

Optimal Taxation with Non-Filers & Imperfect Takeup

Dylan T. Moore

University of Hawai'i Economic Research Organization (UHERO)

Department of Economics

University of Hawai'i at Mānoa

Preliminary Draft

March 28, 2025

Abstract

This paper revisits classic results in optimal income taxation by incorporating non-filers, whose undermining the efficacy of income tax-based redistribution. Under Atkinson-Stiglitz preference assumption, the addition of non-filers rationalizes commodity subsidies as an alternative approach to achieving redistribution. The imperfection of this approach may also rationalize differential commodity taxation/subsidization. Moreover, when filing is an endogenous choice—affected by demogrant incentives—the optimal income tax follows a modified ABC rule, with social marginal utility adjusted to reflect imperfect takeup. These findings reveal that accounting for non-filers can increase or decrease optimal redistribution.

Imperfect takeup of governments transfers—in part due to non-filing of income tax returns—creates a gap between the ground reality of redistribution and the theoretical discussions of redistribution in the optimal tax literature. This paper helps to fill that gap by exploring several extensions of the theory of optimal nonlinear income taxation that account for the possibility that some agents will not receive the transfers to which they are entitled.

I first consider a simple model of joint commodity and income taxation, in a setting where some agents are non-filers. I characterize the jointly optimal choice of a nonlinear tax and multiple linear commodity tax rates, showing that imperfect takeup of income tax-based transfers provides a rationale for commodity subsidization as an alternative to income tax-based redistribution. I further show that even under Atkinson-Stiglitz-type preferences, non-filing may even create a rationale for differential commodity taxation/subsidization, depending on how heterogeneous the non-filer population is.

Next, I consider a simple model of nonlinear income taxation where taxpayers with negative tax liability do not always claim their transfer income, and claiming decisions are endogenous. I show that the optimal nonlinear tax schedule is characterized by a simple modified ABC formula where the social marginal utility of income accounts for both the imperfection and endogeneity of takeup. This model can imply either more or less redistribution will place at the optimum than in the standard model, depending on the pattern of takeup, how takeup responds to the generosity of the transfer, and whether takeup costs are welfare-relevant, introducing a normative ambiguity dimension to the analysis (Reck & Goldin, 2022).

This is a preliminary draft of the paper. In ongoing work I am exploring richer, more realistic extensions of the models discussed here. I am also expanding the results to consider under what circumstances non-filing/imperfect takeup may rationalize predistributive policies such as minimum wages.

1 Exogenous Non-Filing, Subsidies, & Rate Differentiation

The economy consists of two sets of agents: filers and non-filers. Non-filers have a type indexed by $\theta \in \Theta$, and filers have a type indexed by $w \in \mathcal{W}$. The cumulative distribution of filers types is F_w and non-filer types is F_Θ . Let $\mu \in (0, 1)$ be the share of non-filers in the economy.

Filers choose how much income to earn, and how much after-tax income to spend on a set of consumption goods. Non-filers have a fixed budget to spend on the same set of consumption goods. For simplicity of exposition, we will discuss the two-good case.

Filers Let $z(w) \geq 0$ be the taxable income a type w filer selects and let $x_1(w) \geq 0$ and $x_2(w) \geq 0$ be their consumption of the two commodities. The set of choices for a type w filer is, $(z(w), x_1(w), x_2(w))$, the solution to

$$\max_{z, x_1, x_2} \{u(\phi(x_1, x_2), z; w) : (1 + t_1)x_1 + (1 + t_2)x_2 = z - T(z) + n(w)\}, \quad (1)$$

where $T(\cdot)$ is a nonlinear income tax schedule, t_i is the commodity tax rate applied to good $i \in \{1, 2\}$, and $n(w)$ is some untaxed exogenous nonlabor income.

