires less than the standard informational demands, but can also lead to sizable welfare gains amounting to about 10% of the underlying distortion. Finally, we show that, unlike most of the traditional approaches involving randomness, the mechanism we describe can preserve horizontal equity.

Key Words: Optimal income taxation, ex ante randomization, Pareto improvement, lump-sum taxes.
JEL CODES: H20, H21, D82, D63, D31.

1 Introduction

Routinely economics teachers present with enthusiasm the efficiency virtues of lump-sum taxation only to dismiss their practical relevance a few moments later. In this paper we constructively argue that our tendency to essentially ignore lump-sum taxation as a viable policy tool can be highly premature. In fact, we show in a framework with lesser than standard informational constraints that equilibrium outcomes obtained with traditional tax functions can be improved upon by introducing lump-sum taxation. In particular, we show that it is possible, by allowing for the stochastic presence of lump-sum taxation, to improve in the Pareto sense over allocations induced by optimal tax functions of Saez [17] and Diamond [4], Mirrlees [12] without compromising revenue.

In the context of taxation it is normally assumed that individual types are only privately known and consequently tax liabilities must be conditioned on observables such as income. Thus, it is typically assumed that tax liabilities are dictated by a function, $\tau(\,)$, with the actual tax payments being determined by income earned, y, and equal to $\tau(y)$. Unfortunately, this traditional approach invariably leads to welfare losses relative to the first best outcome. Accordingly, following Mirrlees [12] researchers have been searching for tax functions that ensure the highest possible welfare. While this traditional approach has led to profound insights and has generated practical recommendations - Diamond [4] and Saez [17] - it has by definition left us in the second best realm. In this paper, we argue that our choice to restrict our attention to the plasticity of $\tau(\,)$ alone is in fact self-imposed and, more importantly, it is unnecessary as it is possible, as we constructively show, to improve over allocations induced by traditional and possibly optimal tax functions. In fact, we argue that by allowing for a partial presence of lump-sum taxes we can move into second best *plus* world and induce allocations characterized by higher welfare than those induced by traditional tax functions.

Our basic mechanism is extremely simple. In fact, we rely on a specific form of ex-ante randomization to improve in the Pareto sense over the allocation induced by a given tax function, $\tau(\,)$. In our basic context economic agents face lotteries characterized by up-side risk only. Consequently, they can only gain by participating in our mechanism. Specifically, in our context agents choose a lottery from a menu of lotteries

*Please send comments to mkdudek@alumni.caltech.edu

offered by the government. Having selected a lottery - defined over just two tax functions - a given agent awaits uncertainty resolution and then having learned her true tax function decides on her effort, i.e., income earned. In our context, in the case of an unfavorable uncertainty resolution the agent faces the original tax function, $\tau()$. On the other hand when uncertainty is resolved favorably the agent faces a lump-sum tax equal to the amount that she would pay if she faced $\tau()$. Consequently, in our case a given rational agent, in essence, by selecting a given lottery chooses its actual tax liability. In other words, in our framework, in equilibrium, tax liabilities are not random, but naturally marginal tax rates are. In the paper, we construct the relevant menu of lotteries, we explicitly present the relevant probabilities and show that our basic mechanism exists.

The notion of randomness and its potentially improving role in the context of taxation was originally noted by Stiglitz [21] who observed, in particular, that ex-ante randomization can help to convexify the utility possibility frontier and, thus, can be welfare improving. Numerous contributions have followed since. Specifically, Brito et al. [2] show, working with a space of discrete types, that ex-ante randomization can be welfare improving in the context of optimal tax functions. Moreover, Pestieau et al. [23] and more recently and Gauthier and Larroque [9] identify potential gains from randomization in the context of taxation and redistribution. Furthermore, recently in different, but formally related to ours, contexts Baisa [1], and Garratt and Pycia [8] have overturned long-standing profound results by allowing for randomization. Finally, our approach is not the only attempt to improve over allocations induced by traditional tax functions given the original results obtained by Dudek [6] and later popularized by Del Negro et al. [3].

Hellwig [10], in his general contribution, argued that traditional randomization mechanisms in the context of optimal taxation cannot be beneficial as long as the utility function representing the preferences satisfies a technical condition known as weakly decreasing consumption risk aversion. We do not dispute the funding of Hellwig. Accordingly, whenever we discuss optimal taxation we assume that the relevant utility function does not satisfy the condition identified by Hellwig. It is worth noting that the condition of Hellwig is highly intuitive. Still, it is not universal and, in particular, is not satisfied by some highly popular utility functions like that of Saez [17]. Consequently, our procedure can, as we illustrate in this paper, be potentially employed in some highly relevant cases in the context of optimal taxation.

We differ from prior efforts involving ex-ante randomization in a number of critical aspects. First of all, we derive our results in the context of a continuum of types, which makes our results applicable and in fact ready for practical implementation. Secondly, in most contributions researchers limit their attention to local analyses, which are, by definition, limited in scope. Here, perhaps surprisingly, we are able to obtain stronger results by allowing for significantly higher variability in outcomes. In addition, we show that in our framework it is possible to ensure ex-post horizontal equity in the strong form, despite the fact that we allow for randomness in our model.

The procedure outlined in this paper is not only promising because of its potential applicability to optimal taxation functions, but also due to its practical relevance. In particular, we show that the welfare gains generated by our procedure are sizable and can amount to about 10% of the underlying average distortion. Furthermore, our procedure is particularly valuable as its implementation not only can potentially lead to outcomes superior to those induced by the optimal tax functions of Mirrlees, but also it hinges on less demanding informational constraints as the knowledge of the skill distribution is not required in our context. Moreover, the optimal tax functions of Mirrlees are highly sensitive to the shape of the skill distribution. Consequently, practical tax reforms in line with the prescriptions of Mirrlees are risky and very difficult to implement. Naturally, our procedure, as it does not require the knowledge of the distribution of skills, is less risky and, as such, ready for practical implementation.

The paper is organized as follows. In the next section, we outline the basic model. Feasibility of improvement over optimal tax functions of Saez [4] and Diamond [17] is explicitly illustrated in section three. In section four we present a modification that allows to preserve horizontal equity ex-post in the strong form. In the following section we provide estimates of welfare gains stemming from our procedure. A numerical example of applicability of the procedure to optimal tax functions of Mirrlees [12] is presented in section six. Section seven contains conclusions.

2 Basic Model

We start by illustrating our basic findings in a simple, but highly suggestive framework. Specifically, we assume - following Saez [17] - that individual preferences are represented with the following utility function

$$U(C,L) = h(C - \frac{1}{2}L^2), \tag{1}$$

where $h(\,)$ is an increasing and concave function[1], i.e., $h'(\,) > 0$ and $h''(\,) < 0$.

Naturally, we assume that individuals differ with respect to productivity. Individual productivities are distributed on interval $[a_L, a_H]$ with $f()$ being the corresponding *pdf*. An agent with productivity $a \in [a_L, a_H]$, which is not publicly observable, delivers output $y = aL$ if she chooses to supply L units of labor.

Let us now consider tax function $\tau(\,)$. For the time being and only for illustrative purposes, we consider function $\tau(\,)$ to be generic and not optimal in any sense[2]. Now, let y_a denote the level of income earned by an agent with productivity $a \in [a_L, a_H]$ when she faces $\tau(\,)$. Naturally, typically[3] y_a satisfies

$$1 - \tau^{'}(y_a) = \frac{y_a}{a^2}. \tag{2}$$

Finally, let us note that the revenue collected from an agent with productivity a in this case is simply given by $R_a = \tau(y_a)$.

