Economic Quarterly—Volume 93, Number 1—Winter 2007—Pages 77–109

Optimal Nonlinear Income Taxation with Costly Tax Avoidance

Borys Grochulski

The central idea behind an important branch of modern public finance literature is that imperfect government information about taxpayers' individual characteristics limits the economic outcomes attainable by taxation and redistribution policies. This idea, first explored in a seminal article by James Mirrlees (1971), provides a framework for studying the fundamental question of how income should be taxed.[1] In this framework, which has become known as the Mirrlees approach to optimal taxation, an optimal tax system is one that implements the best economic outcome attainable under the constraints imposed by limited physical resources and limited government information. Optimal tax systems derived within the Mirrlees framework contribute to our understanding of the observed tax institutions and can serve as a basis for deriving normative prescriptions for tax policy reforms.

In this article, we use the Mirrlees approach to study the question of optimal income taxation in an environment in which agents can avoid taxation by hiding income. In this environment, the government cannot observe individual income of the agents in the population, but only the income that agents choose to display. Income displayed may be less than actual income. However, the process of income hiding is costly; when income is being concealed, some resources are wasted on income-hiding activities. The concealed income is never observed by the government; it is consumed by the agents in private. True income, therefore, cannot be taxed. Taxes only can be levied on the displayed income.

I would like to thank Marina Azzimonti-Renzo, Brian Minton, Ned Prescott, and Alex Wolman for their helpful comments. The views expressed in this article are those of the author and not necessarily those of the Federal Reserve Bank of Richmond or the Federal Reserve System.

[1] Stiglitz (1987) provides an overview of early contributions to this literature. Recent contributions, which are mostly concerned with dynamic models (e.g., Kocherlakota 2005 and Albanesi and Sleet 2006), are reviewed in Kocherlakota (2006).

The government's objective is to use redistributive taxation to provide agents with insurance against the individual income risk. The income concealment technology available to agents restricts the amount of tax revenue that can be raised and used for redistribution. If the marginal tax rate applied to income level y is higher than the agents' cost to conceal the yth dollar of their income, it is in the best interest of all agents whose true income is y to conceal the last dollar of their income, display income $y - 1$, and incur the concealment cost, rather than to display y fully and pay the high marginal tax. Therefore, if the marginal tax rate on y is too high, no one will display y and the marginal gain in the amount of government revenue raised from y will be zero. Crucial here is the level of the concealment cost. The maximal amount of revenue the government can raise is determined by the structure of the unit income concealment cost across all income levels in the population.

An optimal tax system implements the best scheme for income redistribution among all those feasible under the income concealment technology available to the agents. We characterize optimal income tax structures under a flexible specification of the income concealment cost function. Our main result is that progressive income taxes are optimal in our model when the unit cost of income hiding is increasing with true realized income.

This result contrasts the characterizations of optimal marginal income tax rates obtained in the existing literature. Following Mirrlees (1971), virtually all papers in the private-information-based optimal taxation literature study environments in which agents have private information about their individual productivity.[2] In these environments, each agent's income is the product of his skill and effort. While income is publicly observable, individual skill and effort are not. Taxes, therefore, can be a function of the observed income but cannot be conditioned on the unobservable skill or effort. An important feature of optimal taxes obtained by Mirrlees in this private-skill environment is that the optimal income tax schedule is eventually regressive: marginal income tax rates are decreasing for income levels close to the top of the population distribution of income. This feature of the optimal income tax system in private-skill economies has been shown in subsequent studies to be robust to assumptions about the support of the skill distribution, heterogeneity of labor, and general equilibrium effects (see Stiglitz 1987 for a review).

Our main result demonstrates that the prescriptions for optimal income taxation obtained under the Mirrlees approach are very sensitive to the

[2] Varian (1980) and Albanesi (2006) are exceptions. These papers study optimal tax structures in models with moral hazard, i.e., in situations in which agents can take private actions prior to the resolution of the underlying uncertainty. The environment we study in this article is radically different, since in our model, agents can take a private action (i.e., conceal income) after the uncertainty is realized. Our model is an application, as well as an extension, of the costly state falsification (CSF) model of Lacker and Weinberg (1989). In Section 7, we discuss the relationship between our model and the CSF literature.

exogenous specification of economic fundamentals and informational frictions. If the underlying friction is the unobservability of skill and effort, optimal marginal tax rates eventually have to decrease. If the friction is the possibility of hidden income falsification, then increasing marginal income tax rates may be optimal.

This lack of robustness of the theoretical prescriptions obtained in the Mirrlees approach makes apparent that empirical work is needed to determine what are "the right" frictions—the frictions that could be used to derive useful policy recommendations. This question is beyond the scope of this article. However, the optimality of progressive income taxation obtained in our income falsification environment is consistent with the observed progressivity of income tax systems used in many countries, including the United States.

In addition to the main result, we obtain an auxiliary result, which is more generally useful for studying the environments with costly state falsification, i.e., environments in which it is costly to conceal income. This result identifies subadditivity of the concealment cost function as a sufficient condition for the optimality of no-falsification allocations, in which displayed income coincides with true realized income across the whole support of the income distribution.

Slemrod and Yitzhaki (2002) provide an overview of a large existing literature on tax avoidance and evasion. This literature defines tax evasion as criminal tax avoidance. Tax avoidance, in turn, is defined as taking full advantage of legal methods of reducing tax obligations. The literature on tax avoidance is mainly descriptive (see Stiglitz 1985). Virtually all existing theoretical models of tax evasion are built around the costly state verification model of Townsend (1979). In these models, agents can underreport income and the tax authority can perform an audit, i.e., discover, at a cost, the true realized income. The underreported income, if discovered, is taxed at a penalty rate. Most papers in this literature restrict income tax rates or penalty rates, or both, to be linear in income, some take the penalty rates as exogenous.

This article differs from the papers in this literature in two respects. First, we assume that true realized income can never be discovered by the tax authority, and, therefore, never taxed (thus, there are no penalty tax rates in our model). The interpretation of this assumption is consistent with the literature's notion of tax avoidance, rather than evasion. In our model, income hiding is meant to represent all costly but legal actions that agents take to reduce their tax obligations. In reality, these actions involve shifting income across time and tax jurisdictions, transferring the ownership of productive assets, attributing income to tax-exempt sources, etc. All these activities decrease taxable income, and are, usually, costly. In the model, we abstract from the specific nature of these activities. Instead of introducing them in a specific form, we model tax avoidance indirectly by introducing a general income concealment technology similar to the costly state falsification technology of Lacker and Weinberg (1989).

The modeling methodology is the second important difference between this article and the existing literature on taxation constrained by tax avoidance and evasion. As mentioned before, we use the Mirrlees approach, in which resource feasibility and the underlying friction in the environment (private information) are the only source of restrictions on the set of taxes that can be used by the government. To emphasize, in the Mirrlees approach, no exogenous restrictions on the set of available policy instruments are introduced beyond those implied by the fundamentals of the environment. The existing tax evasion literature, in contrast, introduces exogenous restrictions on income and penalty tax rates.[3]

In order to solve a Mirrlees optimal taxation problem, we go through three main steps. First, we provide a complete specification of all economic fundamentals that constitute the model environment. Second, in the specified environment, we characterize the set of most desirable economic outcomes. Third, we obtain a characterization of optimal tax structures by deriving a tax system that implements an optimal outcome in a market equilibrium of this economy.

This article is organized into seven sections in which we go through the three steps of the Mirrlees optimal taxation problem. Sections 1 through 3 provide necessary definitions. In Section 1, a macroeconomic version of the costly state falsification environment is defined. In Section 2, we specify what constitutes an outcome (allocation) and a best outcome (constrained optimal allocation) in this environment. In Section 3, we provide a formal definition of fiscal implementation of an optimal allocation. In Section 4, we characterize and implement the optimum of a benchmark model in which government information is complete. Section 5 is devoted to characterization of the optimal allocation under costly state falsification, that is, with incomplete government information. Our main result is derived in Section 6, in which we study fiscal implementation of the constrained optimum. In Section 7, we discuss the extent to which our results can be generalized with respect to the considered class of income falsification cost functions. We also discuss the relation of our specification of the falsification technology to the specifications considered in the costly state falsification literature. Section 8 concludes the article.

[3] In introducing exogenous restrictions on the set of tax instruments available to the government, most of the existing tax evasion literature follows the so-called Ramsey approach, in which exogenous restrictions on policy instruments (linearity, most commonly) are imposed. Schroyen (1997) studies a tax evasion model with nonlinear income taxes and exogenous penalties.

1. ENVIRONMENT

Consider a single-period economy with a continuum of ex ante identical agents whose preferences are represented by the expected utility function

$$E\left[u(c)\right],$$

where u is twice continuously differentiable with $u' > 0$, $u'' < 0$.

Agents face idiosyncratic income risk. At the beginning of the period, each agent receives individual income $y \in [y_0, y_1]$. The cumulative distribution function of income is F. Given a law of large numbers, $F(y)$ represents both the ex ante probability of an agent's income realization less than or equal to y, and the ex post fraction of agents whose realized income is less than or equal to y. Aggregate income in this economy, denoted by Y, is equal to the expected value of each agent's individual income, i.e.,

$$Y = E[y] = \int_{y_0}^{y_1} y dF(y).$$

Individual realizations of income y are not immediately observable to the public, but, instead, can be, in part or in whole, privatively concealed before income becomes publicly observable. The process of concealment of income is costly: a fraction of each dollar concealed is lost in the process of hiding it from public view. The remaining fraction of each concealed dollar, denoted by $\lambda(y) \in [0, 1]$, however, remains in hidden possession of an agent and is available for consumption. Note that the cost to conceal a dollar of income can vary with the income level.