A type w filer's problem can be framed as a two-stage problem. In the second stage, the agent chooses commodity consumption conditional on their choice of after-tax income I . Let $\Phi(I)$ represent the indirect

utility the agent gets from this second stage problem for a given value of after-tax income I :

$$\Phi(I) \equiv \max_{x_1, x_2} \{ \phi(x_1, x_2) : (1+t_1)x_1 + (1+t_2)x_2 = I \}. \quad (2)$$

In the first stage, the a type w filer chooses a value of taxable income $z(w)$ which maximizes their utility, anticipating how this choice impacts the value of the third stage problem. Let $V^F(w)$ represent the indirect utility a type w filer gets from this first stage problem:

$$V^F(w) \equiv \max_z \{ u(\Phi(z - T(z) + n(w)), z; w) \}. \quad (3)$$

Non-filers Let $x_1^N(\theta) \geq 0$ and $x_2^N(\theta) \geq 0$ be the consumption of the two commodities for a non-filer of type θ . All non-filers are assumed to earn no income, but rather finance their consumption using some untaxed nonlabor income $n^N(\theta)$. In this section, we shall assume $(x_1^N(w), x_2^N(w))$, the solution to

$$\max_{x_1, x_2} \{ \phi(x_1, x_2) : (1+t_1)x_1 + (1+t_2)x_2 = n^N(\theta) \}. \quad (4)$$

Notice, because the preferences of non-filers have been assumed to be identical to those of filers, we have that the indirect utility of a type θ nonfiler is

$$V^N(\theta) \equiv \Phi(n^N(\theta)).$$

Planner's Problem The social planner's problem is to select $T(\cdot)$, t_1 , and t_2 to maximize the weighted social welfare function

$$(1-\mu) \int \gamma^F(w) V^F(w) dF_w(w) + \mu \int \gamma^N(\theta) V^N(\theta) dF_\theta(\theta)$$

subject to a budget constraint

$$(1-\mu) \left[\int T(z(w)) dF_w(w) + t_1 \bar{x}_1^F + t_2 \bar{x}_2^F \right] + \mu [t_1 \bar{x}_1^N + t_2 \bar{x}_2^N] = 0$$

where average demand for good i by filers is

$$\bar{x}_i^F \equiv \int x_i(w) dF_w(w),$$

and average demand for good i by non-filers is

$$\bar{x}_i^N \equiv \int x_i^N(\theta) dF_\theta(\theta).$$

1.1 Optimal Tax System with Homogeneous Non-labor Income

First, consider the case where $n(w) = n^N(\theta) = 0$. In this scenario, the social planner cannot do anything to affect non-filers. All non filers have no taxable income, receive no transfers, and don't buy any taxed commodities. Hence, the optimal tax system is really the solution to a modified planner's problem which excludes the non-filers' welfare from the objective function. Further note that in this special case, filer preferences satisfy the requirements of the Atkinson-Stiglitz theorem, hence our optimal tax system will consist of a standard ABC formula to characterize the optimal nonlinear income tax schedule, and a zero commodity tax rate.¹

1.1.1 Positive Nonlabor Income

Next, consider the case where $n(w) = n^N(\theta) = n_0 > 0$. In this scenario, the social planner can improve non-filer welfare by subsidizing consumption. All non-filers have an identical income, which they use to buy the taxed commodities. If we subsidize those commodities, we can improve non-filer welfare.

Note that any type w filer such that $z(w) = 0$, will select the same bundle of commodities as a non-filer. This is because both groups have identical preferences over commodities, and identical incomes. Given this, the requirements of the Atkinson-Stiglitz theorem remain satisfied, so we know that optimal commodity taxes will be uniform.

Let us further assume that any such filers achieve the same level utility as non-filers ($V^F(w) = V^N(\theta)$) and that welfare weight assigned to these groups are the same ($\gamma^F(w) = \gamma^N(\theta)$). This is a natural assumption: why should the planner prefer one group of non-working individuals over another if both have identical incomes, simply because one group files a tax return?

Under this assumption, we can find the optimal tax system in two steps. First, set $t_1 = t_2 = 0$, and solve for the optimal schedule under the counterfactual assumption that all the non-filers are filers with types w such that they will always select to earn no income ($z(w) = 0$). This "optimal" schedule, $\tilde{T}(\cdot)$, follows the familiar ABC formula, and will implies some specific value for the "optimal" size of the demogrant: $\tilde{T}(0) < 0$.

¹In fact, the planner is indifferent between this tax system and other tax systems with a uniform commodity tax, with suitably modified income tax rates.