We can describe the allocation that materializes when agents face tax function $\tau(\,)$ with

$$\mathcal{A} = \{h(y_a - \tau(y_a) - \frac{1}{2}(\frac{y_a}{a})^2)|\, a \in [a_L, a_H],\, P(s < a) = \int_{a_L}^{a} f(s)ds\}. \tag{3}$$

Let us now consider a particularly simple modification, which appears to have been overlooked, that allows for a Pareto improvement over allocation $\mathcal{A}$ described with relationship (3). The mechanism relies on a very simple form of ex-ante randomization, which involves only up-side risk with a novel property that the actual tax liabilities are known with certainty. Specifically, we can consider a lottery of the form

$$B_a = \begin{cases} \tau(y) & \text{with probability } \; p(a) \\ \tau(y_a) & \text{with probability } 1 - p(a). \end{cases} \tag{4}$$

Naturally, the above lottery just assigns $\tau(y)$ to a given agent with probability $p(a)$ and tax function $\tau(y_a)$ with probability $1-p(a)$. Note that if a given agent faces lottery B_a then in the case of an unfavorable uncertainty resolution the agent is assigned the original tax function, $\tau(y)$, and in the case of a favorable uncertainty resolution the agent faces a more *favorable* tax function, $\tau(y_a)$, which is flat, i.e., is characterized by a marginal tax of 0.

Let us take a step further and consider a collection of lotteries of the form $\mathcal{B} = \{B_a \,|\, a \in [a_L, a_H]\}$, where of course B_a is described with relationship (4). Furthermore, let us suggest the following sequence of events.

1. The government presents the collection of independent lotteries, $\mathcal{B}$.
2. Agents choose lotteries from the collection of lotteries, $\mathcal{B}$.
3. For each individually chosen lottery the uncertainty is resolved, and each agent is assigned a tax function.
4. Having learned her own tax function each agent decides on her labor supply, i.e., on the level of income earned.

[1] In Saez's specification $h()$ takes the form of the logarithmic function. We show that our results hold for any utility function in Appendix A.

[2] In the subsequent section we illustrate that the procedure can be applied when function $\tau(\,)$ is optimal in the Saez [17] sense.

[3] Occasionally, see Ebert [7], Lollivier and Rochet [11], the condition below is not sufficient.

5. Tax liabilities are paid in line with individually drawn - in Step 3 - tax functions.

Imagine that upon observing the set of lotteries, $\mathcal{B}$, presented by the government an agent with productivity $a \in [a_L, a_H]$ chooses lottery B_b for some $b \in [a_L, a_H]$. Given the choice of lottery B_b there are two possible outcomes. The chosen lottery can be either *lost* or *won*.

If the lottery is lost, which happens with probability $p(b)$, then the original tax function, $\tau(y)$ is presented as the tax function to the agent. Note that by invoking a simple revealed preference argument we can state that in this case the agent having observed the tax function simply chooses to earn y_a and consequently to pay $R_a = \tau(y_a)$ in taxes. The agent's realized utility in this case is given by $U^L_{a,b} = h(y_a - \tau(y_a) - \frac{1}{2}(\frac{y_a}{a})^2)$. Observe that in this case, of unfavorable uncertainty resolution, the agent attains the level of utility that she would have attained had she not participated in the mechanism at all.

If the lottery is *won*, which happens with probability $1 - p(b)$, then tax function $\tau(y_b)$, which is flat, is presented to the agent. Upon observing the tax function in this case the agent chooses to earn[4] $y^*_a = a^2$ and naturally she pays $\tau(y_b)$ in taxes. The agent's realized utility in this case is given by $U^W_{a,b} = h(y^*_a - \tau(y_b) - \frac{1}{2}(\frac{y^*_a}{a})^2)$.

Summarizing we can state that the expected utility of an agent with productivity a upon choosing lottery B_b for $b \in [a_L, a_H]$ is given by

$$EU_{a,b} = p(b)h(y_a - \tau(y_a) - \frac{1}{2}(\frac{y_a}{a})^2) + (1 - p(b))h(y^*_a - \tau(y_b) - \frac{1}{2}(\frac{y^*_a}{a})^2). \tag{5}$$

Presumably, the agent, when presented the above collection of lotteries, is interested in selecting a lottery that maximizes her expected utility. Her optimal choice - assuming sufficient regularity conditions - is defined by $\frac{\partial EU_{a,b}}{\partial b} = 0$, i.e., it is implicitly described with

$$\begin{array}{lr} p'(b)h(y_a - \tau(y_a) - \frac{1}{2}(\frac{y_a}{a})^2) - p'(b)h(y^*_a - \tau(y_b) - \frac{1}{2}(\frac{y^*_a}{a})^2) & + \\ +(1 - p(b))h^{'}(y^*_a - \tau(y_b) - \frac{1}{2}(\frac{y^*_a}{a})^2)(-\tau'(y_b))\frac{\partial y_b}{\partial b} & = 0. \end{array} \tag{6}$$

Relationship (6) defines implicitly the choice, b, of agent a. Note that up to this stage probability function $p(\,)$ has been treated as a parameter. We can, however, try to adjust $p(\,)$ in such a way as to induce agent a to select a lottery that is *intended* for her. In particular, we can try to ensure that it is the case that an agent with productivity a out of all lotteries B_b prefers lottery B_a the most. This occurs, when $\frac{\partial EU_{a,b}}{\partial b}|_{b=a} = 0$, i.e., after rearranging condition (6) with $b = a$, when

$$\frac{-p'(a)}{1 - p(a)} = \frac{h^{'}(y^*_a - \tau(y_a) - \frac{1}{2}(\frac{y^*_a}{a})^2)}{h(y^*_a - \tau(y_a) - \frac{1}{2}(\frac{y^*_a}{a})^2) - h(y_a - \tau(y_a) - \frac{1}{2}(\frac{y_a}{a})^2)}\tau'(y_a)\frac{\partial y_a}{\partial a}, \tag{7}$$

which upon integration becomes

$$log(1 - p(a)) = log(1 - p(n)) + \int_n^a m(s)ds, \tag{8}$$

where $m(\,)$ represents the right-hand side of equation (7) and n is a constant.

Constant n can be determined by requiring that the individual with the highest ability be, in line with the results of Saede [19] and Sadka [16], always presented the winning lottery, i.e., that $p(a_H) = 0$. Therefore, we can finally write $p(a)$ as

$$p(a) = 1 - e^{-\int_a^{a_H} m(s)ds}. \tag{9}$$

Recall that it is always beneficial to pay a given amount of taxes in a lump sum form rather than facing distortive non-zero marginal taxes. Therefore, we can write $h(y^*_a - \tau(y_a) - \frac{1}{2}(\frac{y^*_a}{a})^2) - h(y_a - \tau(y_a) - \frac{1}{2}(\frac{y_a}{a})^2) \geq 0$. Furthermore, given the assumed properties of $h(\,)$ we can state that $h^{'}(y^*_a - \tau(y_a) - \frac{1}{2}(\frac{y^*_a}{a})^2) > 0$. In addition, note that condition (2) implies that $\frac{\partial y_a}{\partial a} = \frac{2a(1-\tau'(y_a))}{1+a^2\tau''(y_a)}$. Moreover, we choose to deal only with

[4] Recall that at this stage we operate in a framework with no income effects. The general case is handled in Appendix A.

tax functions characterized by marginal tax rates between[5] 0 and 1, thus, $2a\tau'(y_a)(1-\tau'(y_a)) \geq 0$. Finally, the second order condition corresponding to condition (2) implies that $1 + a^2\tau''(y_a) \geq 0$. Therefore, we can state that function $m(\,)$ assumes only non-negative values, i.e., $\forall a \in [a_L, a_H]|\, m(a) \geq 0$. Consequently, $\forall a \in [a_L, a_H]|\, -\int_a^{a_H} m(s)ds \leq 0$, and in turn $\forall a \in [a_L, a_H]|\, 0 \leq 1 - e^{-\int_a^{a_H} m(s)ds} \leq 1$, i.e., expression (9) represents a probability[6].

Proposition 1

If the government offers a collection of lotteries, $\mathcal{B}$, with the corresponding probabilities defined with condition (9) then an agent characterized by productivity $a \in [a_L, a_H]$ chooses lottery B_a.

Proof. Direct calculations above.

Before we proceed further, let us note that an agent characterized by productivity a chooses actually lottery B_a. Moreover, in the case when the lottery is lost then the agent is presented with the original tax function $\tau(\,)$, chooses to earn y_a and consequently pays $\tau(y_a)$ in taxes and her realized utility is identical to that realized when there are no lotteries at all and equal to $U^L_{a,a} = h(y_a - \tau(y_a) - \frac{1}{2}(\frac{y_a}{a})^2)$. On the other hand, when the uncertainty is resolved favorably then the agent faces a marginal tax of zero and consequently chooses to earn $y^*_a = a^2$ and again pays only $\tau(y_a)$ in taxes with the realized utility now given by $U^W_{a,a} = h(y^*_a - \tau(y_a) - \frac{1}{2}(\frac{y^*_a}{a})^2)$.