Given this concealment technology, the amount of hidden (i.e., concealed) consumption available to an agent whose realized income is $y \in [y_0, y_1]$ and who displays to the public the amount $\tilde{y} \leq y$ is given by

$$\int_{\tilde{y}}^{y} \lambda(t) dt.$$

The remaining portion of the concealed income

$$\int_{\tilde{y}}^{y} (1 - \lambda(t)) dt = y - \tilde{y} - \int_{\tilde{y}}^{y} \lambda(t) dt \tag{1}$$

is lost as a deadweight cost of falsification. The unconcealed part of income, $\tilde{y}$, becomes public information and, therefore, is subject to social redistribution (i.e., taxation).

2. CONSTRAINED OPTIMUM DEFINED

Since individual income realizations are stochastic and agents are risk-averse, there are welfare gains to be realized from social insurance. Insurance can be provided by committing ex ante to redistribute ex post some resources from those whose realized income is high to those whose income is low. What is the best possible scheme of income redistribution for providing social insurance in this environment?

In this section, we introduce a standard notion of constrained optimality and define constrained optimal social redistribution mechanisms. These mechanisms are defined as solutions to the so-called social planning problem. This section is focused on defining the social planning problem under the possibility of income falsification by the agents. Our discussion of the solution of this problem is deferred to Section 5.

Mechanisms

The social objective is to choose a set of rules governing all interactions between agents so that the final outcome of these interactions is the best possible. Agents possess private information about their income and can take private action (that is, hide income). The rules to be decided on, therefore, have to prescribe how agents are to communicate their private information, what private action they are supposed to take, and, finally, how resources are to be redistributed among the agents. A complete description of these rules is called a mechanism.

In a general form, a mechanism in our environment involves the following stages of interaction:

1. The mechanism itself is committed to by all parties.
2. Agents receive private information.
3. Communication takes place.
4. Agents take private actions.
5. Redistribution takes place.
6. Agents consume.

The social planning problem is to choose a mechanism that leads to a final allocation of consumption that maximizes the ex ante expected utility of each agent in this economy.[4]

The set of mechanisms that can be used is very large. In particular, since communication is costless in our environment, one can use mechanisms with extensive communication between agents. However, essentially all that needs to be communicated is, at most, the agents' private information about their

[4] As all agents are ex ante identical, the expected utility of the representative agent is a natural choice of the social objective function, which is widely used in macroeconomics. In particular, this objective is consistent with the standard notion of Pareto optimality.

realized income. All other communication is superfluous, that is, cannot lead to a welfare gain. This intuition is formalized in a general result called the Revelation Principle. This result states that when searching for an optimal mechanism, it is enough to search among the so-called direct-revelation incentive-compatible (DRIC) mechanisms.

In a direct-revelation mechanism, all that agents communicate is simply their private information, i.e., in our case, the individual realizations of income. A mechanism is incentive compatible (IC) in our environment if, given a recommendation of private action to be taken and a resource redistribution plan, all agents find it optimal to reveal their information truthfully and follow the recommended course of action. The Revelation Principle states that any final allocation that can be attained with some mechanism can also be attained with a DRIC mechanism. Thus, when one searches for an optimal mechanism, it is enough to look at DRIC mechanisms, which we do hereafter.

To summarize, under a DRIC mechanism, six stages of interactions between agents take place according to the following timeline:

1. Society announces the recommended amount of income hiding, $y - \tilde{y}(y)$, for each actual realization of income $y \in [y_0, y_1]$, and commits to a schedule c for redistribution of displayed income $\tilde{y}$, where, for each $\tilde{y} \in [y_0, y_1]$, $c(\tilde{y})$ denotes the amount of resources publicly assigned to each agent who displays income $\tilde{y}$.
2. Agents receive their individual income realizations y.
3. Agents communicate their realizations of y.
4. Agents follow the action recommended by hiding $y - \tilde{y}(y)$ and making $\tilde{y}(y)$ available to the public.
5. Redistribution of the unconcealed income $\tilde{y}$ occurs according to c.
6. Agents with income y consume
$$c(\tilde{y}(y)) + \int_{\tilde{y}(y)}^{y} \lambda(t)dt, \tag{2}$$
where $\int_{\tilde{y}(y)}^{y} \lambda(t)dt$ represents the hidden (not observed by the public) consumption of the unwasted portion of the concealed income.

Incentive Compatibility

Under the Revelation Principle, the choice of the recommendation schedule $\tilde{y}(y)$ is constrained by the requirement of incentive compatibility. Since both the actual income realized and the concealed fraction of it are private information, it is not possible to determine if agents really hide and display the amounts recommended. Thus, the recommendation has to be consistent with agents' self-interest. In order to precisely describe this requirement, let us introduce the following piece of notation. Given that society is committed to redistributing the unconcealed income according to the allocation c, let $\theta_c(y)$ denote the set of income display levels that maximize utility attained by an agent whose true realized income is y. That is,

$$\theta_c(y) = \arg\max_{\theta\in\Theta(y)} u\left(c(\theta) + \int_\theta^y \lambda(t)dt\right), \tag{3}$$

where $\Theta(y)$ is the set of all income levels that an agent whose actual income is y can feasibly declare as his true income without being discovered.[5] In our environment, for each $y \in [y_0, y_1]$, the set $\Theta(y)$ is given by

$$\Theta(y) = \{\tilde{y} \mid \tilde{y} \leq y, \text{ and } \tilde{y} \in \text{supp}F\}.$$

There are two constraints that determine $\Theta(y)$: the individual resource constraint and the so-called support constraint. Since the unconcealed amount $\tilde{y}$ becomes publicly observable, $\tilde{y}$ cannot be larger than y, which is represented by the individual resource feasibility constraint. The set supp$F$ contains all values of income $y \in [y_0, y_1]$ such that the probability of income realization y is strictly positive under the distribution F. As the distribution F and its support are publicly known, an agent declaring an income realization that is impossible under F, clearly, is lying. This is represented by the support constraint.

A recommended income declaration schedule $\tilde{y} : [y_0, y_1] \rightarrow [y_0, y_1]$ and a consumption redistribution allocation $c : [y_0, y_1] \rightarrow \mathbb{R}$ are (jointly) incentive compatible if

$$\tilde{y}(y) \in \theta_c(y) \tag{4}$$

for all $y \in [y_0, y_1]$.

The requirement of incentive compatibility states that a mechanism cannot give any agent an incentive to deviate from the recommended course of action. The fact that $\theta_c(y)$ is not necessarily a singleton [for some consumption allocations, there will be multiple solutions to the maximization problem on the right-hand side of (3)], is not generally considered a problem. If the recommended action is a selection from $\theta_c(y)$, agents have no reason to deviate. We denote the desired selection from $\theta_c(y)$ by $\tilde{y}_c(y)$. This notation explicitly recognizes the fact that incentive compatibility is a joint requirement on the recommended action $\tilde{y}(y)$ and the consumption allocation schedule c.

Under various particular specifications of the income distribution F, the IC requirement (4) can be written out more explicitly. As an example, consider

---

[5] Note that detectable deviations can be deterred by a commitment to punish them strongly enough so that no agent finds it optimal to use them. The set $\Theta(y)$ describes all undetectable deviations, which cannot be deterred in this simple way.

the case in which $\text{supp}F = \{y_0, y_1\}$ with

$$\begin{aligned} \Pr\{y = y_0\} &= F(y_0), \\ \Pr\{y = y_1\} &= 1 - F(y_0), \end{aligned}$$

where $0 < F(y_0) < 1$. In this case, we have $\Theta(y_1) = \{y_0, y_1\}$ while, due to the individual resource constraint $\tilde{y} \leq y$, $\Theta(y_0) = \{y_0\}$. Agents with the low income realization y_0 have no possibility of hiding income, so no IC constraints are required for them. The IC condition (4) for those with high income y_1 is given by

$$u\left(c(\tilde{y}_c(y_1)) + \int_{\tilde{y}_c(y_1)}^{y_1} \lambda(t)dt\right) \geq u\left(c(\theta) + \int_{\theta}^{y_1} \lambda(t)dt\right)$$

for $\theta \in \Theta(y_1) = \{y_0, y_1\}$. Since the utility function u enters both sides of this constraint symmetrically, the above IC condition for utilities is equivalent to the following condition expressed directly in terms of consumption:

$$c(\tilde{y}_c(y_1)) + \int_{\tilde{y}_c(y_1)}^{y_1} \lambda(t)dt \geq c(\theta) + \int_{\theta}^{y_1} \lambda(t)dt$$

for $\theta \in \Theta(y_1) = \{y_0, y_1\}$. This condition is trivially satisfied for $\theta = \tilde{y}_c(y_1)$, which leaves one IC condition for each possible display recommendation $\tilde{y}_c(y_1)$. In particular, for the recommendation of full display, $\tilde{y}_c(y_1) = y_1$, the IC condition is given by

$$c(y_1) \geq c(y_0) + \int_{y_0}^{y_1} \lambda(t)dt,$$

which simply states that for the full display recommendation to be IC, the publicly assigned consumption $c(y_1)$ must be at least equal to the sum of the publicly assigned consumption $c(y_0)$ and the hidden consumption $\int_{y_0}^{y_1} \lambda(t)dt$ that high-income agents can obtain by hiding the amount $y_1 - y_0$.