In the second step, we'll adjust the tax schedule $\tilde{T}(\cdot)$ to implement an identical budget constraint for tax filers under a uniform commodity subsidy. Let t be the optimal uniform commodity tax rate and $T(\cdot)$ be the optimal nonlinear tax schedule. Note that a tax filer's budget constraint in this optimal system is

$$(1+t)x_1 + (1+t)x_2 = n_0 + z - T(z).$$

As Feldstein (1999) notes, this is equivalent to the budget constraint that can be obtained under a zero commodity tax with a slightly adjusted tax schedule

$$\begin{aligned} x_1 + x_2 &= \frac{n_0 + z - T(z)}{1+t} \\ &= z + n_0 - \underbrace{\frac{T(z) + t(z + n_0)}{1+t}}_{\text{alternative tax schedule}} \end{aligned}$$

The optimal nonlinear tax schedule $T(\cdot)$ and uniform commodity tax rate t can be obtained by simply taking the tax schedule from step one, $\tilde{T}(\cdot)$, and asking what alternative tax system satisfies

$$\tilde{T}(z) = \frac{T(z) + t(z + n_0)}{1+t},$$

ensuring all filers face equivalent budget constraints to the original schedule, and $T(0) = 0$, ensuring that a commodity tax is used as the instrument of redistribution rather than a demogrant. Taken together, these requirements imply that

$$T(z) = (1+t)\tilde{T}(z) - t(z + n_0)$$

and

$$t = \frac{\tilde{T}(0)}{n_0 - \tilde{T}(0)} < 0.$$

This new system accomplishes what the step 1 system was trying to do, but could not because of non-filers do not receive the demogrant. The optimal tax system is thus simply a feasible way of implementing the planner's preferred allocation, given the constraints introduced by the non-filers.

While this model is very simple, it captures a critical aspect of optimal taxation with non-filers. In the standard model, under Atkinson-Stiglitz preferences, the planner is indifferent between a range of possible tax systems which combine uniform commodity taxation with a nonlinear income tax, but which all implement the same allocation. But in the model with non-filers, these different tax systems are no longer equivalent, and the planner has a preference for a specific nonlinear tax schedule and a specific uniform commodity

subsidy.

This result has parallels in prior literature and in optimal tax folk wisdom. Informally, the Atkinson-Stiglitz theorem has often been taken to imply that a broad-based commodity tax is good policy, on the basis of enforcement/administrative considerations. A very different justification for resolving the planner’s indifference amongst these systems can be found in Moore (2024) who discusses implications of tourist consumption for optimal taxation in a destination economy. The result of this section offers an alternative resolution, with very different policy implications: broad-based subsidies. To obtain comprehensive policy implications thus might require a model which incorporates both non-filers and these other considerations simultaneously. This is beyond the scope of this paper.

1.2 Optimal Tax System with Heterogeneous Non-labor Income

Let us assume that non-filer non-labor income is heterogeneous, while filer non-labor income is fixed at some value $n^F(w) = n_0$, so that filer preference still satisfy the assumptions that would otherwise result in Atkinson-Stiglitz theorem holding. Heterogeneity of non-filer non-labor income is sufficient to imply that a differentiated commodity tax/subsidy system is preferable to a uniform commodity tax subsidy when one of the two goods is a “necessity” or “luxury” good amongst the non-filers, such that differential commodity taxation allows the planner to better target redistribution to higher marginal utility non-filers (i.e. to put the subsidy on the good disproportionately consumed by the lower income non-filers).

To see why, consider tax reforms that jointly change the tax system in a way that is distribution neutral for filers, in the spirit of ?. Suppose that the economy begins with the tax system $(t_1, t_2, T(\cdot))$. Now, consider a joint tax reform which marginally increases tax rate on good i by $dt_i > 0$ while also marginally changing income tax liability at every income level z by some amount $dT(z)$. The effect of this joint reform on the value of the filers’ second stage problem (equation ??) is:

$$d\Phi(z - T(z) + n_0) = \frac{\partial\Phi(z - T(z) + n_0)}{\partial I} [x_i(z) dt_i + dT(z)]$$

Notice, by setting $dT(z) = -x_i(z) dt_i$ for all z , we obtain a joint reform for which $d\Phi(z) = 0$ for all z .

