

Taxing Financial Transactions

Bruno Biais (HEC)

and

Jean-Charles Rochet (University of Geneva and MIT)

April 2020

Preliminary and incomplete.

Abstract

We examine the capacity of a Financial Transaction Tax (FTT) to generate substantial fiscal revenue without provoking too many distortions. In a model where investors' wealth is imperfectly observable, we show that a FTT is always part of the optimal tax mix under one condition: it must be that richer people are more inclined to trade on financial markets than poorer people. Under this assumption, the financial transactions volume of an individual investor gives a second signal (after capital income) about unobserved wealth. We show that both signals have to be used in the optimal tax mix. If the government could commit not to increase the total tax burden on investors, investors would even benefit from the introduction of a FTT, if it is compensated by an appropriate decrease in the tax on capital income. However by extending the model to include workers, we show that even if a FTT always increases aggregate welfare and workers' utilities, it may be opposed by investors.

1 Introduction

In times of crisis, governments are often tempted to introduce a Financial Transaction Tax (FTT) in order to increase their fiscal revenue. A famous historical example is the UK stamp duty on share transfers. It was enacted in 1694 in order to finance a war against France. It was supposed to stop after four years but it is still in place as we write.

Similarly, FTTs are often envisaged after financial crises, but then with a more "corrective" or even "punitive" objective. For example, following the Global Financial Crisis of 2007-09, G-20 leaders requested the IMF to study how the financial sector could "make a fair and substantial contribution to meeting the costs associated with government interventions to repair it" (Claessens et al 2010, page 2). The idea was to create a Pigouvian tax that would force banks to internalize the negative externalities they may exert on the financial system¹.

And of course, following Tobin (1978), Stiglitz (1998), Summers and Summers (1998), and more recently Davila (2017), there is also a large literature motivating the FTT as a way to limit "excessive" speculation and stabilize financial markets. However there does not seem to be clear evidence that a FTT indeed reduces volatility on financial markets. Moreover, the FTT is criticized by the financial industry for generating sizeable distortions: reducing transaction volume, hindering hedging and slowing down price discovery. Most empirical studies confirm the views of the industry.

We adopt a different viewpoint and study the justification of a FTT in a context where financial markets are perfect but the tax system is imperfect. In our model, the tax system is imperfect because of an unobservable heterogeneity among investors: non financial wealth is not (fully) observable². We also assume that richer people are inclined to make more financial transactions than poorer people. In this context we show that a FTT is always part of the optimal tax mix.

We consider a government that needs to levy taxes in order to fund some given level of public expenditures. In the basic version of the model, there is only one good, that can be consumed or invested. The initial wealth of investors is not observable by the government.³ Investors can liquidate part of their initial wealth and use the proceeds to invest on financial markets, which is observable by the government, and can therefore be taxed.

¹Even though FTTs were enacted in some countries (like France or Italy) after the GFC, the IMF report concluded that such taxes were not an appropriate instrument for financial stability and proposed instead "Financial Stability Contributions" and "Financial Activities Taxes".

²If wealth was fully observable, an inheritance tax would be optimal. On this topic, see Farhi-Werning(2010) and Piketty-Saez (2013).

³For example, an important part of US taxpayers wealth are private businesses, whose revenues are difficult to assess. Zwick (2020) estimates that private businesses account for a large fraction (around 50%) of top wealth and notes that such wealth is difficult to value. There are often no observed transaction prices, so that valuing private business wealth requires self-assessment, and estimation methods that can be imprecise and arbitrary. This leaves ample room for information asymmetries between citizens and tax authorities.

We assume, in addition, that agents' non financial wealth is subject to liquidity shocks. To model this, we assume there is a publicly observable macro-shock, which can be positive or negative. Half the agents have positive exposure to the aggregate shock, while the other half have negative exposure. To hedge this exposure, agents can initially trade in a swap market, with payoffs contingent on the publicly observable shock. Thus, investors that have a positive (resp. negative) exposure to the publicly observable shock hedge by taking a negative (resp. positive) position in the derivative market. We study whether the optimal tax mechanism involves a financial transactions tax, based on hedging-motivated trades in the swap market.

In our model, agents with larger non financial wealth have larger exposure to the shock, and are therefore inclined to take a larger position in the swap market. Thus financial transactions offers a second signal, on top of capital income, about the hidden variable, namely non financial wealth. We show that, if the total fiscal burden on the financial sector is fixed, combining financial transactions taxes and capital income taxes is less distortive than relying on capital income taxation only. Thus, it is optimal to tax financial transactions.