In Section 5, we focus on another special case, namely, the full support case in which $\text{supp}F = [y_0, y_1]$. Our general specification of the IC requirement encompasses both of these extreme cases, as well as a variety of intermediate specifications.

Resource Feasibility

Among the incentive-compatible mechanisms, we are interested in those that are self-financing, or resource feasible. A DRIC mechanism is resource feasible if the promised consumption allocation c can be delivered without using any more resources than those displayed by agents. That is, a DRIC mechanism is resource feasible if the following condition is satisfied:

$$\int_{y_0}^{y_1} \left[c\left(\tilde{y}_c(y)\right) - \tilde{y}_c(y)\right] dF(y) \leq 0. \tag{5}$$

For brevity, a direct-revelation incentive-compatible and resource-feasible mechanism will be called an incentive-feasible (IF) mechanism.

The Social Planning Problem

The social planning problem in our environment is to find a welfare-maximizing incentive-feasible mechanism $(\tilde{y}_c, c)$. This social planning problem can be written concisely as the following mathematical programming problem, which will be referred to as problem SPP:

$$\max_{\tilde{y}_c(y), c(\tilde{y}_c)} \int_{y_0}^{y_1} u\left(c\left(\tilde{y}_c(y)\right) + \int_{\tilde{y}_c(y)}^{y} \lambda(t)dt\right) dF(y),$$

subject to

$$\tilde{y}_c(y) \in \arg\max_{\theta \in \Theta(y)} u\left(c(\theta) + \int_{\theta}^{y} \lambda(t)dt\right) \tag{6}$$

for all y, and

$$\int_{y_0}^{y_1} \left[c\left(\tilde{y}_c(y)\right) - \tilde{y}_c(y)\right] dF(y) \leq 0. \tag{7}$$

A constrained optimal mechanism, or just an optimum, is given by a solution to the planning problem SPP. We will use $(\tilde{y}^*, c^*)$ to denote an optimal mechanism.

Remarks

1. Consumption is delivered to agents in two ways: the publicly observable consumption $c\left(\tilde{y}_c(y)\right)$ and the hidden consumption $\int_{\tilde{y}_c(y)}^{y} \lambda(t)dt$. The public consumption allocation c depends on the true realization of income y only through the displayed amount $\tilde{y}_c(y)$. In general, one could allow for c to be a function of both the reported income y and the displayed income $\tilde{y}$. For incentive reasons, however, the direct dependence of c on y has to be trivial. To see this, let c depend on both $\tilde{y}$ and y, and suppose that there exist realizations y^1 and y^2 in $[y_0, y_1]$ such that

$$c(\tilde{y}_c(y^1), y^1) > c(\tilde{y}_c(y^2), y^2).$$

If $\tilde{y}_c(y^1) = \tilde{y}_c(y^2)$, agents with true realized income y^2 will not reveal it truthfully. Instead, they will report y^1 because this report gives them more final (private plus hidden) consumption:

$$c(\tilde{y}_c(y^1), y^1) + \int_{\tilde{y}_c(y^1)}^{y^2} \lambda(t)dt > c(\tilde{y}_c(y^2), y^2) + \int_{\tilde{y}_c(y^1)}^{y^2} \lambda(t)dt.$$

Thus, $\tilde{y}_c(y^1) = \tilde{y}_c(y^2)$ must imply that $c(\tilde{y}_c(y^1), y^1) = c(\tilde{y}_c(y^2), y^2)$ for all y^1 and y^2 in $[y_0, y_1]$, i.e., the direct dependence of c on the revealed income has to be trivial. Intuitively, given a fixed display amount $\tilde{y}$, announcements of y are "cheap talk," which should be ignored.

2. Given the "cheap talk" property of the revealed income y, the directly revealed information about y is not used in consumption assignment c. Thus, the third stage of the general DRIC mechanism, at which agents reveal their actual income y, can be skipped. We see that, in the CSF model, a direct-revelation mechanism does not have to actually call for direct revelation of the realized uncertainty.

3. Since the utility function u is strictly increasing, a recommendation $\tilde{y}_c(y)$ is incentive compatible if and only if it maximizes the consumption of the agent with realized income y. Thus, the function u can be dropped from the objective on the right-hand side of (6), which makes the IC constraint linear in c.

4. The IF mechanisms discussed above operate under the assumption that society can fully commit ex ante to redistributing resources ex post according to the agreed upon plan, c. This assumption is important. In general, for incentive reasons, it is ex ante optimal for c to redistribute displayed resources $\tilde{y}_c(y)$ partially. Ex post, however, agents cannot hide resources that have already been revealed. At this point, if society could reconsider the allocation policy c, it would prefer to redistribute the revealed income more fully (take more from those who reveal a lot). We assume that society can commit ex ante to not reconsidering c ex post. If it could not commit to c, agents would not display income according to $\tilde{y}_c(y)$, and, in effect, less insurance could be implemented.

3. FISCAL IMPLEMENTATION DEFINED

In the Mirrlees approach to the problem of optimal taxation, an optimal tax system is defined as one that obtains optimal allocation of resources as an equilibrium of a market economy with taxes. Having defined optimal allocations in the previous section, we devote this section to defining equilibrium in a market economy with taxes. In our simple environment, in which income is given to agents exogenously, all redistribution of income is done through taxes and thus there is no need for markets. Thus, the general market/tax mechanism can be specialized to a simple tax mechanism, which we define below.

Let us then assume that the task of implementation of the optimal social redistribution policy, along with the power to tax all unconcealed income, is given to a government. The government chooses a tax function $T : [y_0, y_1] \to \mathbb{R}$, where $T(\tilde{y})$ represents the tax levied on agents whose declared income is

$\tilde{y}$. Taxes can be negative, in which case they represent net transfers from the government to the agents.

The timing of events under a tax mechanism is as follows:

1. The government commits to a tax function T.
2. Agents receive their individual income realizations y.
3. Agents hide the amount $y - \tilde{y}(y)$ and display $\tilde{y}(y)$.
4. Redistribution of the unconcealed income $\tilde{y}$ occurs according to T.
5. Agents with income y whose displayed income is $\tilde{y}(y)$ consume

$$\tilde{y}(y) - T(\tilde{y}(y)) + \int_{\tilde{y}(y)}^{y} \lambda(t)dt,$$

where $\tilde{y}(y) - T(\tilde{y}(y))$ is the after-tax unconcealed income, and $\int_{\tilde{y}(y)}^{y} \lambda(t)dt$ is the hidden consumption of the unwasted portion of the concealed income.

Note that, unlike the direct-revelation mechanism used in the social planning problem, the tax mechanism does not specify any recommendation on what portion of realized income y is to be hidden. In a tax mechanism, agents are simply confronted with a tax schedule T. Agents must make the decision on how much income to hide and how much to display, without any explicit recommendation.

To find out what allocation of consumption is implemented by a tax schedule T, which we need to know in order to evaluate welfare attained by T, we need to predict how much income agents of various income levels will conceal at stage 3 of the above mechanism, given that they know T is the tax schedule they will face at stage 4. This can be done by finding, for each T, the set of solutions to the agents' individual utility maximization problem.

This problem is formulated as follows. Agents of income y choose income displayed $\tilde{y} \leq y$, public consumption c_P, and hidden consumption c_H so as to maximize utility

$$u(c_H + c_P),$$

subject to the budget constraint for hidden consumption

$$c_H \leq \int_{\tilde{y}}^{y} \lambda(t)dt,$$

and the after-tax budget constraint for public consumption, which under the tax function T is given by

$$c_P \leq \tilde{y} - T(\tilde{y}).$$

Let us denote by $\theta_T(y)$ the set of individually optimal display levels for agent y under taxes T, and let $\tilde{y}_T(y)$ be any selection from this set. Clearly, since u is increasing, given an individually optimal displayed income level $\tilde{y}_T(y)$, individually optimal hidden and public consumption levels are such that the two budget constraints are satisfied as equalities. Thus,

$$\theta_T(y) = \arg\max_{\tilde{y} \leq y} u\left(\tilde{y} - T(\tilde{y}) + \int_{\tilde{y}}^{y} \lambda(t)dt\right).$$

A tax schedule T implements an optimal mechanism $(\tilde{y}^*, c^*)$ if the following two conditions are met:

$$\tilde{y}^*(y) \in \theta_T(y) \tag{8}$$

for each $y \in [y_0, y_1]$, and

$$c^*(\tilde{y}) = \tilde{y} - T(\tilde{y}) \tag{9}$$

for each $\tilde{y} \in [y_0, y_1]$.

The first condition in the above definition says that the tax schedule T must be such that the socially optimal hiding policy $\tilde{y}^*$ is individually optimal in the tax mechanism under schedule T. Intuitively, this condition is a form of incentive compatibility requirement on the tax system T. The second condition requires that, for each level of displayed income $\tilde{y}$, the transfers prescribed by T exactly replicate the transfers prescribed by the socially optimal redistribution schedule c^*. We will refer to (9) as the replication condition.

The implementation conditions (8) and (9) guarantee that the hidden and public consumption delivered by the tax mechanism T exactly replicate the hidden and public consumption of the optimal DRIC mechanism $(\tilde{y}^*, c^*)$ for each $y \in [y_0, y_1]$. Therefore, welfare attained by the tax mechanism T is equal to the maximal welfare attainable in this environment. For this reason, a tax system that implements an optimum is called an optimal tax system.