This type of joint reform is distribution neutral from the perspective of filers, as the income tax change is designed to compensate every filer for the loss in purchasing power caused by the increase in the commodity tax rate on good i . Consequently, when we evaluate the first-order welfare effects of this joint reform, we do not need to account for effects on the private utility of residents, but only for effects on revenue obtained from

residents. Further, note that because the reform leaves the value the filers' second stage problem unchanged at every value of z , it also leaves the taxable income choices of filers unchanged.

For non-filers, the first-order welfare of the joint reform is just the standard effects of a change to a linear commodity tax rate. Thus, the planner's FOC for this joint reform is

$$-\mu \left[\int \gamma^N(\theta) \frac{\partial \Phi(\theta)}{\partial n} x_i^N(\theta) dF_\theta(\theta) \right] + \lambda \left. \frac{dR}{dt_i} \right|_{dT(z)=-x_i(z)dt_i} = 0$$

where the revenue effect is

$$\begin{aligned} \left. \frac{dR}{dt_i} \right|_{dT(z)=-x_i(z)dt_i} &= \underbrace{\mu \bar{x}_i^N + \mu \int \left[t_i \left. \frac{\partial x_i^N(\theta)}{\partial t_i} \right|_u + t_{-i} \left. \frac{\partial x_{-i}^N(\theta)}{\partial t_i} \right|_u \right] dF_\theta(\theta)}_{\text{nonfiler substitution effects}} \\ &\quad - \underbrace{\mu \int x_i^N(\theta) \left[t_1 \frac{\partial x_1^N(\theta)}{\partial I} + t_2 \frac{\partial x_2^N(\theta)}{\partial I} \right] dF_\theta(\theta)}_{\text{nonfiler income effects}} \\ &\quad + \underbrace{(1-\mu) \int \left[t_i \left. \frac{\partial x_i(w)}{\partial t_i} \right|_{u,z=z(w)} + t_{-i} \left. \frac{\partial x_{-i}(w)}{\partial t_i} \right|_{u,z=z(w)} \right] dF_w(w)}_{\text{filer substitution effects}} \end{aligned} \quad (5)$$

Applying standard properties of Hicksian demand functions to the filers' compensated responses, we can simplify this to:²

$$\begin{aligned} \left. \frac{dR}{dt_i} \right|_{dT(z)=-x_i(z)dt_i} &= \mu \bar{x}_i^N - \underbrace{\mu \left[\frac{t_i}{1+t_i} - \frac{t_{-i}}{1+t_{-i}} \right] \int x_i^N(\theta) \varepsilon_i^N(\theta) dF_\theta(\theta)}_{\text{nonfiler substitution effects}} \\ &\quad - \underbrace{\mu \int x_i^N(\theta) \alpha(\theta) dF_\theta(\theta)}_{\text{nonfiler income effects}} \\ &\quad - \underbrace{(1-\mu) \left[\frac{t_i}{1+t_i} - \frac{t_{-i}}{1+t_{-i}} \right] \int x_i(w) \varepsilon_{i|z}(w) dF_w(w)}_{\text{filer substitution effects}} \end{aligned}$$

²Note that

$$(1+t_i) \left. \frac{\partial x_i(w)}{\partial t_i} \right|_{u,z} + (1+t_{-i}) \left. \frac{\partial x_{-i}(w)}{\partial t_{-i}} \right|_{u,z} = 0.$$

Combined with Slutsky symmetry, this implies that

$$\left. \frac{\partial x_{-i}(w)}{\partial t_i} \right|_{u,z} = -\frac{1+t_i}{1+t_{-i}} \left. \frac{\partial x_i(w)}{\partial t_i} \right|_{u,z}. \quad (6)$$