Relation with the literature: A key argument in favour of financial transaction taxes dates back to Keynes (1936). In line with Keynes (1936), Tobin (1978) and then Stiglitz (1989) called for financial transaction taxes, arguing they would curb speculation, and thus reduce excess volatility. In response to this argument, financial institutions, as well as many economists (such as Cochrane, 2013), claim that financial transaction taxes would on the contrary decrease market quality, reduce liquidity and make prices less informative. Thus, most of the debate revolves on whether financial transaction taxes increase or decrease financial markets efficiency. Subrahmaniam (1998), Dow and Rahi (2000) and Sorensen (2019) have contributed to the analysis of this Pigovian aspect of financial transaction taxes. Dow and Rahi (2000) and Sorensen (2019) offer models in which all agents have well defined preferences and make optimal decisions, so that welfare can be analysed. The model in Dow and Rahi (2000) involves risk averse agents trading for informational or hedging motives. As in Subrahmaniam (1998), in response to the tax informed agents scale back their trades, but Dow and Rahi (2000) show that this can lead to larger (after tax) profits for informed traders than when there is no tax.⁴ The consequences of financial transaction taxes have been documented empirically, e.g., Colliard and Hoffman (2017) and Gomber et al (2015) find a decrease in market liquidity after the introduction of the financial transaction tax in France in 2012.

While the literature discussed above focuses on financial transactions taxes and their consequences for financial markets liquidity (and often finds financial transactions taxes reduce liquidity), we take a different approach. We consider financial transactions taxes along with other taxes, such as capital income taxes

⁴The effect of the tax on hedgers depend on its consequences on hedging opportunities. By making prices less informative, the tax can actually increase the expected utility of hedgers, because of the Hirshleifer (1971) effect. Thus, the tax can lead to a Pareto improvement.

and labor income tax. In this context, we ask whether the optimal taxation scheme involves only capital and labor income tax, or also involves financial transactions taxes.

Thus, our analysis fits in the optimal mechanism approach to taxation initiated by Mirrlees (1971) and Diamond and Mirrlees (1978). In our analysis, the fundamental friction is asymmetric information about agents' endowments, and the tax scheme is designed optimally to mitigate the distortions induced by information asymmetry.⁵

2 Basic model

Our basic model has three dates $t = 0, 1, 2$, features investors and a government, and involves only one good, that can be consumed or invested. After analyzing this basic model, we will enrich it by adding another category of agents, workers, who provide a second good, labor, used in the production process.

There is a mass one continuum of heterogeneous investors. At time 0, an investor of "type" y receives an initial endowment y of the good. These initial endowments are distributed over some interval $[y_{\min}, y_{\max})$ with c.d.f. F and density f . Denote by e the total endowment at time 0:

$$e = \int_{y_{\min}}^{y_{\max}} y dF(y).$$

At time 0, the consumption good can be stored for consumption at time 1 or invested in a long term technology producing, at time 2, $R > 1$ units of good per unit of investment. The storage technology is available to all investors, but the long term technology is managed by competitive firms without initial resources.

At time 1, there is a publicly observable macro-shock, ε_M , which can take the value $+1$ or -1 with equal probability. Half of the investors of each type y have positive exposure to this shock, while the other half have negative exposure. Thus, at time 1, the type y investors with positive exposure receive $\varepsilon_M \sigma$, while those with negative exposure receive $-\varepsilon_M \sigma$. Since the mass of investors with positive exposure to the shock is equal to the mass of investors with negative exposure, the aggregate exposure of the economy is 0. Thus, there is no aggregate risk, and each individual shock can be interpreted as a pure liquidity shock. Individual liquidity shocks are denoted ε . For simplicity we assume that whether an agent has positive or negative exposure to the macro-shock is publicly observable. By symmetry, we can assume that individual allocations only depend on individual shocks and not on ε_M , which will therefore not appear in the sequel. Finally, we assume that $0 < \sigma < 1$ so that individual liquidity shocks increase with, but never exceed, initial endowments.

⁵This is in line with Mirrlees (1971), whose approach is summarized as follows by Golosov et al (2006): "Rather than starting with an exogenously restricted set of tax instruments, Mirrlees's (1971) starting point is an information friction that endogenizes the feasible set of tax instruments."

This framework enables us to capture in a very simple way the two basic functions of financial markets: channeling savings to investment opportunities (with return R), and insuring investors against liquidity shocks (ε).

The time- t consumption of an investor of type y hit by a liquidity shock ε is denoted by $C_t^\varepsilon(y)$. All investors have the same preferences: an investor's expected utility at time 0 is

$$E_\varepsilon [u_1(C_1^\varepsilon(y)) + u_2(C_2^\varepsilon(y))],$$

where utility functions u_1 and u_2 are concave increasing.

Finally, the government raises taxes at time 2 in order to finance public expenditures G , which we take as given for the moment. In addition to the resource constraint, the government is subject to the "safety net constraint" that no citizen get utility level below a minimum, which we denote u_{\min} . Thus, the safety net constraint is

$$E_\varepsilon [u_1(C_1^\varepsilon(y)) + u_2(C_2^\varepsilon(y))] \geq u_{\min}, \forall y. \quad (1)$$

We start by analyzing the case in which initial endowments y are publicly observable, so that the government can use non distortionary wealth taxes $T(y)$ (first best). Then we turn to the case where the government cannot observe initial endowments (second best). Finally, we study how the second best can be implemented with markets and taxes.