Also, transfers implemented by an optimal T are budget feasible for the government. Since an optimal mechanism $(\tilde{y}^*, c^*)$ is resource feasible, we have that

$$\begin{aligned} 0 &\leq -\int_{y_0}^{y_1} \left[c\left(\tilde{y}^*(y)\right) - \tilde{y}^*(y)\right] dF(y) \\ &= -\int_{y_0}^{y_1} \left[\tilde{y}^*(y) - T\left(\tilde{y}^*(y)\right) - \tilde{y}^*(y)\right] dF(y) \\ &= \int_{y_0}^{y_1} T(\tilde{y}_T(y)) dF(y), \end{aligned}$$

which means that net tax revenue is nonnegative under T. The inequality above follows from (7), i.e., the fact that $(\tilde{y}^*, c^*)$ is resource feasible. The first equality follows from the implementation condition (9), and the second from the implementation condition (8).

In this article, we are interested in a characterization of a tax system T that is optimal in the environment defined in Section 2, in which hiding of income is costly.

4. SOLVING THE FULL INFORMATION BENCHMARK CASE

Before we proceed to the optimal taxation problem with income hiding, we describe in this short section the solution to the optimal taxation problem in an environment in which income hiding is not possible. This case serves as a benchmark against which we can compare optimal allocations and tax systems obtained in environments with income falsification.

When income cannot be hidden, we can think of it as being public information as soon as it is realized. Thus, there is no private information to be communicated, nor is there any private action to be taken. The only object that needs to be specified by a mechanism in the full information case is the allocation of consumption $c(y)$ for each realized income level $y \in [y_0, y_1]$. Resource feasibility is the sole constraint that the consumption allocation c has to satisfy. Namely,

$$\int_{y_0}^{y_1} c(y)dF(y) \leq Y.$$

What allocation of consumption is optimal under full information? Clearly, it is the full-insurance allocation under which all income risk is insured and thus all agents' consumption is the same, i.e.,

$$c(y) = c^{FI}$$

for all $y \in [y_0, y_1]$. Why? Any allocation with unequal consumption can be improved upon, since all agents have the same preferences with marginal utility decreasing in consumption. It is socially beneficial to redistribute a unit of consumption from those who have more to those who have less because the utility gain to the poorer caused by such a transfer is larger than the utility loss to the richer and, hence, the total social welfare is increased. Under full information, such a transfer, as self-financing, is feasible. Thus, the optimal redistribution scheme is to allocate consumption equally to all.

What is the maximal level c^{FI} of the same-for-all consumption that can be attained? The resource feasibility constraint implies that

$$Y = \int_{y_0}^{y_1} c^{FI}dF(y) = c^{FI}, \tag{10}$$

that is, each agent's consumption equals per capita income.

How can this allocation be implemented with a tax system? Since, obviously, there is no hidden consumption in the public information case, each agent's final consumption is simply equal to the consumption of publicly assigned resources. Therefore, the only condition for implementation is the condition

$$c^{FI} = y - T(y)$$

for all $y \in [y_0, y_1]$. Using (10), we conclude that the tax system

$$T(y) = y - Y$$

is optimal in the full information benchmark.

In the full information benchmark case, the optimal marginal tax rate is 100 percent. Agents with the realized income y_0 pay a tax of $y_0 - Y$, which is a negative number, i.e., they receive a transfer. As realized income increases in the population, the size of the transfer from the government to the agents decreases 1 to 1 with income. The agents whose realized income is exactly equal to the average income Y pay zero. All income above the average level Y is taxed out. The implemented distribution of consumption is uniform: all agents consume Y.

5. CHARACTERIZING CONSTRAINED OPTIMAL ALLOCATIONS

As the first step toward a characterization of optimal taxes in the class of environment with private income and private action, we characterize in this section optimal allocations of those environments.

We start out by noting that the full information optimum cannot be achieved in the private information case when the cost of falsification is less than 100 percent. For the full information optimum to be implementable, it must be the case that

$$c^{FI} \geq c^{FI} + \int_{\theta}^{y} \lambda(t)dt$$

for all y and all $\theta \in \Theta(y)$, which, given that λ is nonnegative, is true only if $\lambda(t) = 0$ for all $t \in \operatorname{supp} F$.

Intuitively, the full-insurance allocation c^{FI} does not give agents any incentive to display income, as consumption publicly assigned to agents is independent of income they display. If λ is not identically equal to zero, agents can benefit from hiding income. Since low declared income does not cause any loss of publicly assigned consumption, all agents will display the lowest possible income realization, i.e., y_0. The promise of $c^{FI} = Y$ for each agent will be impossible to fund with the total displayed income of $y_0 < Y$. This makes the full information optimum infeasible outside of the trivial case in which $\lambda(y) = 0$ for all y.

In contrast, the no-redistribution allocation $c(\tilde{y}) = \tilde{y}$ can always be implemented with the recommendation for all agents to display all income. Thus, we see how the need to provide incentives puts a limit on the amount of redistribution (i.e., social insurance) that can be implemented when income falsification is possible.

What is the maximal amount of social insurance that can be provided when agents can falsify income? To answer this question, we need to solve the social planning problem SPP.

A No-Falsification Theorem

Problem SPP (defined on page 86) is not very convenient to work with, since it involves the recommendation $\tilde{y}_c$, which depends on the allocation c. Each time we want to evaluate welfare generated by a candidate redistribution policy c, we need to specify an incentive-compatible display recommendation $\tilde{y}_c$. Optimization, therefore, takes place jointly over the choice of $c(y)$ and $\tilde{y}_c(y)$ for all $y \in [y_0, y_1]$. The social planning problem would be much simpler if we could fix a display recommendation and search only over the consumption allocations c. It turns out that, in the class of economies we consider, such a simplification is possible.

Suppose that we confine attention to mechanisms that recommend full display of income for all realizations of $y \in [y_0, y_1]$. We call such mechanisms no-falsification (NF) mechanisms. Below, we prove a result that states that when searching for an optimal mechanism, confining attention to NF mechanisms is without loss of generality. The Revelation Principle implies that limiting attention to incentive-feasible mechanisms is without loss of welfare. We show something stronger: it is without loss of welfare to confine attention to those IF mechanisms that are NF mechanisms.

This result, which we call a no-falsification theorem, significantly simplifies the social planning problem SPP. It implies that the recommendation $\tilde{y}_c$ can be taken to be $\tilde{y}_c(y) = y$ for all y, independently of the candidate allocation c. This greatly reduces the dimensionality of the social planning problem, as now optimization is only over the allocation functions c.

Formally, we will say that an incentive-feasible mechanism $(\tilde{y}_c, c)$ is a no-falsification mechanism if

$$\tilde{y}_c(y) = y$$

for all $y \in [y_0, y_1]$.

An incentive-feasible, no-falsification (IFNF) mechanism can be expressed simply as an allocation function c, with $\tilde{y}_c$ implicitly specified as $\tilde{y}_c(y) = y$ for all y.

The main result of this section is the following:

Theorem. *For any IF mechanism $(\tilde{y}_c, c)$, there exists an IFNF mechanism $\hat{c}$ that delivers the same social welfare as $(\tilde{y}_c, c)$.*

Proof. Let $(\tilde{y}_c, c)$ be an IF and resource-feasible mechanism. Define an IFNF mechanism $\hat{c}$ as follows:

$$\hat{c}(y) = c\left(\tilde{y}_c(y)\right) + \int_{\tilde{y}_c(y)}^{y} \lambda(t)dt.$$

We first show that $\hat{c}$ is incentive compatible. Suppose it is not. Then, the incentive compatibility constraint (6) has to be violated at $\hat{c}$, which means that there exist y and $z \in \Theta(y)$ such that

$$\hat{c}(y) < \hat{c}(z) + \int_z^y \lambda(t)dt.$$

Substituting for $\hat{c}(y)$ and $\hat{c}(z)$ from the definition of $\hat{c}$, the above is equivalent to

$$\begin{aligned} c\left(\tilde{y}_c(y)\right) + \int_{\tilde{y}_c(y)}^{y} \lambda(t)dt &< c\left(\tilde{y}_c(z)\right) + \int_{\tilde{y}_c(z)}^{z} \lambda(t)dt + \int_z^y \lambda(t)dt \\ &= c\left(\tilde{y}_c(z)\right) + \int_{\tilde{y}_c(z)}^{y} \lambda(t)dt, \end{aligned} \tag{11}$$

where the equality follows from the additivity of the definite integral with respect to the limits of integration. Denoting $\tilde{y}_c(z)$ by x, we rewrite the above inequality as

$$c\left(\tilde{y}_c(y)\right) + \int_{\tilde{y}_c(y)}^{y} \lambda(t)dt < c\left(x\right) + \int_x^y \lambda(t)dt.$$

Note now that $x \in \Theta(y)$. Indeed,

$$x = \tilde{y}_c(z) \in \operatorname{supp} F,$$

because $x = \tilde{y}_c(z) \in \Theta(z)$, and

$$x \leq z \leq y, \tag{12}$$

because $z \in \Theta(y)$ and $x \in \Theta(z)$. But this contradicts the incentive compatibility of the mechanism $(\tilde{y}_c, c)$, because x is a feasible display level for an agent with realized income y that provides strictly more consumption (i.e., also utility) than the recommended display level $\tilde{y}_c(y)$. This contradiction implies that $\hat{c}$ is incentive compatible.