Observe that in either case the agent pays $\tau(y_a)$ in taxes, i.e., the government collects the same amount of revenue as in the case when lotteries are not available, i.e., the procedure is revenue neutral on individual basis, not to mention globally. Moreover, the ex-post realized utility is never lower and sometimes is higher than the realized utility when there are no lotteries at all. Furthermore, an agent characterized by productivity a by choosing optimally lottery B_a does not face any uncertainty with regard to the amount of tax liabilities. Irrespective of the actual outcome the amount paid is always equal to $\tau(y_a)$. Naturally, the marginal rate itself is uncertain and in particular is positive when the lottery is lost and is zero when the lottery is won.

3 Optimal Tax Functions: Diamond and Saez - a Theoretical Example

In this section, we illustrate that it can be possible to improve in Pareto sense over allocations induced by the optimal tax functions as described by Saez [17] and Diamond [4] by adopting the basic lottery based procedure described in the previous section.

Recall that the standard optimal taxation problem is typically described by

$$max_{\{\tau(\,)\}} W = \int_{a_L}^{a_H} G(U_a) f(a) da \tag{10}$$

subject to

$$i)\ U_a = h(y_a - \tau(y_a) - \frac{1}{2}(\frac{y_a}{a})^2) \tag{11}$$

$$ii)\ h'(y_a - \tau(y_a) - \frac{1}{2}(\frac{y_a}{a})^2)(1 - \tau'(y_a) - \frac{y_a}{a^2}) = 0 \tag{12}$$

$$iii)\ \int_{a_L}^{a_H} \tau(y_a) f(a) da = R, \tag{13}$$

[5] See a series of papers by Ruiz del Portal [13, 14, 15] who argues that a focus on such a range need not be appropriate.

[6] Note that it could be the case that $p(a_H) = 0$ and $\forall a \in [a_L, a_H)|\, p(a) = 1$ when $\int_a^{a_H} m(s)ds = \infty$. We deal with this issue in subsequent sections.

where R denotes the relevant revenue requirement, $G()$ is normally assumed to be an increasing and concave function reflecting the degree of aversion towards inequality of the policy maker, and where, again, $f(\,)$ denotes the *pdf* corresponding to the distribution of skills. To avoid the discussion of trivial cases we assume that $R > a_L^2 \int_{a_L}^{a_H} f(a)da$, i.e., that it is impossible to satisfy the revenue requirement by taxing all agents a uniform lump-sum tax.

Optimal taxation problems, originally described by Mirrlees [12], defined with conditions (10), (11), (12), and (13) can be approached using a variety of methods. Unfortunately, significant progress - contributions of Diamond [4] and Saez [17] - in characterizing the properties of the optimal tax function has only been possible in the case of utility functions characterized by the absence of income effects and when the distribution of skills is unbounded and thick-tailed. Accordingly, in this section[7] we consider the case when the distribution of types is unbounded, i.e., when $a_H = \infty$, which paradoxically is easier to handle in our context.

We illustrate by considering an explicit example that our approach when applied to the optimal tax functions as described by Diamond [4] and Saez [17] can lead to a Pareto improvement while ensuring revenue neutrality. First, we assume that $h(x) = -e^{-x}$, i.e., we choose a convenient form for $h(\,)$, but we impose no restrictions on $G()$ itself leaving a fair degree of freedom for the overall aversion towards inequality.

Given our assumptions thus far we can express the relevant probability characterizing our lotteries given with expression (9) as $p(a) = 1 - e^{-\int_a^\infty m(s)ds}$ where

$$m(s) = \frac{e^{-(y_s^* - \tau(y_s) - \frac{1}{2}(\frac{y_s^*}{s})^2)}}{-e^{-(y_s^* - \tau(y_s) - \frac{1}{2}(\frac{y_s^*}{s})^2)} + e^{-(y_s - \tau(y_s) - \frac{1}{2}(\frac{y_s}{s})^2)}} \tau'(y_s) \frac{\partial y_s}{\partial s}, \tag{14}$$

which can be, recall that the first order condition (2) implies that $s - \frac{y_s}{s} = s\tau'(y_s)$, further reduced to

$$m(s) = \frac{1}{e^{\frac{1}{2}[s\tau'(y_s)]^2} - 1} \tau'(y_s) \frac{\partial y_s}{\partial s}. \tag{15}$$

Recall that $m(\,)$ is non-negative. Therefore, it is necessarily true that $\int_a^\infty m(s)ds \geq 0$, and, thus it must be that $p(a) = 1 - e^{-\int_a^\infty m(s)ds} \in [0, 1]$, i.e., $p(a)$ can always be interpreted as probability.

Recall that the ingenious approach of Saez [17] led him to a conclusion that the optimal asymptotic marginal tax rates are in fact positive when the distribution of income/skills is heavy tailed, e.g., takes the Pareto form. Thus, if we choose to identify function $\tau(\,)$ with the optimal tax function of Saez [17] then we can state that for large s we have $\tau'(y_s) \to \tau_{ES} > 0$. Furthermore, given that $\tau_{ES} > 0$, it must be the case that for s sufficiently large it must be necessarily true that $(y_s)^N < 2(e^{\frac{1}{2}[s\tau'(y_s)]^2} - 1)$, which in turn allows us to establish, details are provided in Appendix B, that

$$\forall a > a_L \exists a_N | \int_a^\infty m(s)ds < \int_a^{a_N} m(s)ds + \frac{1}{N-1} \frac{2}{\tau(y_{a_N})^{N-1}} < \infty, \tag{16}$$

as, naturally $\int_a^{a_N} m(s)ds$ is finite given that $a > a_L$.

Finally, we can, given that condition (16) holds, state that

$$\forall a > a_L |\, p(a) = 1 - e^{-\int_a^\infty m(s)ds} < 1, \tag{17}$$

i.e., that for all $a > a_L$ and, in particular, for a sufficiently large the losing lotteries are never assigned with certainty when one adopts our lottery-based procedure to the optimal tax function of Saez.

Observe that by not participating in or by opting out of our lottery procedure an agent whose productivity is equal to a attains the utility level of $U_a = -e^{-(y_a - \tau(y_a) - \frac{1}{2}(\frac{y_a}{a})^2)}$. On the other hand, by participating and choosing her lottery optimally the agent can in the case of an unfavorable uncertainty resolution secure utility level of $U_a^L = -e^{-(y_a - \tau(y_a) - \frac{1}{2}(\frac{y_a}{a})^2)}$, which is exactly equal to the level of utility materialized when lotteries are not available or not chosen. Naturally, when uncertainty is resolved favorably then the realized level of utility is given by $U_a^W = -e^{-(\frac{1}{2}a^2 - \tau(y_a))}$, which is strictly higher than the level of utility attained when there are no lotteries present.

Consequently, we can state that by participating in our lottery-based procedure agents are induced by probabilities defined with condition (17) to select lotteries intended for them and, thus, generate exactly the

[7] We postpone the more typical cases with bounded distribution of skills to a subsequent section.

same amount of revenue as in the case when lotteries are not available. Furthermore, ex-post the level of realized utility is never lower than the level of utility attained when the described mechanism is not available. Finally, in the case of a favorable uncertainty resolution, which occurs with a non-zero probability, note that condition (17) is satisfied, the realized level of utility is higher than the level of utility when lotteries are not available. Consequently, we can conclude that in the case when $h(x) = -e^{-x}$ our procedure leads to a Pareto improvement over allocations induced by the optimal tax functions of Saez. Furthermore, we can add that the welfare level as described with condition (10) is increased irrespective of the actual functional form of $G(\,)$ as in the worst case scenario the level of realized utility when lotteries are present is the same as the level of utility realized when there are no lotteries at all.