3 First best

3.1 Optimal allocation

In the absence of informational constraints, Pareto optimal allocations are obtained by maximizing weighted sums of utilities for arbitrary weights $\alpha(y) \geq 0$:⁶

$$\max E_{y,\varepsilon} \alpha(y) [u_1(C_1^\varepsilon(y)) + u_2(C_2^\varepsilon(y))], \quad (2)$$

subject to the safety net constraint (??)⁷, and the resource constraint (taking into account that aggregate investment is equal to aggregate endowment e minus aggregate consumption $E_{y,\varepsilon} [C_1^\varepsilon(y)]$)

$$E_{y,\varepsilon} [RC_1^\varepsilon(y) + C_2^\varepsilon(y)] \leq Re - G. \quad (3)$$

Denoting by λ the multiplier associated with the resource constraint, and by $\mu(y)$ the multiplier associated with the safety net constraint, the Lagrangian is

$$E_{y,\varepsilon} [(\alpha(y) + \mu(y))(u_1(C_1^\varepsilon(y)) + u_2(C_2^\varepsilon(y)) - \lambda(RC_1^\varepsilon(y) + C_2^\varepsilon(y)) - \mu(y)u_{\min})]$$

The optimal allocation is pinned down by the first order condition with respect to $C_1^\varepsilon(y)$:

$$(\alpha(y) + \mu(y))u_1'(C_1^\varepsilon(y)) = \lambda R, \quad (4)$$

⁶We normalize weights so that $E_y[\alpha(y)] = 1$.

⁷We rule out the Pareto optima that do not satisfy this constraint.

and with respect to $C_2^\varepsilon(y)$:

$$(\alpha(y) + \mu(y))u_2'(C_2^\varepsilon(y)) = \lambda. \quad (5)$$

They imply that consumption at times 1 and 2 does not depend on ε , and thus can be written $C_1(y)$ and $C_2(y)$. Since there is no aggregate risk, liquidity shocks are fully insured in the first best, so that consumption at times 1 and 2 is independent of ε .

Moreover, dividing (??) by (??), we have

$$\frac{u_1'(C_1(y))}{u_2'(C_2(y))} = R. \quad (6)$$

Condition (??) states that in the first best, the marginal rate of substitution between consumption at time 1 and consumption at time 2 equals the marginal rate of transformation, R . To complete the characterization of the first best, note that when the safety net constraint binds, we have:

$$u_1(C_1(y)) + u_2(C_2(y)) = u_{\min} \quad (7)$$

In this case consumption at time 1 and 2, which we denote by (C_1^{\min}, C_2^{\min}) , is pinned down by (??) and (??). Using condition (??), the fact that $\mu(y) \geq 0$ implies

$$\alpha(y) \leq \alpha_{\min} \equiv \frac{\lambda}{u_2'(C_2^{\min})},$$

i.e., condition (??) binds for agents with low Pareto weights. For the other agents, for which $\alpha(y) > \alpha_{\min}$, $u_2'(C_2(y)) = \frac{\lambda}{\alpha(y)}$.

3.2 Implementation

As implied by the second welfare theorem, the first best allocations can be implemented by wealth taxes $T(y)$ (which can be interpreted as personalized lump sum transfers, or reallocation of initial endowments) and complete financial markets. In our simple set-up, only two markets are needed at $t = 0$: a market for bonds issued by firms and repaid at $t = 2$, and a market for swaps contingent on the macro-shock.

An investor buying one unit of the swap contract receives p units of the good at $t = 1$ if $\varepsilon = -1$ and pays one unit if $\varepsilon = 1$. Thus, p is the swap rate. Perfect competition for funding among firms implies the equilibrium return on bonds is R . Symmetry of liquidity shocks implies that the equilibrium swap rate must be $p = 1$. We denote by $S(y)$ the savings (number of bonds purchased) of type y investor, and by $\Delta(y)$ the payment received from his or her swap position when hit by a negative shock σy . By symmetry, the investor has to pay $\Delta(y)$ when hit by a positive shock. The indirect utility function of a type y investor can thus be written as

$$U(y) = \max_{S, \Delta} E_\varepsilon [u_1(y + \varepsilon\sigma(y) - S(y) - \varepsilon\Delta(y)) + u_2(RS - T(y))].$$

The first order condition with respect to $\Delta(y)$ yields

$$E_\varepsilon[-\varepsilon u'_1(C_1^\varepsilon(y))] = 0.$$

That is:

$$u'_1(C_1^+(y)) = u'_1(C_1^-(y)). \quad (8)$$

Condition (??) states that time-1 consumption is independent of ε . There is complete insurance of liquidity shocks, as requested by condition (??). The first order condition with respect to $S(y)$ yields equation (??), the first best optimality condition for savings and investment. Individual savings, swap positions and taxes are deduced easily

$$S(y) = y - C_1(y), \Delta(y) = \sigma(y), T(y) = RS(y) - C_2(y). \quad (9)$$

Proposition 1 *Any first best allocation can be implemented by a wealth tax $T(y)$ and two competitive financial markets operating at $t = 0$, a bond market and a swap market.*

4 Second best

We now turn to the case in which the government cannot observe investors' initial endowments y . In this case, allocations must satisfy incentive compatibility constraints.