Welfare generated by $\hat{c}$ is equal to welfare generated by $(\tilde{y}_c, c)$, as, by definition of $\hat{c}$, both mechanisms deliver the same consumption to agents of all income levels. Note that $\hat{c}$ delivers publicly the same consumption that $(\tilde{y}_c, c)$ delivers as a sum of public and hidden consumption for each realization of income y.

It remains to be shown that $\hat{c}$ is resource feasible. The resources needed to deliver $\hat{c}$ are

$$\begin{aligned}
\int_{y_0}^{y_1} \hat{c}\,(y)\; dF(y) &= \int_{y_0}^{y_1} \left[c\,(\tilde{y}_c(y)) + \int_{\tilde{y}_c(y)}^{y} \lambda(t)dt \right] dF(y) \\
&= \int_{y_0}^{y_1} \left[c\,(\tilde{y}_c(y)) - \tilde{y}_c(y) + \tilde{y}_c(y) + \int_{\tilde{y}_c(y)}^{y} \lambda(t)dt \right] dF(y) \\
&= \int_{y_0}^{y_1} \left[c\,(\tilde{y}_c(y)) - \tilde{y}_c(y) \right] dF(y) \\
&\quad + \int_{y_0}^{y_1} \left[\tilde{y}_c(y) + \int_{\tilde{y}_c(y)}^{y} \lambda(t)dt \right] dF(y) \\
&\leq \int_{y_0}^{y_1} \left[\tilde{y}_c(y) + \int_{\tilde{y}_c(y)}^{y} \lambda(t)dt \right] dF(y) \qquad (13) \\
&\leq \int_{y_0}^{y_1} \left[\tilde{y}_c(y) + y - \tilde{y}_c(y) \right] dF(y) \qquad (14) \\
&= Y,
\end{aligned}$$

where (13) follows from (5), that is, the fact that $(\tilde{y}_c, c)$ is resource feasible, and (14) from the fact that

$$\lambda(t) \leq 1$$

for all t. Since $\hat{c}$ is an incentive-compatible, no-falsification mechanism, agents display all income truthfully. Thus, since the amount of resources available for redistribution under $\hat{c}$ is

$$\int_{y_0}^{y_1} y\, dF(y) = Y,$$

$\hat{c}$ is resource feasible and the proof is complete.

Remarks

1. Note in the last step of the preceding proof that, in a large class of environments, inequality (14) is strict. When $\tilde{y}_c(y) < y$ for some y such that $\lambda(y) < 1$, under the mechanisms $(\tilde{y}_c(y), c)$, agents engage in a wasteful activity of hiding income. Under the NF mechanism $\hat{c}$, this inefficiency is eliminated. Therefore, whenever $\lambda < 1$, NF mechanisms are not merely as good as falsification mechanisms, but strictly better.

2. A key step in showing the incentive compatibility of the NF mechanism $\hat{c}$ is the equality in (11). This equality holds true because the cost of hiding the amount $y - x$ is equal to the sum of costs of hiding $y - z$ and $z - x$. The proof of our no-falsification result would not go through if the cost of hiding $y - x$ were strictly larger than the sum of costs of hiding $y - z$ first and $z - x$ next. In Section 7, we discuss an example of

such an environment. There, also, we discuss how our no-falsification theorem is related to the results of Lacker and Weinberg (1989).

3. Another important step in the proof involves showing that $\tilde{y}_c(z) \in \Theta(y)$. This holds because in our environment it is possible to hide the whole income. Suppose that there is an upper bound on the proportion of income that can be hidden. Say only 20 percent of actually realized income can be hidden. With this bound in place, it may be impossible to display $\tilde{y}_c(z)$ when true income is y because, despite being a feasible display, for the true income z, $\tilde{y}_c(z)$ may be less than 80 percent of y, which means that it is not a feasible display for the true income y. Clearly, this will be the case when $0.8z \leq \tilde{y}_c(z) < 0.8y$. Thus, under such a partial concealment technology, our no-falsification theorem fails. This point follows from an insight of Green and Laffont (1986).

Simplifying the Social Planning Problem

By our no-falsification theorem, we hereafter confine attention to NF mechanisms without loss of generality. The incentive compatibility constraint (4) of an NF mechanism

$$y \in \theta_c(y),$$

for all $y \in [y_0, y_1]$ can be equivalently written as

$$c(y) \geq c(\theta) + \int_{\theta}^{y} \lambda(t)dt \tag{15}$$

for all $y \in [y_0, y_1]$ and all $\theta \in \Theta(y)$.

The resource feasibility constraint (5) under an NF mechanism simplifies to

$$\int_{y_0}^{y_1} c\,(y)\,dF(y) \leq Y. \tag{16}$$

Under an NF mechanism, all consumption is public (as no resources are hidden). Welfare attained by an IFNF mechanism c, therefore, is given simply by

$$\int_{y_0}^{y_1} u\,(c\,(y))\,dF(y). \tag{17}$$

The social planning problem SPP restricted to the class of no-falsification mechanisms, thus, is to find a schedule $c(y)$ so as to maximize social welfare (17) subject to incentive compatibility (15) and resource feasibility (16).

This formulation of the social planning problem is much simpler (as no function $\tilde{y}_c$ is involved). It will be useful, however, to simplify it even further.

Simplifying the IC Constraints

When supp$F$ contains many points, the number of constraints in the condition (15) is large, as incentive compatibility of c needs to be checked for all y in supp$F$ and all $\tilde{y}$ in supp$F$ below y. This is true, in particular, for the case of full support, that is, if

$$\text{supp}F = [y_0, y_1].$$

In this section, we show how, in the full support case, incentive compatibility conditions (15) can be equivalently expressed with a smaller number of so-called local IC constraints. Replacing the global conditions (15) with the local constraints defined below does not alter the requirement of incentive compatibility, but the social planning problem is simpler to handle when local constraints are used.

Define the local IC constraints as

$$dc(y) \geq \lambda(y)dy \tag{18}$$

for all $y \in [y_0, y_1]$. The notation $dc(y)$ stands for the change in c when y is changed infinitesimally (similar to the notation $dF(y)$ we have already used to denote integration with respect to differences in the distribution function F). If c is differentiable, the above condition reduces to

$$c'(y) \geq \lambda(y)$$

for all $y \in [y_0, y_1]$.

Intuitively, the local IC constraints prevent agents from hiding small amounts of output. Take an agent whose realized income is y. The recommended display under an NF mechanism c is to hide nothing. The agent considers a small deviation from no-falsification, which means hiding a small amount of income, dy. The private benefit of doing so comes in the form of a small amount $\lambda(y)dy$ of resources available to the agent for hidden consumption. The local IC constraint (18) requires that the loss in publicly assigned consumption resulting from this underreporting, $dc(y)$, be large enough to at least offset the agent's gain in hidden consumption.

We show now that, if F has full support, the global IC constraints (15) and local IC constraints (18) are equivalent.[6]

If c satisfies the global constraints (15), it must also satisfy the local constraints (18). The global IC constraint (15) at y with display level θ is given by

$$c(y) - c(\theta) \geq \int_{\theta}^{y} \lambda(t)dt.$$

Taking the limit $\theta \to y$, we obtain (18) at y.

[6] In order to avoid technical detail, the argument is presented quite informally.

The local constraints, in turn, guarantee the incentive compatibility of allocation c in the global sense. To see this, fix an arbitrary y and $\theta \leq y$, both in $[y_0, y_1]$. The local IC constraints imply that for all $t \in [\theta, y]$ we have

$$0 \leq dc(t) - \lambda(t)dx.$$

By the positivity of the operation of integration, we, thus, have

$$\begin{aligned} 0 &\leq \int_\theta^y dc(t) - \int_\theta^y \lambda(t)dt \\ &= c(y) - c(\theta) - \int_\theta^y \lambda(t)dt, \end{aligned}$$

which shows that the global IC constraint is satisfied for y and θ. Since the choice of y and θ was arbitrary, the same is true for all y and $\theta \leq y$ in $[y_0, y_1]$ and, thus, all IC constraints (15) are satisfied.

Having shown that local IC constraints are necessary and sufficient for incentive compatibility of an IFNF mechanism c, we can express the social planning problem simply as follows: find an allocation c that maximizes social welfare in the class of all allocations that are resource feasible and locally incentive compatible. The reduction of the original planning problem SPP to this form is going to pay off now in that the solution to the reduced-form problem will be easy to find.

Solving the Social Planning Problem

Intuitively, the local IC constraints (18) put a lower bound on how flat the distribution of consumption can be. At the full-insurance allocation, consumption distribution c^{FI} is completely flat:

$$dc^{FI}(y) = 0.$$

If this distribution cannot be achieved, due to $\lambda(y) > 0$, the best distribution that can be implemented is the one that comes as close to c^{FI} as possible. Thus, intuition says that the best among all IC allocations should be the one at which the slope of $c(y)$ is as small as possible at all levels of y. Given the lower bound imposed by the local IC constraints, this means that the slope of c at y should be equal to $\lambda(y)$, for all y.