4 Horizontal Equity

It is customary to self-impose an additional constraint on problems that deal with policy issues. Specifically, in the context of taxation we normally choose to require that identical agents be treated equally. Observe that our basic lottery-based mechanism described above ensures that agents of identical abilities pay identical amounts in taxes. Specifically, recall that in our model agents whose productivity is equal to a select optimally lottery B_a and irrespective of the actual uncertainty resolution always pay $\tau(y_a)$ in taxes. In this limiting sense agents of the same ability are always treated in the same manner, i.e., we can state that in this narrow sense horizontal equity is preserved. Unfortunately, in our basic setup despite the fact that agents of a given ability pay always identical amounts in taxes it is not true that ex post agents of the same ability attain the same level of utility. In fact, agents of ability a whose individual lotteries are resolved favorably attain a level of utility $U^W_{a,a} = h(y^*_a - \tau(y_a) - \frac{1}{2}(\frac{y^*_a}{a})^2)$, which naturally exceeds the level of utility attained by agents whose individual lotteries are resolved unfavorably, $U^L_{a,a} = h(y_a - \tau(y_a) - \frac{1}{2}(\frac{y_a}{a})^2)$. Naturally, given that $U^W_{a,a} > U^L_{a,a}$ we cannot claim that our basic lottery based mechanism leads to ex post horizontal equity in the strong form. Nevertheless, below, we constructively present a minor modification of our mechanism that ensures an identical level of utilities to agents of the same ability, i.e., guarantees that horizontal equity in the strong form is attained in the ex post sense.

Let us now consider a slightly modified version of our lottery based mechanism. In particular, let us consider lotteries of the form

$$B^{'}_a = \begin{cases} \tau(y) - \phi(a) & \text{with probability } p(a) \\ \tau(y_a) + \psi(a) & \text{with probability } 1 - p(a), \end{cases} \tag{18}$$

and let us allow agents to select lotteries from the following collection of lotteries $\mathcal{B}^{'} = \{B^{'}_a \mid a \in [a_L, a_H]\}$.

An agent with productivity a who chooses lottery $B^{'}_b$ faces tax function $\tau(y) - \phi(b)$ when she loses her lottery. Naturally, in this case, she chooses to earn y_a as $\phi(b)$ is just a constant. The agent faces tax function $\tau(y_b) + \psi(b)$ when she wins her lottery. In this case, she chooses to earn y^*_a. Accordingly, she attains the expected utility of

$$EU_{a,b} = p(b)h(y_a - \tau(y_a) + \phi(b) - \frac{1}{2}(\frac{y_a}{a})^2) + (1 - p(b))h(y^*_a - \tau(y_b) - \psi(b) - \frac{1}{2}(\frac{y^*_a}{a})^2).$$

The relevant efficiency condition, $\frac{\partial EU_{a,b}}{\partial b} = 0$, for $b = a$ becomes

$$\begin{gathered} p'(a)[h(y_a - \tau(y_a) + \phi(a) - \tfrac{1}{2}(\tfrac{y_a}{a})^2)+ \\ -h(y^*_a - \tau(y_a) - \psi(a) - \tfrac{1}{2}(\tfrac{y^*_a}{a})^2)]+ \\ +p(a)h'(y_a - \tau(y_a) + \phi(a) - \tfrac{1}{2}(\tfrac{y_a}{a})^2)\phi'(a)+ \\ +(1 - p(a))h'(y^*_a - \tau(y_a) - \psi(a) - \tfrac{1}{2}(\tfrac{y^*_a}{a})^2)(-\tau'(y_a)\tfrac{\partial y_a}{\partial a} - \psi'(a)) = 0. \end{gathered} \tag{19}$$

Note that up to this stage we have not put any restrictions on functions $\phi(\,)$ and $\psi(\,)$. We choose now functions in such a way as to ensure the same level of utility ex post irrespective of the actual uncertainty resolution. In fact, we require that $\phi(\,)$ and $\psi(\,)$ be such so that we have

$$y_a - \tau(y_a) + \phi(a) - \frac{1}{2}(\frac{y_a}{a})^2 = y^*_a - \tau(y_a) - \psi(a) - \frac{1}{2}(\frac{y^*_a}{a})^2, \tag{20}$$

which naturally is equivalent to, recall that $y_a^* = a^2$,

$$\phi(a) + \psi(a) = \frac{1}{2}[a\tau'(y_a)]^2. \tag{21}$$

Naturally, condition (20) implies that $h(y_a - \tau(y_a) + \phi(a) - \frac{1}{2}(\frac{y_a}{a})^2) = h(y_a^* - \tau(y_a) - \psi(a) - \frac{1}{2}(\frac{y_a^*}{a})^2)$. Thus, now condition (19) can be simplified to

$$\begin{aligned} p(a)h'(y_a - \tau(y_a) + \phi(a) - \tfrac{1}{2}(\tfrac{y_a}{a})^2)\phi'(a) + \\ +(1-p(a))h'(y_a^* - \tau(y_a) - \psi(a) - \tfrac{1}{2}(\tfrac{y_a^*}{a})^2)(-\tau'(y_a)\tfrac{\partial y_a}{\partial a} - \psi'(a)) = 0. \end{aligned} \tag{22}$$

Furthermore, given that condition (20) holds we naturally have $h'(y_a - \tau(y_a) + \phi(a) - \frac{1}{2}(\frac{y_a}{a})^2) = h'(y_a^* - \tau(y_a) - \psi(a) - \frac{1}{2}(\frac{y_a^*}{a})^2)$. Hence, condition (22) becomes

$$p(a)\phi'(a) - (1-p(a))\psi'(a) = (1-p(a))\tau'(y_a)\frac{\partial y_a}{\partial a}. \tag{23}$$

Let us now impose further restrictions on $\phi(\ )$ and $\psi(\ )$. In particular, let us require that $p(a)\phi(a) = (1-p(a))\psi(a)$, which when combined with condition (21) yields $\phi(a) = (1-p(a))\frac{1}{2}[a\tau'(y_a)]^2$ and $\psi(a) = p(a)\frac{1}{2}[a\tau'(y_a)]^2$. Consequently, condition (23) simplifies to

$$\frac{-p'(a)}{1-p(a)} = \frac{\tau'(y_a)\frac{\partial y_a}{\partial a}}{\frac{1}{2}[a\tau'(y_a)]^2}, \tag{24}$$

which upon integrating and imposing boundary condition $p(a_H) = 0$ yields[8]

$$p(a) = 1 - e^{-\int_a^{a_H} \frac{\tau'(y_s)\frac{\partial y_s}{\partial s}}{\frac{1}{2}[s\tau'(y_s)]^2} ds}. \tag{25}$$

Note that the expression given above describing the relevant probability does not involve the distribution of skills nor does it rely on the form of $G(\ )$ and, thus, the implementation of our procedure is less informationally demanding than the implementation of optimal tax schedules.

Naturally, by construction probabilities described with condition (25) induce agents with productivity a to select lottery $B_a^{'}$. When the uncertainty is resolved unfavorably an agent whose productivity is equal to a faces tax function $\tau(y) - (1-p(a))\frac{1}{2}[a\tau'(y_a)]^2$, earns y_a, pays $\tau(y_a) - (1-p(a))\frac{1}{2}[a\tau'(y_a)]^2$ in taxes and attains a utility level of $U_a^{Loss} = h(y_a - \tau(y_a) + (1-p(a))\frac{1}{2}[a\tau'(y_a)]^2 - \frac{1}{2}(\frac{y_a}{a})^2)$. In the case of a favorable uncertainty resolution the tax liabilities are dictated with $\tau(y_a) + p(a)\frac{1}{2}[a\tau'(y_a)]^2$, the agent earns $y_a^* = a^2$, and attains utility level of $U_a^{Win} = h(y_a^* - \tau(y_a) - p(a)\frac{1}{2}[a\tau'(y_a)]^2 - \frac{1}{2}(\frac{y_a^*}{a})^2)$.

By construction, definitions of $\phi(\ )$ and $\psi(\ )$, it is necessarily true that $U_a^{Loss} = U_a^{Win}$, i.e., ex-post all agents of productivity a attain the same level of utility. Clearly, our mechanism ensures that identical agents are treated equally and in all states of nature attain identical levels of happiness. In other words, the specific mechanism described in this section ensures that horizontal equity in the strong form is satisfied ex-post.