4.1 Incentive compatible mechanisms

By the revelation principle, we can restrict attention to direct mechanisms, mapping reported types \hat{y} into net savings after swap payoffs $S^\varepsilon(\hat{y}) = S(y) + \varepsilon\Delta(\hat{y})$ (which, along with endowments determine time-1 consumption), and time-2 consumption $C_2^\varepsilon(\hat{y})$. The savings of the investors are invested in the productive technology.

If an investor of type y reports \hat{y} , his or her time 1 consumption is $[y + \sigma(y)\varepsilon - S^\varepsilon(\hat{y})]$ and time 2 consumption is $C_2^\varepsilon(\hat{y})$. The indirect utility function of investor y is

$$U(y) = \max_{\hat{y}} E_\varepsilon[u_1(y + \sigma(y)\varepsilon - S^\varepsilon(\hat{y})) + u_2(C_2^\varepsilon(\hat{y}))].$$

The incentive compatibility condition is that truthful reporting is optimal

$$y \in \arg \max_{\hat{y}} E_\varepsilon[u_1(y + \sigma(y)\varepsilon - S^\varepsilon(\hat{y})) + u_2(C_2^\varepsilon(\hat{y}))]. \quad (10)$$

By the envelope theorem, the incentive compatibility condition implies that the derivative of the value function of the agent is

$$U'(y) = E_\varepsilon[(1 + \sigma\varepsilon)u'_1(C_1^\varepsilon(y))]. \quad (11)$$

Since $\sigma < 1$, $U'(y) \geq 0$, i.e., the value function U of the agent is increasing in his or her initial endowment y . By concavity of u_1 , investors' preferences satisfy the single crossing property with respect to initial endowments y and net savings $S^\varepsilon(\hat{y})$ in each state ε . Thus implementable functions $S^\varepsilon(\cdot)$ are necessarily non decreasing. The next lemma show that the converse is true.

Lemma 1: *A direct mechanism satisfies the incentive compatibility condition if and only if the associated indirect utility function satisfies (??) and $S^\varepsilon(y)$ is non decreasing in y for all ε .*

4.2 Optimality conditions

Second best allocations are obtained by maximizing weighted sums of investors' utilities under the feasibility and incentive compatibility constraints. Hereafter, we study the relaxed problem, where the incentive compatibility constraint (??) is replaced by the weaker envelope condition (??). We will check ex-post that the solution of the relaxed problem is such that net savings do increase in wealth, so that the mechanism is indeed incentive compatible. The relaxed problem is:

$$\max_{C_1^\varepsilon(y), C_2^\varepsilon(y)} E_{\varepsilon, y} \alpha(y) [u_1(C_1^\varepsilon(y)) + u_2(C_2^\varepsilon(y))],$$

subject to the resource constraint (with multiplier λ)

$$E_{\varepsilon, y} [RC_1^\varepsilon(y) + C_2^\varepsilon(y)] \leq eR - G, \quad (12)$$

the envelope condition (with multiplier $\beta(y)$)

$$U'(y) = E_\varepsilon [(1 + \varepsilon\sigma)u'(C_1^\varepsilon(y))], \quad (13)$$

the condition defining U (with multiplier $\gamma(y)$)

$$U(y) = E_{\varepsilon, y} [u_1(C_1^\varepsilon(y)) + u_2(C_2^\varepsilon(y))], \quad (14)$$

and the safety net constraint $U(y) \geq u_{\min}$.

Condition (??) implies that U is increasing. Thus the safety net constraint reduces to

$$U(y_{\min}) \geq u_{\min}.$$

Denoting the multiplier of the safety net constraint by μ , the Lagrangian is:

$$\begin{aligned} \mathcal{L} = & E_{\varepsilon, y} \{ (\alpha(y) + \gamma(y))(u_1(C_1^\varepsilon(y)) + u_2(C_2^\varepsilon(y))) \\ & - \lambda [RC_1^\varepsilon(y) + C_2^\varepsilon(y) - eR + G] \\ & + \beta(y)[U'(y) - (1 + \varepsilon\sigma)u'(C_1^\varepsilon(y))] \\ & - \gamma(y)U(y) + \mu[U(y_{\min}) - u_{\min}] \} \end{aligned}$$

Maximizing \mathcal{L} with respect to consumption and rent is a combination of a variation calculus problem (in $U(y)$ and $U'(y)$) and a pointwise optimization problem

(in $C_1^\varepsilon(y)$ and $C_2^\varepsilon(y)$). The first order conditions of the pointwise maximisation problem with respect to $C_1^\varepsilon(y)$ and to $C_2^\varepsilon(y)$ are

$$u_1'(C_1^\varepsilon(y)) = \frac{\lambda}{\alpha(y) + \gamma(y)}R + \frac{\beta(y)}{\alpha(y) + \gamma(y)}(1 + \varepsilon\sigma)u_1''(C_1^\varepsilon(y)), \quad (15)$$

and

$$u_2'(C_2^\varepsilon(y)) = \frac{\lambda}{\alpha(y) + \gamma(y)}, \quad (16)$$

respectively. Note that C_2 does not depend on ε , hence we can drop this index. Dividing (15) by (16) yields

$$\frac{u_1'(C_1^\varepsilon(y))}{u_2'(C_2^\varepsilon(y))} = R + \frac{\beta(y)}{\lambda}(1 + \varepsilon\sigma)u_1''(C_1^\varepsilon(y)). \quad (17)$$

The full characterization of second best allocations is provided in Appendix 2. In the next subsection we discuss the properties of these allocations.