Similarly, non-filer compensated responses satisfy

$$\left. \frac{\partial x_{-i}^N(\theta)}{\partial t_i} \right|_u = -\frac{1+t_i}{1+t_{-i}} \left. \frac{\partial x_i^N(\theta)}{\partial t_i} \right|_u.$$

where the compensated elasticity of demand for good i is

$$\varepsilon_{i|z}(w) \equiv -\frac{1+t_i}{x_i(w)} \frac{\partial x_i(w)}{\partial t_i} \Big|_{u,z}$$

for filers, is

$$\varepsilon_i^N(\theta) \equiv -\frac{1+t_i}{x_i^N(\theta)} \frac{\partial x_i^N(\theta)}{\partial t_i} \Big|_u$$

for nonfilers, and where the income effect for nonfilers is

$$\alpha(\theta) \equiv t_1 \frac{\partial x_1^N(\theta)}{\partial I} + t_2 \frac{\partial x_2^N(\theta)}{\partial I}.$$

The planner's FOC can thus be simplified to

$$1 - \bar{g} - \text{Cov}\left(\frac{x_i^N(\theta)}{\bar{x}_i^N}, g(\theta)\right) = \left[\frac{t_i}{1+t_i} - \frac{t_{-i}}{1+t_{-i}}\right] \left(\mathcal{E}_i^N + \frac{1-\mu}{\mu} \frac{\bar{x}_i^F}{\bar{x}_i^N} \mathcal{E}_{i|z}^F\right) \quad (7)$$

where the social marginal utility of income for nonfilers is

$$g(\theta) \equiv \frac{\gamma^N(\theta) \frac{\partial \Phi(\theta)}{\partial n}}{\lambda} + \alpha(\theta),$$

aggregate compensated demand elasticities for good i are

$$\mathcal{E}_i^N \equiv \frac{\int x_i^N(\theta) \varepsilon_i^N(\theta) dF_\theta(\theta)}{\bar{x}_i^N}$$

and

$$\mathcal{E}_{i|z}^F \equiv \frac{\int x_i(w) \varepsilon_{i|z}(w) dF_w(w)}{\bar{x}_i^F}.$$

In addition, note that the average social welfare weight for nonfilers is $\bar{g} = \int g(\theta) dF_\theta(\theta)$ and the average nonfiler consumption of good i is $\bar{x}_i^N \equiv \int x_i^N(\theta) dF_\theta(\theta)$.

Considering equation 7 for all $i \in \{1, 2\}$, we can obtain an intuitive condition for the optimal level of commodity tax differentiation

$$\frac{t_1}{1+t_1} - \frac{t_2}{1+t_2} = \frac{-\text{Cov}\left(g(\theta), \frac{x_1^N(\theta)}{\bar{x}_1^N} - \frac{x_2^N(\theta)}{\bar{x}_2^N}\right)}{\mathcal{E}_1^N + \mathcal{E}_2^N + \frac{1-\mu}{\mu} \left(\frac{\bar{x}_1^F}{\bar{x}_1^N} \mathcal{E}_{1|z}^F + \frac{\bar{x}_2^F}{\bar{x}_2^N} \mathcal{E}_{2|z}^F\right)}. \quad (8)$$

This condition has the intuitive implication that the tax rate on good 1 should be higher if—amongst nonfilers—”relatively high consumption” of good 1 is negatively correlated with social marginal utility of

income.

Note, the definition of “relatively high consumption” is somewhat subtle here. The question is not whether households with higher social welfare weights spend a lower share of their income on good 1, but rather whether they tend to have a lower value of $\frac{x_1^N(\theta)}{\bar{x}_1^N} - \frac{x_2^N(\theta)}{\bar{x}_2^N}$: the difference between normalized consumption of good 1 and 2 (relative to the average consumption of each good). Importantly, variation in consumption patterns amongst the filers plays no role in equation 8, and properties of filer demand do not determine whether tax rate differentiation is optimal, or the direction of optimal differentiation. Filer demand only plays a role in the optimal magnitude of tax rate differentiation.

Ultimately, the result is quite straightforward. Amongst filers, all redistribution can be conducted using the nonlinear income tax. Redistribution from filers to nonfilers can be conducted using the same type of tax structure discussed in the homogeneous nonlabor income case: a uniform commodity subsidy. But when non-labor income is heterogeneous amongst non-filers, the planner will also want to redistribute amongst nonfilers. Commodity tax differentiation is the only way they can accomplish this.