This intuition is correct as can be seen from the following argument. Suppose to the contrary there exists an optimal allocation c such that

$$dc(y') > \lambda(y')dy' \tag{19}$$

for some $y' \in [y_0, y_1]$. Consider also an alternative allocation $\bar{c}$, which is identical to c except in a small neighborhood of y', where $\bar{c}$ prescribes a little more redistribution than c. More income redistribution at y' means that $\bar{c}$ grows more slowly in the neighborhood of y' than does c, that is, $d\bar{c}(y') < dc(y')$. With a sufficiently small increase in the amount redistributed, however, the differential $d\bar{c}(y')$ can be made arbitrarily close to $dc(y')$. In particular, given that at c the local incentive constraint around y' is slack, the increase in the amount redistributed can be made sufficiently small so as to have

$$d\bar{c}(y') \geq \lambda(y')dy',$$

which means that $\bar{c}$ is incentive compatible. As under any NF allocation, agents hide no income under $\bar{c}$, so the amount available for redistribution under $\bar{c}$ is Y. Since $\bar{c}$ uses the same amount of resources as c, Y is sufficient to fund the total consumption promised by $\bar{c}$, so $\bar{c}$ is resource feasible. Also, as $\bar{c}$ provides marginally more consumption to agents with higher marginal utility, it generates higher social welfare than c. This contradicts the optimality of c.

The above argument implies that any optimal allocation, denoted by c^*, must satisfy

$$dc^*(y) = \lambda(y)dy \tag{20}$$

for all $y \in [y_0, y_1]$, i.e., all local IC constraints are binding at a solution to social planning problem.

Note now that the binding local IC constraints pin down the optimal allocation up to a constant. Integrating (20) we get

$$\int_{y_0}^{y} dc^*(t) = \int_{y_0}^{y} \lambda(t)dt,$$

that is,

$$c^*(y) = c^*(y_0) + \int_{y_0}^{y} \lambda(t)dt.$$

This formula tells us a lot about the structure of optimal allocation of consumption. It is optimal to assign to an agent with realized income y only as much consumption as he could guarantee himself by declaring the lowest income realization, y_0. The incentive to display y fully is delivered by publicly giving the agent exactly as much as what he could get by hiding $y - y_0$. The amount of this "incentive payment" is equal to $\int_{y_0}^{y} \lambda(t)dt$.

The constant $c^*(y_0)$ can be obtained from the resource feasibility constraint:

$$\begin{aligned}
Y &= \int_{y_0}^{y_1} c^*(y)dF(y) \\
&= \int_{y_0}^{y_1} [c^*(y_0) + \int_{y_0}^{y} \lambda(t)dt]dF(y) \\
&= c^*(y_0) + \int_{y_0}^{y_1} \int_{y_0}^{y} \lambda(t)dtdF(y) \\
&= c^*(y_0) + \int_{y_0}^{y_1} (1 - F(t))\lambda(t)dt,
\end{aligned}$$

which implies that

$$c^*(y_0) = Y - \int_{y_0}^{y_1} (1 - F(t))\lambda(t)dt. \tag{21}$$

The optimal amount of consumption assigned to an agent at the very bottom of income distribution is equal to what it would be in the full-insurance case ($c^{FI} = Y$), less the average incentive payment made to other agents whose income exceeds the low realization y_0.

Since the constant $c^*(y_0)$ is uniquely determined in (21), the optimal allocation c^* is uniquely pinned down as

$$c^*(y) = Y - \int_{y_0}^{y_1} (1 - F(t))\lambda(t)dt + \int_{y_0}^{y} \lambda(t)dt \tag{22}$$

for all $y \in [y_0, y_1]$. Consumption assigned to agents with income y is equal to the average income, minus the average population incentive payment, plus the incentive payment specific to agents of income y.

As an example, consider the special case in which the cost of hiding income is independent of income, i.e., take $\lambda(y) = \lambda$ for all y. In this case, we get

$$\begin{aligned} c^*(y_0) &= Y - \lambda \int_{y_0}^{y_1} (1 - F(t))dt \\ &= Y - \lambda(Y - y_0) \\ &= \lambda y_0 + (1 - \lambda)Y, \end{aligned}$$

and

$$\begin{aligned} c^*(y) &= \lambda y_0 + (1 - \lambda)Y + \lambda(y - y_0) \\ &= \lambda y + (1 - \lambda)Y. \end{aligned}$$

The optimal assignment of consumption, in this case, does not depend on the income distribution F. Consumption assigned to agents with income y is a weighted average of their income y and the average income Y, where the weight assigned to the average income is equal to the per-dollar income falsification cost $1 - \lambda$. In particular, when this cost is 100 percent, the full-insurance allocation $c^{FI} = Y$ is implementable. If this cost is zero, no social insurance can be implemented, and the no-redistribution allocation $c(y) = y$ is optimal.

6. FISCAL IMPLEMENTATION OF THE CONSTRAINED OPTIMUM

The no-falsification theorem does not only help solve the social planning problem, but also makes fiscal implementation of the optimum straightforward.

In order to implement an IF mechanism $(\tilde{y}_c, c)$, we need to find an income tax schedule $T\colon [y_0, y_1] \to \mathbb{R}_+$ that satisfies two implementation conditions:

the incentive compatibility condition (8) and the transfer replication condition (9). If the mechanism to be implemented is a no-falsification mechanism, however, the incentive compatibility condition follows from the transfer replication condition, and thus only one simple condition has to be checked.

Therefore, a tax schedule T implements the optimal IFNF mechanism c^*, if and only if

$$c^*(y) = y - T(y)$$

for all $y \in [y_0, y_1]$.

This condition uniquely pins down the optimal tax schedule, which will be denoted by T^*. Substituting for $c^*(y)$ from (22), we get

$$T^*(y) = y - \int_{y_0}^{y} \lambda(t)dt - Y + \int_{y_0}^{y_1} (1 - F(t))\lambda(t)dt$$

for all $y \in [y_0, y_1]$.

As we see, the structure of the optimal tax system T^* is determined by the unit income falsification cost function $1 - \lambda$. Optimal marginal income taxes are given by

$$dT^*(y) = (1 - \lambda(y))dy.$$

At all points of continuity of λ, we, thus, have

$$\frac{d}{dy}T^*(y) = 1 - \lambda(y),$$

that is, the optimal marginal income tax rate applied to the income level y is equal to the per-dollar income falsification cost at y.

Since our model does not put any restrictions on the shape of the function λ, a large class of tax schedules T is consistent with optimality under some specification of λ. In particular, if the unit cost of income falsification $1-\lambda(y)$ is increasing in y, progressive taxation of income is optimal in our model. Clearly, it is easy to provide a specification for the function λ that generates an optimal tax system that is piecewise-linear, similar to the income tax schedule currently used in the United States.

What, in conclusion, does our model suggest about why we observe progressive taxation in many countries, including the United States? Our model shows that, if the cost of falsification is increasing in income, it is optimal to tax higher income at a higher rate because in this way, the maximal amount of desirable social insurance can be provided without pushing people into wasteful tax avoidance activities. In this sense, our model provides a possible explanation for the observed progressivity of income taxes.

7. SOME EXTENSIONS AND ALTERNATIVE SPECIFICATIONS

The no-falsification property is a key feature of the optimal mechanism for the provision of social insurance in the class of environments we have considered so far. In this section, we study the extent to which our no-falsification theorem can be extended to environments with more general falsification cost technologies.

The class of falsification cost functions that we considered so far consists of all functions that can be expressed as the definite integral (1). We have demonstrated that a useful no-falsification theorem holds for all such cost functions. The proof of this theorem uses the additivity property of the definite integral. It turns out, however, that this proof goes through under a weaker condition of subadditivity of the falsification cost function. Therefore, the no-falsification result extends to a larger class of environments than merely those in which the falsification cost function can be expressed as an integral of the form given in (1).

We identify subadditivity of the falsification cost function as a key condition for the no-falsification result as well as for the optimality of no-falsification mechanisms. In the first subsection, we show that subadditivity is sufficient for the no-falsification result, which implies that no-falsification mechanisms are optimal whenever the falsification cost function is subadditive. In the second subsection, we show that no-falsification mechanisms are not optimal in general. We give an example of a falsification technology under which all no-falsification mechanisms are welfare-dominated by a mechanism that uses falsification.

In the third and final subsection, we discuss the relation between our model and the costly state falsification literature.

A Generalized No-Falsification Theorem

Our no-falsification theorem can be extended to any subadditive cost function $\psi : D \to \mathbb{R}_+$, where

$$D = \left\{(y, x) \in [y_0, y_1]^2 \mid x \leq y\right\},$$

and where subadditive means that for all $x \leq z \leq y$, $x \geq y_0$, $y \leq y_1$, we have

$$\psi(y, x) \leq \psi(y, z) + \psi(z, x).$$

In fact, under this more general specification of the falsification cost function, our proof goes through without change. In particular, for any IF mechanism $(\tilde{y}_c, c)$ we define the no-falsification mechanism $\hat{c}$ as

$$\hat{c}(y) = c\left(\tilde{y}_c(y)\right) + y - \tilde{y}_c(y) - \psi(y, \tilde{y}_c(y)).$$

As we work step-by-step through the original proof, it follows that $\hat{c}$ is always at least as good as $(\tilde{y}_c, c)$ for any subadditive cost function ψ.

The class of subadditive cost functions contains many flexible specifications. Therefore, by our no-falsification theorem, the class of environments in which the NF mechanisms are optimal is fairly large.