4.3 Properties of the second best

First, note that the cross derivative of the Lagrangian with respect to $C_1^\varepsilon(y)$ and ε is $-\beta(y)\sigma u_1''(C_1^\varepsilon(y))$. Since u_1 is concave and $\sigma > 0$, has the same sign as $\beta(y)$. Hence, $C_1^+(y) > C_1^-(y)$ if and only if $\beta(y) > 0$. We state this result in our next proposition.

Proposition 2

1. *In any second best allocation where the incentive compatibility condition binds, investors are imperfectly insured against liquidity shocks at time 1.*
2. *If the multiplier $\beta(y)$ of the incentive compatibility condition is positive, the consumption of investor y at time 1 is larger after a positive shock than after a negative shock, i.e.,*

$$C_1^+(y) > C_1^-(y). \quad (18)$$

3. *Time 2 consumption, however, is not exposed to the liquidity shock.*

There is only one second best allocation for which incentive compatibility conditions are nowhere binding. It is implemented by a poll tax $T = G$ charged on all investors.⁸ We rule it out by assuming that it does not satisfy the safety net constraint, either because G or u_{min} are relatively high, or because y_{min} is low.

⁸This is the only first best allocation that is incentive compatible.

Condition (??) in Proposition 2 states that when the incentive compatibility condition binds, insurance of liquidity shocks is imperfect, in contrast with the first best (see (??)). The optimal trade-off between rent extraction and efficiency implies that investors are not perfectly insured against liquidity shocks. This is because liquidity shocks are increasing in the variable of adverse selection, as $\sigma > 0$.

Finally, $\beta(y)$ is positive when only the downward incentive compatibility condition is binding at y , i.e., investor of type y is only attracted by the allocation of the immediately poorer investor. Then, under-insurance of liquidity shocks allows to reduce marginal rent $U'(y)$.

4.4 Implementation

In this subsection, we determine the tax systems that are needed to implement second best allocations. First, note that taxing only ex post financial income $RS - \varepsilon\Delta$ is never optimal because this violates the condition that C_2 should not depend on ε . Second, we show below that taxes should not only depend on ex ante financial income RS but also on financial transactions Δ . Indeed, consider a general tax function $T(RS, \Delta)$. To implement a particular second best allocation, the indirect utility function must be such that

$$U(y) = \max_{S, \Delta} E[u_1(y(1 + \varepsilon\sigma) - S - \Delta\varepsilon) + u_2(RS - T(RS, \Delta))].$$

The first order condition with respect to Δ is:

$$\frac{\partial T}{\partial \Delta} = -E_\varepsilon \left[\varepsilon \frac{u'_1(C_1^\varepsilon(y))}{u'_2(C_2(y))} \right] = -\frac{\beta(y)}{\lambda} E_\varepsilon [(\varepsilon + \sigma)u''_1(C_1^\varepsilon(y))], \quad (19)$$

which is different from 0 when $\beta(y) \neq 0$. Thus taxes must depend on financial transactions Δ as well as gross financial income RS .

The first order condition with respect to RS is:

$$\frac{\partial T}{\partial RS} = 1 - \frac{1}{R} E_\varepsilon \left[\frac{u'_1(C_1^\varepsilon(y))}{u'_2(C_2(y))} \right] = -\frac{\beta(y)}{\lambda R} E_\varepsilon [(1 + \varepsilon\sigma)u''_1(C_1^\varepsilon(y))]. \quad (20)$$

This shows that, like the partial derivatives w.r.t. Δ , the partial derivative w.r.t. RS has the same sign as $\beta(y)$. Thus we have established:

Proposition 3 *No second best allocation can be implemented without a FTT. Moreover, the two marginal tax rates w.r.t. S and Δ have the same sign as $\beta(y)$.*

The intuition behind this proposition is that Δ and S provide two different signals on the hidden variable y . It would be sub-optimal to use only one. Note also that, in our set-up, financial markets are complete, because investors can trade bonds and swaps to generate all possible state-contingent payoffs. Yet,

investors do not hedge completely their liquidity shocks, because taxes make financial transactions costly.

When transactions volumes $\Delta(y)$ and $S(y)$ are increasing functions of y , which we will check ex post in all our solved examples, these two functions can be inverted: there is a unique y that corresponds to a particular swap position Δ and another unique y that corresponds to a particular purchase of bonds S . In this case, any second best allocation can be implemented by a separable tax schedule $T(RS)$ and $\tau(\Delta)$. The first order conditions derived above completely determine, on the relevant ranges, the derivatives of T and τ that are needed to implement a particular second best allocation. The intercept $T(0) + \tau(0)$ is determined by the budget constraint of the government.