For completeness let us note that the total revenue collected from agents of productivity a when there are no lotteries is given by $R_a = \tau(y_a)f(a)$. On the other hand, when lotteries are available the revenue collected, given the law of large numbers, can be expressed as $R_a^{'} = \{p(a)[\tau(y_a) - \phi(a)] + (1-p(a))[\tau(y_a) + \psi(a)]\}f(a) = \{\tau(y_a) - p(a)\phi(a) + (1-p(a))\psi(a)\}f(a) = \{\tau(y_a) - p(a)(1-p(a))\frac{1}{2}[a\tau'(y_a)]^2 + (1-p(a))p(a)\frac{1}{2}[a\tau'(y_a)]^2\}f(a) = \tau(y_a)f(a)$, i.e., $R_a = R_a^{'}$. Clearly, the above described procedure not only ensures horizontal equity ex-post, but also is revenue neutral.

Finally, let us note that by participating in the above-described mechanism agents of productivity a attain with certainty the level of utility of $U_a^{LOT} = h(y_a - \tau(y_a) + (1-p(a))\frac{1}{2}[a\tau'(y_a)]^2 - \frac{1}{2}(\frac{y_a}{a})^2)$, which is never lower than the level of utility, $U_a = h(y_a - \tau(y_a) - \frac{1}{2}(\frac{y_a}{a})^2)$, attained when there are no lotteries at all. This indicates that agents when offered an opportunity to participate in our lottery based mechanism will do so voluntarily.

[8]Naturally, the expression below remains practically relevant only when $\int_a^{a_H} \frac{\tau'(y_s)\frac{\partial y_s}{\partial s}}{\frac{1}{2}[s\tau'(y_s)]^2} ds < \infty$. We deal with the opposite case in a subsequent section.

5 Welfare Gains - Lower Bound

It is customary for governments to rely on piece-wise linear tax functions to secure their revenue needs. Only recently, governments, given the results of Diamond [4] and Saez [17], started to contemplate replacing the traditional tax functions with the optimal tax functions. Undoubtedly, a move towards the optimal tax function, if implemented, must lead to welfare gains. Unfortunately, such a move requires the knowledge of the distribution of skills and the preferences of the policy maker. Furthermore, the political will plays a role as some constituents must suffer a reduction in utility if one adopts the optimal tax function of Saez/Diamond. The situation is much simpler if one considers a reform based on lotteries described in this paper, as implementation does not require the knowledge of $G(\,)$ or $f(\,)$. Furthermore, our procedure is by construction Pareto improving and thus all political constraints are void as well. In other words, the implementation of our procedure is simple. This just leaves the question of practical benefits from implementing the procedure. In this section we assess the quantitive benefits that could potentially accrue should the procedure be implemented.

To gauge the welfare gains that may stem from implementing our procedure we rely on the utility function of Saez [17], which takes the form

$$U_S(C,L) = h(C - \frac{1}{1+\frac{1}{\eta}}L^{1+\frac{1}{\eta}}). \tag{26}$$

In this case, the magnitude of the tax distortion, when the marginal tax rate is constant, measured by the dead weight burden is $D_a = a^{1+\eta}[\frac{1}{1+\eta}(1-(1-\tau)^{1+\eta}) - \tau(1-\tau)^{\eta}]$ and the relevant probability takes the form $p(a) = 1 - (\frac{a}{a_H})^{w_p}$, where $w_p = \frac{(1+\eta)^2\tau(1-\tau)^{\eta}}{1-(1-\tau)^{1+\eta}-(1+\eta)\tau(1-\tau)^{\eta}}$.

Now, borrowing again from the previous section it can be shown that the welfare gains of an agent whose productivity is equal to a are equivalent to a monetary transfer of $(1-p(a))D_a$, i.e., our randomizing mechanism eliminates proportion $1-p(a)$ of the underlying tax distortion for an agent of productivity a.

In reality, tax functions typically take the form of piece-wise linear tax functions and, thus, are not differentiable everywhere. Still, our procedure can easily extended to such cases utilizing an approach outlined in Dudek [6]. Here, for simplicity we focus on a subcategory of tax payers and assess the welfare gains of high-income earners, i.e., agents who, in particular, face the highest and constant marginal tax rate τ. Now, following Saez [17], we assume that the distribution of skill is Paretian with $f(a) = Aa^{-m}$ being the corresponding *pdf*. Under these assumptions, the average reduction in tax distortion in the highest income bracket, assuming that agents with productivities $a \in [\bar{a},\, a_H]$ self select themselves and pay the highest marginal tax rate τ, takes the form

$$RD = \frac{\int_{\bar{a}}^{a_H}(1-p(a))D_a f(a)da}{\int_{\bar{a}}^{a_H} D_a f(a)da} = \frac{2+\eta-m}{2+\eta-m+w_p}\,\frac{1-(\frac{\bar{a}}{a_H})^{2+\eta-m+w_p}}{1-(\frac{\bar{a}}{a_H})^{2+\eta-m}}. \tag{27}$$

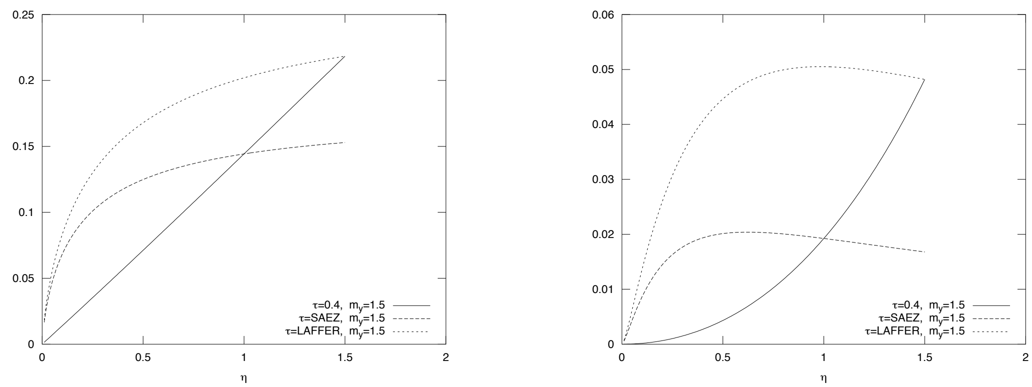
Furthermore, there is, again as noted by Saez [17], a simple relationship between income and skill distribution when the distribution of skills is assumed to be Paretian. In particular, in such a case, the distribution of income is Paretian as well with $g(y) = By^{-m_y}$ being the corresponding *pdf*, where $m = m_y + \eta(m_y - 1)$. Moreover, an agent whose productivity is $a \in [\bar{a},\, a_H]$ earns $y = (1-\tau)a^{1+\eta}$. Therefore, we can simplify expression (27) to

$$RD = \frac{2-m_y}{2-m_y+\frac{w_p}{1+\eta}}\,\frac{1-(\frac{\bar{y}}{y_H})^{2-m_y+\frac{w_p}{1+\eta}}}{1-(\frac{\bar{y}}{y_H})^{2-m_y}}. \tag{28}$$

It is straightforward to establish that the relative welfare gains depend on the values of four parameters: m_y, the parameter characterizing the distribution of income, the actual marginal tax rate, τ, and the width of the highest income tax bracket as measured by ratio $\frac{\bar{y}}{y_H}$, i.e., the ratio of the lowest to highest income earned, and on the value of elasticity η.

In general, the value of RD is sensitive to the values of the underlying parameters and can assume substantial magnitudes. Still, recognizing the limitations of our approach, we are interested in providing meaningful estimates of the welfare gain. Accordingly, in our simulations we set $m_y = 1.5$, i.e., to a relatively low value, which according to Diamond and Saez [5] corresponds to the current empirical estimate in the

US. Furthermore, we set the ratio of the lowest to the highest income earned in the highest tax bracket to 0.0001. In addition, we consider three scenarios for the tax rate τ. Specifically, first we examine what the implied gains are when $\tau = 0.4$, which again as noted by Saez [17] roughly corresponds to the actual rate faced by high-income earners in the US. Secondly, we estimate the gains when the actual tax rate is equal to the optimal rate of Saez, i.e., when $\tau = \frac{1}{1+m_y\eta}$. Finally, we provide estimates of welfare gains when the tax rate is equal to the simple static revenue maximizing rate, i.e., when $\tau = \frac{1}{1+\eta}$. We present our estimates in the figure (1) below.