However, the equity benefits of commodity tax differentiation must be weighed against the efficiency costs that differentiation creates. The denominator on the righthand side of 8 measures the magnitude of these costs. The magnitude of optimal differentiation declines the greater are these costs. This attenuation is also increasing in the ratio of filer to nonfiler average demand for each good— $\frac{\bar{x}_1^F}{\bar{x}_1^N}$ and $\frac{\bar{x}_2^F}{\bar{x}_2^N}$ —and decreasing in the nonfiler share μ . This happens because the equity benefits of differentiation are exclusive to the nonfiler population, so the larger the relative size of filer consumption, the more weight is placed on the efficiency costs of differentiation stemming from filers.

2 Endogenous Filing & Optimal Redistribution

Let us consider a different simple model from the one discussed in section 2. Here we assume:

- (i) the planner cannot levy commodity taxes, only an income tax;
- (ii) there exists a set of “potential non-filers”, only some of whom fail to file their taxes;
- (iii) the share of “potential non-filers” who file is increasing in the incentive to file;
- (ii) and, all “potential non-filers” earn zero income.

In this model, the incentive to file taxes stems from the receipt of a demogrant, G . Potential non-filers who file their taxes so not change their labor supply, continuing to earn nothing, but will have higher welfare those who do not file because they receive a demogrant.

Let the indirect utility of type w filers be

$$V^F(w) \equiv \max_z \{u(z - T(z) + G, z; w)\},$$

let the indirect utility of type θ potential non-filers who file be

$$V^{N,F}(\theta) \equiv V^N(G; \theta),$$

and let the indirect utility of type θ potential non-filers who do not file be

$$V^{N,N}(\theta) \equiv V^N(0; \theta) + \rho\xi$$

where ξ is the cost of filing, and $\rho \in [0, 1]$ is a parameter which determines whether the cost of filing is *normatively relevant* to the planner's calculus. Varying ρ will allow us to consider the differing implications of assuming non-filing is a rational response to the cost of filing, or of assuming it is a mistake of the agent.

Regardless of our normative assumptions about the cost of filing, we assume that it does influence agent behavior. We assume that each type θ potential non-filers draws a cost of filing ξ from some cost distribution and only files their tax return if

$$V^N(G; \theta) - V^N(0; \theta) > \xi.$$

We then let $\pi(G; \theta)$ denote the probability a type θ potential non-filers files their tax return. If we assume that $V^N(G; \theta)$ is strictly increasing in G , and this implies that $\pi(G; \theta)$ is increasing in G .

The social planner's problem in this case is to select a non-linear income tax schedule $T(\cdot)$ to maximize the weighted social welfare function

$$(1 - \mu) \int \gamma^F(w) V^F(w) dF_w(w) \\ + \mu \int \gamma^N(\theta) [\pi(G; \theta) V^{N,F}(\theta) + (1 - \pi(G; \theta)) V^{N,N}(\theta)] dF_\theta(\theta)$$

subject to a budget constraint

$$(1 - \mu) \int T(z(w)) dF_w(w) = G \left[1 - \mu + \mu \int \pi(G; \theta) dF_\theta(\theta) \right].$$

Consider what happens in this model as the share of non-filers (μ) rises. The impact on the optimal size of the demogrant, G , is ambiguous. On the one hand, as non-filers become a more important, this may lead the planner to favor a higher demogrant, in order to encourage more (potential) non-filers to file their taxes. That is, the demogrant is useful not only to increase the welfare of current filers, but also to extend redistribution to non-filers by getting them to file. On the other hand, this also makes redistributing to low income taxpayers more costly than it otherwise would be, since in some cases our choices are now between giving them nothing, or giving them a larger amount than might otherwise be optimal.

The key novel sufficient statistics for determining the optimal level of redistribution in this model are the responsiveness of non-filing to the demogrant, $\pi(G; \theta)$, across different types of potential non-filers. Recall, non-filers may differ in their non-labor income. If endogenous filing responses are higher amongst non-filers with relatively lower non-labor income, this favors a more redistributive optimal schedule. If the opposite is true, this will attenuate the optimal level of redistribution. Thus, ideally we would want to have information about not only responsiveness of filing to the incentive to file, but also how that responsiveness varies with consumption or wealth (some measure of differences in baseline welfare across households).