In order to explore the optimal tax scheme further, we are going to make additional assumptions on distributions and utility functions that imply that optimal tax rates are constant. This will illustrate the comparative statics properties of optimal taxes and facilitate the study of the political economy of the FTT, the final objective of this paper.

5 An example where optimal tax rates are constant

From now on, we focus on a particular specification, which we call the linear-exponential case, where optimal tax rates are constant.

5.1 The linear-exponential case

First we assume that the utility of consumption at date 2 is linear:

$$u_2(c) \equiv c, \tag{21}$$

and we take $\alpha(y) \equiv 1$. Second we assume that the distribution of y is truncated exponential with parameter $\mathcal{A} \in R^{*+}$. Thus $y_{\max} \rightarrow \infty$, $f(y) = \frac{1}{\mathcal{A}} \exp \frac{y_{\min} - y}{\mathcal{A}}$, and

$$\frac{1 - F(y)}{f(y)} = \mathcal{A}.$$

Note that $\mathcal{A} = e^{-y_{\min}}$, which implies that parameter \mathcal{A} measures unobserved heterogeneity in wealth. Because of quasi linearity of preferences, the condition that characterizes second best allocations simplifies to:

$$u_1'(C_1^\varepsilon(y)) = R + \frac{\beta(y)}{\lambda} (1 + \sigma\varepsilon) u_1''(C_1^\varepsilon(y)). \tag{22}$$

Moreover, we show in the appendix that with a truncated exponential distribution, the multiplier of the incentive constraint is a constant:

$$\beta(y) \equiv (\lambda - 1) \mathcal{A} > 0. \tag{23}$$

5.2 Optimal consumption

The above results imply that optimal time 1 consumption is independent of y in the linear-exponential case:

Proposition 4 *In the linear-exponential case, optimal time-1 consumption is independent of y :*

$$C_1^\varepsilon(y) \equiv \mathcal{C}\left(\frac{\lambda-1}{\lambda}\mathcal{A}(1+\sigma\varepsilon)\right), \quad (24)$$

where $\mathcal{C}(\mathcal{A})$ is defined implicitly by:

$$u_1'(C(\mathcal{A})) = R + Au_1''(C(\mathcal{A})).$$

The following lemma is proved in the appendix.

Lemma 5 *The function \mathcal{C} is increasing. When u_1 is CRRA, it is also concave.*

The fact that \mathcal{C} is increasing confirms the property established more generally above: consumption is higher after a positive shock than after a negative one.

5.3 When there is no FTT

To better understand the role of financial transaction taxes in our economy, consider the “third best” case in which the financial transactions tax is constrained to be 0. Then, type y faces the following problem:

$$\max_{C_1^\varepsilon} E_\varepsilon [u(C_1^\varepsilon(y)) + [y(1+\sigma\varepsilon) - C_1^\varepsilon(y)]R(1-t)].$$

The first order condition is

$$u'(C_1^\varepsilon(y)) = R(1-t).$$

So the agent has the same consumption after good and bad liquidity shocks, in contrast with the second best. Moreover the agent’s time-1 consumption is independent of y . We therefore omit its arguments and denote it by C_1^{TB} .

Proposition 6 (Third best): *When the FTT is constrained to be zero in the linear-exponential case, the optimal consumption is independent of ε :*

$$C_1^{TB} \equiv \mathcal{C}\left(\frac{\lambda_0-1}{\lambda_0}\mathcal{A}\right),$$

where λ_0 is defined implicitly by the budget constraint of the government:

$$u_1(C_1^{TB}) + R(e - C_1^{TB}) - Au_1'(C_1^{TB}) - u_{min} = G.$$

5.4 Optimal tax rates

The optimal tax rates are also constant and easy to compute:

Proposition 7 *In the linear-exponential case, the tax rates that implement the second best allocation are constant:*

$$\tau'(\Delta(y)) \equiv \tau = \frac{t^+ - t^-}{2}, \quad (25)$$

and

$$T'(RS(y)) \equiv t = \frac{t^+ + t^-}{2}, \quad (26)$$

where $t^\varepsilon = t^*(\frac{\lambda-1}{\lambda}\mathcal{A}(1 + \sigma\varepsilon))$, and $t^*(\mathcal{A}) \equiv (1 - \frac{u'_1(C(\mathcal{A}))}{R})$.

Without a FTT (third best), the optimal tax rate on capital income is $t^*(\frac{\lambda_0-1}{\lambda_0}\mathcal{A})$.

In the appendix, we prove the following lemma.