(a) The Reduction in the Average Distortion as a Function of Elasticity η.

(b) The Benefit as a Fraction of the GDP as a Function of Elasticity η.

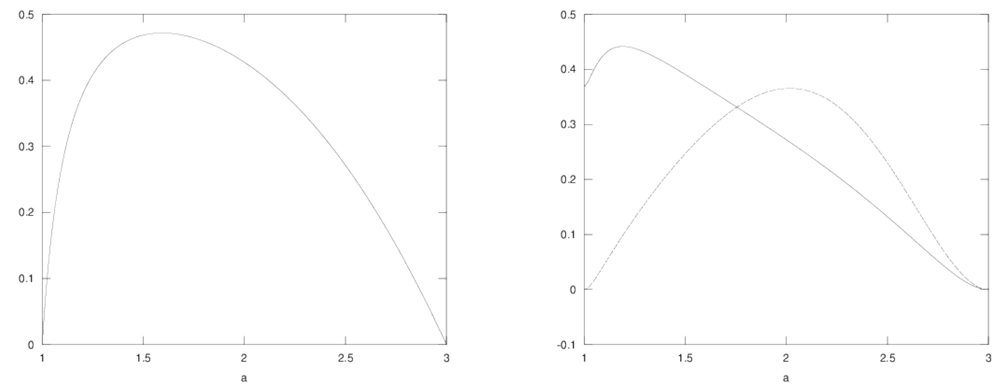
Figure 1: The Expected Benefit for High-Income Earners.

The figures above reveal that the values of relative gains are sensitive to the values of the underlying parameters. Nevertheless, the actual benefits can be significant and, in particular, exceed 10% of the magnitude of the average of the underlying dead weight burden. Equivalently, we can state that the average welfare gains can be as high as 5% of the GDP for high-income earners.

6 Optimal Tax Functions: Mirrlees - a Numerical Example

In this section we illustrate, by providing an explicit numerical example, that our mechanism can be applicable to solutions to optimal taxation problems of Mirrlees. Specifically, we consider an optimal taxation problem of Mirrlees defined with equations (10), (11), (12), and (13), where $a_L = 1$, $a_H = 3$, $f(a) = \frac{1}{a_H - a_L}$, $G(U) = U$, $h(x) = -e^{-Bx}$, $B = 3$, and $R = 0$. Such a problem can be solved with numerical techniques. The optimal marginal tax rates are presented in figure (2a).

Below we illustrate that our randomizing procedure can be applied to the solution of the Mirrlees problem outlined above, i.e., we show that with our method one can, in principle, generate a Pareto improvement overall allocations induced by optimal tax functions of Mirrlees.



(a) The Optimal Marginal Tax Rates

(b) Solution to Condition (35).

Figure 2: Numerical Solution to the Optimal Taxation Problem defined with conditions (10), (11), (12), and (13), where $a_L = 1$, $a_H = 3$, $f(a) = \frac{1}{a_H - a_L}$, $G(U) = U$, $h(x) = -e^{-Bx}$, $B = 3$, and $R = 0$.

Recall that solutions to such problems involve, as visualized in figure (2a) and formally established by Sadka [16] and Saede [19], zero marginal tax rates at both ends of the skill distribution.[9], i.e., $\tau^{'}_M(y_{a_L}) = 0$ and $\tau^{'}_M(y_{a_H}) = 0$. Unfortunately, it turns out that the fact that the optimal marginal tax rate at the highest level of income is equal to zero precludes our basic lottery based mechanism from being applied directly as the integrand, $m(\,)$, defining the relevant probabilities becomes undefined, in particular, at the highest skill level. The presence of poles/singularities of $m(\,)$ at the highest skill/income level implies that the relevant probabilities are equal to 1, i.e., the procedure in its basic form leads to no welfare improvement. Naturally, a modification is necessary if one is interested in improving over the allocation induced the optimal tax function, $\tau_M(\,)$. We describe such a modification next.

We choose to borrow heavily from section 4 and assume that agents again can choose lotteries from the set described with expression (18). Again, as established in section 4 horizontal equity is preserved and agents in equilibrium select lotteries intended for them when $\phi(a) + \psi(a) = \frac{1}{2}[a\tau^{'}_M(y_a)]^2 = D_a$ and the relevant probability $p(a)$ is defined with condition (23), with $\tau(\,)$ replaced with $\tau_M(\,)$, which now can be written as

$$p(a)\phi'(a) - (1 - p(a))\psi'(a) = (1 - p(a))\tau^{'}_M(y_a)\frac{\partial y_a}{\partial a}. \tag{29}$$

Furthermore, let us assume that $\phi(a) = (1 - p(a))D_a - \lambda p(a) + \omega$, and $\psi(a) = p(a)D_a + \lambda p(a) - \omega$, where $\lambda > 0$ and ω are constants. Now, condition (29) defining $p(a)$ becomes

$$\frac{-p(a)}{1 - p(a)} = \frac{\tau^{'}_M(y_a)\frac{\partial y_a}{\partial a}}{D_a + \lambda} = m(a). \tag{30}$$

Note that $m(a)$ in the above condition is always nonnegative and it does not have a pole at $a = a_H$ as

[9] When agents at the bottom of the skill distribution choose to be inactive the optimal marginal tax rate is in fact positive. The presence of this possibility has no bearing on our results.

$\lambda > 0$. Thus, we can integrate the above condition and express the relevant probability as

$$p(a) = 1 - e^{-\int_a^{a_H} m(s)ds}. \tag{31}$$

Naturally, in this case $m(a)$ is bounded for all $a \in [a_L, a_H]$. Hence, we can conclude that $0 < p(a) < 1$ for all $a < a_H$ and $p(a_H) = 0$.

The level of realized utility of individuals who choose to participate in our randomizing mechanism is given by $U_a^R = h(y_a - \tau_M(y_a) - \frac{1}{2}(\frac{y_a}{a^2})^2 + (1-p(a))D_a - \lambda p(a) + \omega)$ irrespective of the actual uncertainty resolution. Agents who choose not to participate in our mechanism secure utility $U_a = h(y_a - \tau_M(y_a) - \frac{1}{2}(\frac{y_a}{a^2})^2)$. Consequently, we can conclude that the participation constraint is given by

$$(1 - p(a))D_a - \lambda p(a) + \omega \geq 0. \tag{32}$$

It is straightforward to show that if an agent of productivity a chooses to participate then agents with higher productivities participate as well. Thus, the set of participating agents $\mathcal{P}$ takes the form $\mathcal{P} = [a_I, a_H]$, where, naturally, condition (32) holds with equality for $a = a_I$.

Recall that in section 4 we required that the total revenue collected when our randomizing mechanism is chosen from an agent whose productivity is equal to a be the same as the revenue collected when there is no randomization at all. To attain that objective we set $\phi(\,)$ and $\psi(\,)$ accordingly. In this section, we impose a milder restriction and just require that the total revenue collected when the randomizing mechanism is present be the same as the total revenue collected when there is no randomization at all. A few steps of algebra reveal that the revenue neutrality is achieved as long as the following condition is met

$$\int_{a\in\mathcal{P}} p(a)\phi(a)f(a)da = \int_{a\in\mathcal{P}} (1 - p(a))\psi(a)f(a)da, \tag{33}$$

which reduces to

$$\lambda \int_{a_I}^{a_H} p(a)f(a)da = \omega \int_{a_I}^{a_H} f(a)da. \tag{34}$$

Now, the issue of a possibility of a Pareto improvement over the allocation induced by the optimal tax function of Mirrlees, $\tau_M(\,)$, amounts to a technical requirement of finding values of λ, a_I, and ω that satisfy simultaneously condition (34), condition (32) with equality with $p(a)$ being given with condition (31). In fact, conditions (34) and (32), with $a = a_I$, can be combined to yield

$$D_{a_I} = \lambda \frac{\int_{a_I}^{a_H} [p(a_I) - p(a)]f(a)da}{(1 - p(a_I)) \int_{a_I}^{a_H} f(a)da}. \tag{35}$$

Therefore, to prove that there is a possibility of a Pareto improvement over the allocations induced by the optimal tax function of Mirrlees, $\tau_M(\,)$, it suffices to find values of λ and a_I that satisfy condition (35) where $p(a)$ is given with condition (31).