Is subadditivity of the cost function ψ necessary for the no-falsification result? When the cost function ψ is not subadditive, as mentioned already in Remark 2 on page 94, our proof of the no-falsification theorem does not go through because from the supposition that the NF mechanism $\hat{c}$ is not IC, it no longer follows that the mechanism $(\tilde{y}_c, c)$ is not IC. To see this, note that the fact that there exists $z \in \Theta(y)$, such that

$$\hat{c}(y) < \hat{c}(z) + y - z - \psi(y, z)$$

implies that

$$\begin{aligned} c\left(\tilde{y}_c(y)\right) + y - \tilde{y}_c(y) - \psi(y, \tilde{y}_c(y)) &< c\left(\tilde{y}_c(z)\right) + z - \tilde{y}_c(z) \\ &\quad -\psi(z, \tilde{y}_c(z)) + y - z - \psi(y, z) \\ &= c\left(\tilde{y}_c(z)\right) + y - \tilde{y}_c(z) \\ &\quad - \left[\psi(y, z) + \psi(z, \tilde{y}_c(z))\right] \end{aligned}$$

but does not, in general, imply that

$$c\left(\tilde{y}_c(y)\right) + y - \tilde{y}_c(y) - \psi(y, \tilde{y}_c(y)) < c\left(\tilde{y}_c(z)\right) + y - \tilde{y}_c(z) - \psi(y, \tilde{y}_c(z)).$$

This last implication fails when

$$\psi(y, \tilde{y}_c(z)) > \psi(y, z) + \psi(z, \tilde{y}_c(z)),$$

that is, when the cost of two piecemeal falsifications is smaller than the cost of making the same falsification in one big step.

We see that when the falsification cost function is not subadditive, there are IC allocations of final (private plus hidden) consumption that cannot be achieved with an NF mechanism. This, in itself, does not imply that NF mechanisms are sub-optimal. It is possible that the allocations that are not implementable without falsification are welfare dominated by allocations that can be implemented in an NF mechanism. In the next subsection, however, by means of an example, we show that, in general, this is not true. In fact, under some income-hiding cost functions, mechanisms that prescribe falsification are optimal.

Optimality of Falsification Mechanisms

In this subsection, we specify a particular falsification cost function and derive the best NF mechanism. Then we provide an example of a falsification mechanism that welfare dominates the best NF mechanism in this environment.

Consider the following falsification cost function:

$$\psi(y, x) = \max\{y - x - \delta, 0\} \tag{23}$$

for all $(y, x) \in D$. Under this specification, the first δ dollars of income can be hidden costlessly, while the resource cost of hiding anything in excess of δ is 100 percent. Clearly, this cost function is not subadditive.

What is the best no-falsification mechanism under this cost function? We see that an allocation c is consistent with no-falsification if and only if

$$dc(y) \geq dy \tag{24}$$

for all y. Indeed, if $dc(y) < dy$, agents can benefit from hiding up to δ dollars of income because their hidden consumption increases one-to-one with every dollar hidden while their public consumption decreases at a slower rate for falsifications smaller than δ. Clearly, if $dc(y) \geq dy$, then no agent benefits from hiding income and, thus, no-falsification is incentive compatible. Also, it is clear that among all allocations satisfying (24), the one at which all constraints (24) bind, provides the most insurance and, hence, the highest ex ante social welfare among all NF mechanisms. Thus, the best NF mechanism, denoted as c^{NF}, satisfies

$$dc^{NF}(y) = dy$$

for all y. Integrating, we get

$$c^{NF}(y) - c^{NF}(y_0) = y - y_0$$

for all y. Resource feasibility implies that

$$c^{NF}(y_0) = y_0.$$

Under the falsification cost (24), therefore, the best NF mechanism coincides with the no-insurance allocation

$$c^{NF}(y) = y$$

for all y. Intuitively, since small falsifications are costless to agents at all income levels, full display of income is incentive compatible only when there is no redistribution (taxation) of the displayed income, which means that no insurance of the individual income risk is possible.

Now consider the following falsification mechanism $(\tilde{y}_{\bar{c}}(y), \bar{c})$:

$$\begin{aligned} \tilde{y}_{\bar{c}}(y) &= \max\{y - \delta, y_0\}, \\ \bar{c}(y) &= \bar{c} \end{aligned}$$

for all y. In this mechanism, the recommendation function $\tilde{y}_{\bar{c}}(y)$ says that agents should hide δ and display $y - \delta$, or, if $y - \delta < y_0$, agents should display the lowest income realization y_0 and hide $y - y_0$. The redistribution function $\bar{c}$ simply assigns a constant amount of resources to all agents, regardless of their displayed income.

It is easy to see that this mechanism is IC. First, no one has an incentive to hide less than the recommended amount, because the public consumption allocation $\bar{c}(y) = \bar{c}$ does not reward agents who display larger income. Second, hiding more than δ for agents with income $y > y_0 + \delta$ yields no additional hidden consumption because the marginal cost of hiding is 1 for all income hidden in excess of δ. Finally, hiding more than $y - y_0$ when $y < y_0 + \delta$ violates the support condition $\tilde{y} \in \text{supp}F$. Thus, $(\tilde{y}_{\bar{c}}(y), \bar{c})$ is IC.

The mechanism $(\tilde{y}_{\bar{c}}(y), \bar{c})$ is also resource feasible if we set

$$\begin{aligned} \bar{c} &= \int_{y_0}^{y_1} \tilde{y}_{\bar{c}}(y) dF(y) \\ &= y_0 F(y_0 + \delta) + \int_{y_0+\delta}^{y_1} (y - \delta) dF(y). \end{aligned}$$

With this choice of $\bar{c}$, the mechanism $(\tilde{y}_{\bar{c}}(y), \bar{c})$ is incentive feasible. Assuming $\delta < y_1 - y_0$, that is, that not all income in excess of y_0 can be hidden at zero cost, we have

$$\bar{c} > y_0.$$

Under this falsification mechanism, the final consumption $\bar{c}^{ph}(y)$ provided to an agent with income y,

$$\bar{c}^{ph}(y) = \begin{cases} y + \bar{c} - y_0 & \text{if} \quad y < y_0 + \delta, \\ \bar{c} + \delta & \text{if} \quad y \geq y_0 + \delta \end{cases} \tag{25}$$

is the sum of the public consumption $\bar{c}$ and the hidden consumption $y - \max\{y - \delta, y_0\}$.

Clearly, the best NF mechanism $c^{NF}(y) = y$ and the mechanism $(\tilde{y}_{\bar{c}}(y), \bar{c})$ do not provide the same allocation of final consumption, and the consumption profile $\bar{c}^{ph}(y)$ cannot be replicated by an NF mechanism. It is not immediately clear, however, that the falsification mechanism $(\tilde{y}_{\bar{c}}(y), \bar{c})$ welfare-dominates the best NF mechanism $c^{NF}(y) = y$, as agents at the top of the distribution of realized income are worse off under $(\tilde{y}_{\bar{c}}(y), \bar{c})$, relative to the no-insurance allocation $c^{NF}(y) = y$. The following argument shows that the best NF mechanism is in fact suboptimal.

Denote by $G(c)$ the cumulative distribution function of the distribution of final consumption $\bar{c}^{ph}$ provided by the mechanism $(\tilde{y}_{\bar{c}}(y), \bar{c})$. That is,

$$G(c) = \Pr\left\{y : \bar{c}^{ph}(y) \leq c\right\} = F(\bar{c}^{ph}(c)).$$

Using (25), the formula for G can be explicitly written out as

$$G(c) = \begin{cases} 0 & \text{if} \quad c < \bar{c}, \\ F(c - \bar{c}) & \text{if} \quad c \in [\bar{c}, \bar{c} + \delta), \\ 1 & \text{if} \quad c > \bar{c} + \delta. \end{cases} \tag{26}$$

The cumulative distribution function of consumption provided by the no insurance mechanism $c^{NF}(y) = y$ is simply given by F. In this notation, the best NF mechanism is welfare dominated by $(\tilde{y}_{\bar{c}}(y), \bar{c})$ if and only if

$$\int_{y_0}^{y_1} u(c) dF(c) < \int_{y_0}^{y_1} u(c) dG(c). \tag{27}$$

Given that all income hiding that takes place under $(\tilde{y}_{\bar{c}}(y), \bar{c})$ is costless (i.e., no resources are wasted in the process of falsification), both consumption allocations use the same amount of resources

$$\int_{y_0}^{y_1} c\,dG(c) = \int_{y_0}^{y_1} c\,dF(c) = Y, \tag{28}$$

which means that G and F are two distributions with the same mean value, Y. Thus, given that u is strictly concave, the welfare domination condition (27) is literally equivalent to the second-order stochastic domination of distribution G over distribution F.[7] It is a standard result (see, for example, Mas-Colell, Whinston, and Green 1995) that G second-order stochastically dominates F if and only if

$$\int_{y_0}^{c} [F(t) - G(t)]\,dt \geq 0 \tag{29}$$

for all $c \in [y_0, y_1]$. We now show that this condition is satisfied.

From (26), we get that the difference $F(t) - G(t)$ is positive for $t \leq \bar{c} + \delta$ and then negative for $t > \bar{c} + \delta$. The integral on the left-hand side of (29) is, therefore, first increasing and then decreasing. Integrating (28) by parts we get

$$\int_{y_0}^{y_1} [F(t) - G(t)]\,dt = 0.$$

Also, naturally, we have

$$\int_{y_0}^{y_0} [F(t) - G(t)]\,dt = 0.$$

These end-point conditions and the fact that the integral on the left-hand side of (29) is first increasing and then decreasing imply that the integral on the left-hand side of (29) is everywhere positive. Thus, (29) is satisfied for all $c \in [y_0, y_1]$, and G does second-order stochastically dominate F.