Lemma 8 *The function t^* is increasing. When u_1 is CRRA, it is also concave.*

The auxiliary function $t^*(\cdot)$ plays an important role in the sequel. It can be interpreted as the tax rate t on capital income that maximizes fiscal revenue when there is no FTT. In our model, this fiscal revenue is

$$u_1(C_1) + R(e - C_1 - Au'_1(C_1)) - u_{min},$$

where C_1 is defined implicitly by $u'_1(C_1) = R(1 - t)$. $t^*(\mathcal{A})$ can be viewed as the rate that corresponds to the top of the Laffer curve. It is easy to see that it is increasing in \mathcal{A} , which measures unobservable heterogeneity among investors. When there is no such heterogeneity, $\mathcal{A} = 0$ and $t^*(0) = 0$: the government can extract maximal revenue without any distortion. When \mathcal{A} increases, the tax rate that maximizes fiscal revenue increases.

The formulas derived in Proposition 7 have a natural interpretation in the light of function $t^*(\cdot)$. First, the FTT rate τ is positive because

$$t^+ = t^*(\frac{\lambda-1}{\lambda}\mathcal{A}(1 + \sigma)) > t^- = t^*(\frac{\lambda-1}{\lambda}\mathcal{A}(1 - \sigma)).$$

Second, when $\sigma = 0$, there is no point in having a FTT ($\tau = 0$) and the optimal tax on capital income equals $t_0 = t^*(\frac{\lambda_0-1}{\lambda_0}\mathcal{A})$, where λ_0 is the shadow cost of public funds in the third best allocation. When G becomes very big, this multiplier goes to infinity, and the tax rate converges to $t^*(\mathcal{A})$, the maximum of the Laffer curve. When G is smaller, the tax rate is also smaller, because the government trades off aggregate surplus with fiscal revenue.

5.5 The political economy of the FTT

Proposition 7 states that, in the linear-exponential case, optimal tax rates are constant. Now, the expected utility of type y is:

$$U(y) = u_{min} + \int_{y_{min}}^y U'(z)dz.$$

By the envelope condition $U'(z) = E_\varepsilon [(1 + \sigma\sigma\varepsilon)u'(C_1^\varepsilon(z))]$, which is a constant. Thus, the indirect utility of type y is affine in y and equal to

$$U(y) = u_{\min} + (y - y_{\min})E_\varepsilon [(1 + \sigma\varepsilon)u'(c_1^\varepsilon)]. \quad (27)$$

Taking the expectation of (??) over y , the second-best utilitarian welfare is

$$\phi^{SB} = u_{\min} + (e - y_{\min})E_\varepsilon [(1 + \sigma\varepsilon)u'(C_1^\varepsilon)]. \quad (28)$$

Similarly, in the third best case (no FTT) we have:

$$U^{TB}(y) = u_{\min} + (y - y_{\min})R(1 - t).$$

Integrating over all agents

$$\phi^{TB} = u_{\min} + (e - y_{\min})R(1 - t).$$

$U^{TB}(y)$ is affine in y , with intercept u_{\min} , and going through (e, ϕ^{TB}) . Since $\phi^{TB} < \phi^{SB}$, the utility level achieved by all agents (except the poorest investor y_{\min}) is strictly larger with financial transaction taxes than without, i.e. the outcome with FTT Pareto dominates the outcome without FTT.

Thus we have established:

Proposition 9 *In the linear-exponential case, the utility of all investors is higher with a FTT than without.*

For all investors, the loss from imperfect insurance is more than compensated by the gain from a lower income tax. We prove a complementary result in the appendix:

Proposition 10 *When u_1 is CRRA:*

- *The FTT reduces the shadow cost of public funds: $\lambda < \lambda_0$.*
- *The tax rate on capital income is lower with a FTT: $t < t_0$.*
- *Savings are higher with a FTT.*

It is obvious that aggregate surplus is (strictly) higher with two instruments rather than one. But in our model, investors are unanimous: each of them gains from the FTT. So if the government could commit not to increase the total tax burden on investors, they would all agree to the FTT. The opposition of investors to the FTT may come from government's inability to commit. If we extend the model by introducing workers and/or endogenizing G , investors may lose from the FTT because the government may use the opportunity of the introduction of the FTT to increase G or reduce taxes on labor. Thus the political acceptability of the FTT by investors may depend on their expectations on future government decisions. To clarify this, we extend our model to labor taxation.

6 Extending the analysis to labor taxation

6.1 The complete model with labor

For simplicity, we still focus on the linear-exponential model for investors. In addition to capital, we consider another input: labor, supplied by a mass-one population of workers. We use the extensive model of Choné-Laroque (2011): each worker can supply $\ell \in [0, 1]$ units of labor at time 1, yielding ℓw units of consumption good at time 2. Workers differ in terms of their disutility from labor, which is denoted by x and is distributed on $[0, x_{\max}]$, with density h and c.d.f. H . The utility of worker of type x is

$$V(x) = c(x) - x\ell(x),$$

where $c(x)$ is the worker's time 2 consumption. Because workers' preferences are linear, the only relevant Pareto function is utilitarian welfare:

$$W = E_x V(x) + E_y U(y). \quad (29)$$

6.2 First best

In the first best, x and y are observable, and W is maximized, subject to the safety net constraint