The left-hand and the right-hand sides of equation (35) are plotted in figure (2b). The graph reveals that indeed there is a solution to equation (35). Specifically, $a_I = 1.75552$ solves equation (35) when $\lambda = 0.2 \max(D_a) = 0.073109$. Therefore, we can conclude that it is possible to improve, while preserving revenue, in the Pareto sense over the allocation induced by the optimal tax function of Mirrlees, $\tau_M(\,)$, with the corresponding optimal marginal tax rates presented in figure (2a), by offering lotteries with the relevant probabilities defined with condition (31) where $\lambda = 0.073109$ and $\omega = 0.054420$, with agents with productivities higher than $a_I = 1.75552$ voluntarily participating in the mechanism.

Naturally, our numerical example presented above indicates that our basic randomizing mechanism can occasionally be applied to generate a Pareto improvement over allocations induced by optimal tax functions of Mirrlees. Furthermore, one can expect that our randomizing mechanism can be applied more broadly if one is interested only in welfare improvement rather than in a more demanding requirement of Pareto improvement.

7 Conclusions

We constructively show that our routine inclination to dismiss lump-sum taxes as a viable policy option can be highly premature. It turns out that lump-sum taxes need not be ignored as a policy tool even in the context of asymmetric information. While we naturally concede that the first best outcome is not attainable given the standard information constraints, we show that it is possible to implement - what we refer to as - the second best *plus* allocation by relying on lump-sum taxes assigned to economic agents with some - endogenously determined - probability. We explicitly outline a framework where agents face ex-ante randomness and pay lump-sum taxes with some probability. We show that such a simple modification is revenue neutral and allows to exploit additional efficiency gains resulting form the probabilistic presence of lump-sum taxes. Our approach is general in nature. In fact, it relies on lesser informational demands than those of standard optimal taxation problems. Furthermore, at the technical level our approach involves a continuum of agents making it suitable for practical considerations. Finally, we illustrate that our approach can be applicable to improve over allocations induced by the optimal tax functions of Saez [17], Diamond [4], and Mirrlees [12].

The procedure described in the text can be extended to other contexts with the presence of imperfect information. In general, in any mechanism that involves screening it is possible to consider ex-ante randomization that involves the original underlying mechanism and just the payoff/payment generated in equilibrium in the original mechanism. Such as simple modification can, as shown in this paper, be feasible - with the equilibrium probabilities being well defined - and more importantly can lead to additional welfare gains that are unattainable in deterministic setups. We want to emphasize that procedures outlined in this paper can be potentially applicable as in most practical cases they involve only upside risk and can be offered - and likely be accepted - as a supplementary option to the procedures currently in place.

A Appendix

In this appendix we show that our basic mechanism can be applied in the context of any utility function not just the one that we work with in the main part of the paper. Let us consider a utility function of the form $U = u(C, L)$, which naturally given our standard assumptions can be expressed as $U = u(y - \tau(y), \frac{y}{a})$, where $\tau(\,)$ is a given tax function. Let y_a denote the optimal choice given tax function $\tau(\,)$. Naturally, we have

$$u_1(y_a - \tau(y_a), \frac{y_a}{a})(1 - \tau^{'}(y_a)) + u_2(y_a - \tau(y_a), \frac{y_a}{a})\frac{1}{a} = 0. \tag{36}$$

Let us denote with $y_a^*(T)$ the optimal choice of an agent whose productivity is equal to a when she faces a lump sum tax of T. Naturally, $y_a^*(T)$ solves $max_{\{y\}} u(y - T, \frac{y}{a})$ and is implicitly given by

$$u_1(y_a^*(T) - T, \frac{y_a^*(T)}{a}) + u_2(y_a^*(T) - T, \frac{y_a^*(T)}{a})\frac{1}{a} = 0. \tag{37}$$

Let us now assume that an agent whose productivity is equal to a chooses lottery B_b for some $b \in [a_L, a_H]$ described with condition (4). Note that such an agent faces $\tau(\,)$ when lottery is resolved unfavorably and faces a lump-sum tax of $\tau(y_b)$ when the lottery is resolved favorably. Naturally, the agent chooses to earn y_a when she loses and chooses to earn $y_a^*(\tau(y_b))$ when she wins. Her expected payoff can be expressed as

$$EU_{a,b} = p(b)u(y_a - \tau(y_a), \frac{y_a}{a}) + (1 - p(b))u(y_a^*(\tau(y_b)) - \tau(y_b), \frac{y_a^*(\tau(y_b))}{a}). \tag{38}$$

Naturally, the agent wants to maximize her expected payoff, which requires $\frac{\partial UE_{a,b}}{\partial b} = 0$. Furthermore, we naturally want the agent to be choosing a lottery that is intended for her, i.e., we want to the efficiency condition, $\frac{\partial UE_{a,b}}{\partial b} = 0$, to hold when $b = a$, which is equivalent to

$$\begin{gathered} p'(a)u(y_a - \tau(y_a), \tfrac{y_a}{a}) - p'(a)u(y_a^*(\tau(y_a)) - \tau(y_a), \tfrac{y_a^*(\tau(y_a))}{a}) \\ (1 - p(a))[u_1(y_a^*(\tau(y_a)) - \tau(y_a), \tfrac{y_a^*(\tau(y_a))}{a}) \\ (\tfrac{\partial y_a^*(\tau(y_a))}{\partial \tau(y_a)}\tau'(y_a)\tfrac{\partial y_a}{\partial a} - \tau'(y_a)\tfrac{\partial y_a}{\partial a}) + \\ u_2(y_a^*(\tau(y_a)) - \tau(y_a), \tfrac{y_a^*(\tau(y_a))}{a})\tfrac{1}{a}\tfrac{\partial y_a^*(\tau(y_a))}{\partial \tau(y_a)}\tau'(y_a)\tfrac{\partial y_a}{\partial a}] = 0. \end{gathered} \tag{39}$$

Now, utilizing the first order condition, (37), we can rewrite the above expression as

$$\begin{aligned} p'(a)u(y_a - \tau(y_a), \tfrac{y_a}{a}) - p'(a)u(y_a^*(\tau(y_a)) - \tau(y_a), \tfrac{y_a^*(\tau(y_a))}{a}) + \\ (1-p(a))u_1(y_a^*(\tau(y_a)) - \tau(y_a), \tfrac{y_a^*(\tau(y_a))}{a})(-\tau'(y_a)\tfrac{\partial y_a}{\partial a}) \quad &= 0, \end{aligned} \tag{40}$$

which naturally reduces to

$$\frac{-p'(a)}{1-p(a)} = \frac{u_1(y_a^*(\tau(y_a)) - \tau(y_a), \frac{y_a^*(\tau(y_a))}{a})\tau'(y_a)\frac{\partial y_a}{\partial a}}{u(y_a^*(\tau(y_a)) - \tau(y_a), \frac{y_a^*(\tau(y_a))}{a}) - u(y_a - \tau(y_a), \frac{y_a}{a})}, \tag{41}$$

which upon integrating and imposing condition that $p(a_H) = 0$ becomes

$$p(a) = 1 - e^{-\int_a^{a_H} m(s)ds}, \tag{42}$$

where

$$m(s) = \frac{u_1(y_s^*(\tau(y_s)) - \tau(y_s), \frac{y_s^*(\tau(y_s))}{s})\tau'(y_s)\frac{\partial y_s}{\partial s}}{u(y_s^*(\tau(y_s)) - \tau(y_s), \frac{y_s^*(\tau(y_s))}{s}) - u(y_s - \tau(y_s), \frac{y_s}{s})}. \tag{43}$$

Note that typically marginal tax rates are assumed to be positive. Furthermore, income earned should be a non-decreasing function of the skill level. Naturally, we also assume that consumption is attractive, i.e., $u_1(\ ) > 0$. Moreover, it is always beneficial to pay $\tau(y_s)$ lump-sum rather than in a distortionary manner, i.e., it must be $u(y_s^*(\tau(y_s)) - \tau(y_s), \frac{y_s^*(\tau(y_s))}{s}) \geq u(y_s - \tau(y_s), \frac{y_s}{s})$. Thus, function $m(s)$ given with condition (43) assumes only non-negative values. Therefore, $\forall a \in [a_L, a_H] | \int_a^{a_H} m(s)ds \geq 0$, and consequently $\forall a \in [a_L, a_H] | \, p(a) \in [0, 1]$, i.e., $p(\ )$ can be in fact interpreted as probability. Hence, we can conclude that our procedure described in the main part of the paper is applicable to general utility functions.