Intuitively, the falsification mechanism $(\tilde{y}_{\bar{c}}(y), \bar{c})$ dominates the no-insurance allocation $c^{NF}(y) = y$ because it manages to provide some social insurance. At $(\tilde{y}_{\bar{c}}(y), \bar{c})$, consumption provided to those with the lowest income y_0 is larger than $c^{NF}(y_0) = y_0$, as $\bar{c} > y_0$. Also, the consumption profile $\bar{c}^{ph}(y)$ is everywhere at least weakly flatter than the no-insurance consumption profile $c^{NF}(y) = y$, but not flatter than the full-insurance profile, at which $c(y)$ is constant. We see then that $(\tilde{y}_{\bar{c}}(y), \bar{c})$ delivers a consumption profile intermediate between the no-insurance and full-insurance allocations.

[7] By definition, distribution G second-order stochastically dominates distribution F if, under any strictly concave utility function, G delivers larger expected utility than F, which is exactly what our condition (27) requires.

Thus, $(\tilde{y}_{\bar{c}}(y), \bar{c})$ welfare-dominates the no-insurance allocation, that is, the best allocation among all attainable with an NF mechanism.

Relation to the CSF Literature

In the original paper introducing the costly state falsification (CSF) model, Lacker and Weinberg (1989) (hereafter LW) study a class of falsification cost functions ψ in which the cost of falsification depends only on the amount hidden. In particular, conditional on the amount hidden, the falsification cost does not depend on the actual income realization y. More precisely, the class of falsification cost functions considered in LW consists of such falsification cost functions ψ for which there exists a function $g : \mathbb{R} \to \mathbb{R}_+$ with $g(0) = 0$ such that

$$\psi(y, x) = g(y - x) \tag{30}$$

for all $x \leq y$, $x \geq y_0$, $y \leq y_1$.

Following LW, a number of papers in economics and finance have used the CSF model in a variety of applications. These include managerial incentives and asset pricing (Lacker, Levy, and Weinberg 1990), optimal insurance contract design (Crocker and Morgan 1998), managerial compensation (Crocker and Slemrod 2005, forthcoming), investor protection law and growth (Castro, Clementi, and MacDonald 2004), and optimal dynamic capital structure of the firm (DeMarzo and Sannikov 2006). All of these papers consider the LW specification of falsification cost function (30).

This article differs from these papers in two respects. First, this article is, to our knowledge, the first to apply the CSF model to the problem of optimal redistributive taxation. Second, the class of integral falsification cost functions that we consider is different from the LW class, which means that this article studies a version of the CSF model that has not been previously studied in the literature.

In the remainder of this subsection, we discuss the relationship between the LW class of falsification cost functions and the class of cost functions we study in this article. The class considered in this article consists of all functions ψ that admit the integral representation (1), i.e., such functions ψ for which there exists a function $\lambda : [y_0, y_1] \to [0, 1]$ such that

$$\psi(y, x) = \int_x^y (1 - \lambda(t))dt,$$

for all $x \leq y$, $x \geq y_0$, $y \leq y_1$.

Neither the LW class nor our class of falsification cost functions is more general than the other. Clearly, the integral cost function representation we consider is not a special case of the LW specification, as in our model the cost of hiding a fixed amount of income can depend on the realized income level, y. The LW specification is not a special case of the integral representation, either. A key property of the integral representation is additivity. The LW specification encompasses nonadditive cost functions, for example, the nonadditive cost function $\psi(y, x) = \max\{y - x - \delta, 0\}$ considered in the previous subsection admits the LW representation with $g(h) = \max\{h - \delta, 0\}$, where $h = y - x$ is the amount hidden.

These two classes of cost functions are not disjoint, for example, the constant per dollar cost function belongs to both of them. Clearly, if in the integral representation λ is constant, then

$$\int_x^y (1 - \lambda(t))dt = (1 - \lambda) \int_x^y dt = (1 - \lambda)(y - x) = g(x - y),$$

where $g(h) = (1 - \lambda)h$. Also, there are cost functions ψ that do not belong to either of the two classes. An example is the function

$$\psi(y, x) = \max\{y - x - \delta(y), 0\},$$

with $\delta(y)$ a nonconstant function of y.

8. CONCLUSION

In this article we follow the Mirrlees approach to the question of optimal income taxation. This question is studied in an environment in which agents can avoid taxes by concealing income. The structure of the optimal income tax schedule is determined by the properties of the income concealment technology. The main result obtained shows that, if the cost of concealment is increasing with income, it is optimal to tax higher income at a higher marginal rate because, in this way, the maximal amount of desirable social insurance can be provided without pushing people into wasteful tax avoidance activities. In this sense, our model provides a possible explanation for the progressivity of income taxes that we observe in many countries, including the United States.

As an auxiliary result, we prove a no-falsification theorem for the class of CSF environments in which the concealment technology is characterized by subadditivity of the concealment cost function. We demonstrate that, in this class of environments, it is without loss of generality to restrict attention to mechanisms that recommend full display of all realized income for agents of all income levels. This result can be useful more generally, that is, in different applications of the CSF model.

Several possible lines of extension of our model are worth mentioning. First, in contrast to the Mirrlees environments, the realized (pre-concealment) income is exogenous in our model. In particular, pre-concealment income does not respond to taxation. In a richer environment, the falsification effect that we study in this article would be only one of several forces shaping optimal tax structures. Second, the class of income falsification technologies considered in the model is large, which allows for a large variety of tax structures to be consistent with optimality under some concealment technology. Grounding the model more fundamentally in technology could provide sharper predictions about the structure of optimal taxes. Third, and related, falsification technology is taken as exogenous in the model. In particular, it cannot be affected by the government. The results obtained could change if the scope of tax avoidance activities available to the agents is explicitly modeled as dependent on government policy.

REFERENCES

Albanesi, S. 2006. "Optimal Taxation of Entrepreneurial Capital Under Private Information." NBER Working Paper No. 12212.

Albanesi, S., and C. Sleet. 2006. "Dynamic Optimal Taxation with Private Information." *Review of Economic Studies* 73 (1): 1–30.

Castro, R., G. L. Clementi, and G. MacDonald. 2004. "Investor Protection, Optimal Incentives, and Economic Growth." *Quarterly Journal of Economics* 119 (3): 1,131–75.

Crocker, K. J., and J. Morgan. 1998. "Is Honesty the Best Policy? Curtailing Insurance Fraud through Optimal Incentive Contracts." *Journal of Political Economy* 106 (2): 355–75.

Crocker, K. J., and J. B. Slemrod. 2005. "Corporate Tax Evasion with Agency Costs." *Journal of Public Economics* 89 (9–10): 1,593–610.

Crocker, K. J., and J. B. Slemrod. Forthcoming. "The Economics of Earnings Manipulation and Managerial Compensation." *RAND Journal of Economics*.

DeMarzo, P., and Y. Sannikov. 2006. "Optimal Security Design and Dynamic Capital Structure in a Continuous-Time Agency Model." *Journal of Finance* 61 (6): 2,681–2,724.

Green, J. R., and J. J. Laffont. 1986. "Partially Verifiable Information and Mechanism Design." *Review of Economic Studies* 53 (3): 447–56.

Kocherlakota, N. R. 2005. "Zero Expected Wealth Taxes: A Mirrlees Approach to Dynamic Optimal Taxation." *Econometrica* 73 (5): 1,587–621.

Kocherlakota, N. R. Forthcoming. "Advances in Dynamic Optimal Taxation." *Advances in Economics and Econometrics, Theory and Applications: Ninth World Congress of the Econometric Society*.

Lacker, J. M., and J. A. Weinberg. 1989. "Optimal Contracts under Costly State Falsification." *Journal of Political Economy* 97 (6): 1,345–63.

Lacker, J. M., R. J. Levy, and J. A. Weinberg. 1990. "Incentive Compatible Financial Contracts, Asset Prices, and the Value of Control." *Journal of Financial Intermediation* 1 (1): 31–56.

Mas-Colell, A., M. D. Whinston, and J. R. Green. 1995. *Microeconomic Theory*. New York: Oxford University Press.

Mirrlees, J. A. 1971. "An Exploration in the Theory of Optimum Income Taxation." *Review of Economic Studies* 38 (114): 175–208.

Schroyen, F. 1997. "Pareto Efficient Income Taxation Under Costly Monitoring." *Journal of Public Economics* 65 (3): 343–66.

Slemrod, J., and S. Yitzhaki. 2002. "Tax Avoidance, Evasion, and Administration." In *Handbook of Public Economics*, vol. 3, eds. A. J. Auerbach and M. Feldstein. Amsterdam: North-Holland.

Stiglitz, J. E. 1985. "The General Theory of Tax Avoidance." *National Tax Journal* 38 (3): 325–38.

Stiglitz, J. E. 1987. "Pareto Efficient and Optimal Taxation and the New Welfare Economics." In *Handbook of Public Economics,* vol. 2*,* eds. A. J. Auerbach and M. Feldstein. Amsterdam: North-Holland.

Townsend, R. M. 1979. "Optimal Contracts and Competitive Markets with Costly State Verification." *Journal of Economic Theory* 21 (2): 265–93.

Varian, H. R. 1980. "Redistributive Taxation as Social Insurance." *Journal of Public Economics* 14 (1): 49–68.