B Appendix

In this appendix we establish the validity of the bound expressed with condition (16) in the main text. First, note that $m(s)$ defined in the main text can be rewritten as

$$m(s) = \frac{(y_s)^N}{e^{\frac{1}{2}[s\tau'(y_s)]^2} - 1} \frac{\tau'(y_s)\frac{\partial y_s}{\partial s}}{(y_s)^N} \tag{44}$$

for any N.

Let us note that by assumption we have $lim_{s\to\infty}\tau'(y_s) = \tau_{ES} > 0$. Therefore by continuity we necessarily have that $lim_{s\to\infty} e^{\frac{1}{2}[s\tau'(y_s)]^2} - 1 = e^{\frac{1}{2}[s\tau_{ES}]^2} - 1$. Therefore, for s sufficiently large it must be

$$e^{\frac{1}{2}[s\tau_{ES}]^2} - 1 < 2(e^{\frac{1}{2}[s\tau'(y_s)]^2} - 1). \tag{45}$$

Moreover, for any N it is necessarily true that for some a_N it must be, recall that $y_s \leq s^2$,

$$\forall s > a_N | \, (y_s)^N < e^{\frac{1}{2}[s\tau_{ES}]^2} - 1, \tag{46}$$

which naturally, given inequality (45), implies that

$$\forall s > a_N | \, \frac{(y_s)^N}{e^{\frac{1}{2}[s\tau'(y_s)]^2} - 1} < 2. \tag{47}$$

Therefore, we have

$$\forall s > a_N | \, m(s) = \frac{(y_s)^N}{e^{\frac{1}{2}[s\tau'(y_s)]^2} - 1} \frac{\tau'(y_s)\frac{\partial y_s}{\partial s}}{(y_s)^N} < 2\frac{\tau'(y_s)\frac{\partial y_s}{\partial s}}{(y_s)^N}. \tag{48}$$

Furthermore, given that $\tau(y_s) < y_s$ we can naturally write

$$\forall s > a_N | \, m(s) < 2\frac{\tau'(y_s)\frac{\partial y_s}{\partial s}}{(\tau(y_s))^N}. \tag{49}$$

Thus, we can write

$$\forall a > a_L | \int_a^{\infty} m(s)ds = \int_a^{a_N} m(s)ds + \int_{a_N}^{\infty} m(s)ds < \int_a^{a_N} m(s)ds + 2\int_{a_N}^{\infty} \frac{\tau'(y_s)\frac{\partial y_s}{\partial s}}{(\tau(y_s))^N} ds, \tag{50}$$

which naturally reduces to

$$\forall a > a_L | \int_a^{\infty} m(s)ds < \int_a^{a_N} m(s)ds + \frac{1}{N-1}\frac{2}{\tau(y_{a_N})^{N-1}}, \tag{51}$$

as presumably $lim_{s\to\infty}\tau(y_s) = \infty$.

Furthermore, both $\int_a^{a_N} m(s)ds$ and $\frac{1}{N-1}\frac{2}{\tau(y_{a_N})^{N-1}}$ are finite as we assume that $N > 1$. Therefore, we can write $\forall a > a_L | \int_a^{\infty} m(s)ds < \infty$, which establishes condition (16) in the main text.

References

[1] Baisa, Brian (2013). "Auction Design without Quasilinear Preferences," forthcoming Theoretical Economics.

[2] Brito, Dagobert L., Jonathan H. Hamilton, Steven M. Slutsky, and Joseph E. Stiglitz, (1995). "Randomization in optimal income tax Schedules," Journal of Public Economics, 56, pp. 189-223.

[3] Del Negro, Marco, Fabrizio Perri and Fabiano Schivardi, (2010), "Tax Buyouts," Journal of Monetary Economics, 57, pp. 576–595.

[4] Diamond, Peter A., (1998). "Optimal Income Taxation: An Example with a U-Shaped pattern of Optimal analysis Tax Rates," American Economic Review, 88, pp. 83-95.

[5] Diamond, Peter A., and Emmanuel Saez (2011). "From Basic Research to Policy Recommendations: The Case for a Progressive Tax," Journal of Economic Perspectives, 25, pp. 165-190.

[6] Dudek, Maciej K. (2008). "On the Welfare Improving Trade in Marginal Tax Rates," Available at SSRN: http://ssrn.com/abstract=1185577 or http://dx.doi.org/10.2139/ssrn.1185577

[7] Ebert, Udo, (1992). "A Reexamination of the Optimal Nonlinear Income Tax," Journal of Public Economics, 49, pp. 47-73.

[8] Rod Garratt and Marek Pycia, (2014), "Efficient Bilateral Trade," UCLA working paper.

[9] Gauthier, Stéphane and Guy Larroque (2015), "Redistribution by Means of Lotteries," Documents de Travail du Centre d'Economie de la Sorbonne, CES Working Papers.

[10] Hellwig, Martin (2007). "The Undesirability of Randomized Income Taxation under Decreasing Risk Aversion," Journal of Public Economics, 91, pp. 791-816.

[11] Lollivier, Stefan and Jean-Charles Rochet, (1983). "Bunching and Second Order Conditions: a Note on Optimal Tax Theory," Journal of Economic Theory, 31, pp. 392-400.

[12] Mirrlees, James A., (1971). "An Exploration in the Theory of Optimal Income Taxation," Review of Economic Studies, 38, pp. 175-208.

[13] Ruiz del Portal, Xavier. "The problem of optimum income taxation: a remark on its monotonicity constraint." Journal of Public Economic Theory 9.2 (2007): 265-283.

[14] Ruiz del Portal, Xavier. "Is the optimal income tax regressive?." Economics Letters 100.3 (2008): 402-404.

[15] Ruiz del Portal, Xavier. "On the qualitative properties of the optimal income tax." Mathematical Social Sciences 59.3 (2010): 288-298.

[16] Sadka, Efraim, (1976). "On Income Distribution, Incentive Effects and Optimal Income Taxation," Review of Economic Studies, 42, pp. 261-268.

[17] Saez, Emmanuel, (2001). "Using Elasticities to Derive Optimal Income tax Rates," Review of Economic Studies, 68, pp. 205-229.

[18] Saez, Emmanuel, (1999). "A Characterization of the Income Tax Schedule Minimizing Deadweight Burden," Chapter 2, MIT Ph.D. Thesis.

[19] Seade, Jesus K., (1977). "On the Shape of Optimal Tax Schedules," Journal of Public Economics, 7, pp. 203-236.

[20] Seade, Jesus K., (1982). "On the Sign of the Optimum Marginal Income Tax," Review of Economic Studies, 49, pp. 637-643.

[21] Stiglitz, Joseph E., (1982). "Self-Selection and Pareto Efficient Taxation," Journal of Public Economics, 17, pp. 213-240.

[22] Stiglitz, Joseph E., (1987). "Pareto Efficient and Optimal Taxation and the New New Welfare Economics," Handbook of Public Economics, Volume 2, edited by A.J. Auerbach and M. Feldstein, Elsevier Science Publishers B.V. (North Holland)

[23] Pestieau Pierre, Uri Possen, and Steven Slutsky, (2002), "Randomization, Revelation, and Redistribution in a Lerner World," Economic Theory. 20, pp. 539–553.

[24] Piketty, Thomas, (1993), "Implementation of First-Best Allocations via Generalized Tax Schedules," Journal of Economic Theory, 61, pp. 23-41.