$$U(y), V(x) \geq u_{\min}, \forall x \text{ and } \forall y, \quad (30)$$

and the feasibility constraint

$$E_{y,\varepsilon} C_2^\varepsilon(y) + E_x c(x) + G \leq w E_x \ell(x) + R(e - E_{y,\varepsilon} C_1^\varepsilon(y)), \quad (31)$$

where the left-hand-side is the aggregate consumption at time 2, while the right-hand-side is the aggregate output at time 2. The optimality conditions are that risk-sharing and production be efficient, i.e.,

$$C_1^\varepsilon(y) \equiv u_1'^{-1}(R) \equiv C_{FB}$$

as in the basic model, and

$$\ell(x) = 1(x \leq w).$$

The shadow cost of public funds is $\lambda = 1$. The utilities of the agents are

$$\begin{aligned} U(y) &= u_1(C_{FB}) + R(y - C_{FB}) - T_I(y) \\ V(x) &= (w - x)1(x \leq w) - T_w(x). \end{aligned}$$

Taxes are set to implement the desired surplus sharing rule, under the safety net constraint and the government budget constraint

$$E_y T_y(y) + E_x T_w(x) \geq G.$$

The program has a solution if and only if

$$u_1(C_{FB}) + R(e - C_{FB}) + \int (w - x)1(x \leq w) dH(x) - 2u_{\min} \geq G.$$

6.3 Second best

In the second best, the optimal mechanism maximizes utilitarian welfare, (??), subject to the safety net constraint (??), the feasibility constraint (??) and the incentive compatibility conditions, which, under the usual monotonicity conditions, boil down to the envelope conditions:

$$U'(y) = E_\varepsilon [(1 + \sigma y \varepsilon) u'_1(C_1^\varepsilon(y))], \quad (32)$$

and

$$V'(x) = -\ell(x). \quad (33)$$

Relying on the linear-exponential specification, we obtain the following lemma:

Lemma 11 *In the linear- exponential case with investors and workers, the optimal mechanism solves*

$$\max E_\varepsilon [(1 + \sigma \varepsilon) u'_1(C_1^\varepsilon) \mathcal{A}] + E_x \left[\ell(x) \frac{H(x)}{h(x)} \right],$$

subject to the feasibility constraint

$$E_\varepsilon [u_1(C_1^\varepsilon) + R(e - C_1^\varepsilon) - (1 + \sigma \varepsilon) \mathcal{A} u'_1(C_1^\varepsilon)] + E_x \left[\left(w - x - \frac{H(x)}{h(x)} \right) \ell(x) \right] \geq G + 2u_{\min}.$$

The lemma implies that the Lagrangian is separable and writes as

$$\mathcal{L} = \mathcal{L}_I + \mathcal{L}_w,$$

where

$$\mathcal{L}_I = \lambda E_{y,\varepsilon} [u_1(C_1^\varepsilon(y)) + R(e - C_1^\varepsilon(y))] + (1 - \lambda) E_{\varepsilon,y} [(1 + \sigma \varepsilon) u'_1(C_1^\varepsilon(y)) \mathcal{A}],$$

and

$$\mathcal{L}_w = \lambda E_x [(w - x) \ell(x)] + (1 - \lambda) E_x \left[\ell(x) \frac{H(x)}{h(x)} \right].$$

Thus, the only link between the problem concerning investors and the problem concerning workers is the shadow cost of public funds λ . Consequently, for a given value of λ , investors' time-1 consumption is independent of y , but impacted by ε , so that $c_1^+ > c_1^-$.

In \mathcal{L}_I the first term is the surplus associated with investors

$$\phi_I(c_1^\varepsilon) = E_{y,\varepsilon} [u_1(C_1^\varepsilon(y)) + R(e - C_1^\varepsilon)],$$

while the second term is the corresponding rent

$$\mathcal{R}_I(c_1^\varepsilon) = E_{\varepsilon,y} [(1 + \sigma y \varepsilon) \mathcal{A} u'_1(c_1^\varepsilon(y))].$$

Similarly, in \mathcal{L}_w the first term is the surplus associated with workers

$$\phi_w = E_x [(w - x) \ell(x)],$$

while the second term is the corresponding rent

$$\mathcal{R}_w = E_x \left[\ell(x) \frac{H(x)}{h(x)} \right].$$

Because \mathcal{L}_w is linear in $\ell(x) \in [0, 1]$, the optimal individual labor supply is either 0 or 1. Moreover, the implementability condition implies $\ell(x)$ must be decreasing. Hence, we can state our next lemma:

Lemma 12 *The optimal mechanism involves a threshold value of the disutility from work, x^* , such that*

$$\ell(x) = 1_{x \leq x^*}.$$

The lemma implies that the Lagrangian writes as

$$\mathcal{L}_w = \lambda \int_0^{x^*} (w - x) dH(x) + (1 - \lambda) \int_0^{x^*} H(x) dx = \lambda \phi_w(x^*) + (1 - \lambda) \mathcal{R}_w(x^*).$$

Thus, the optimal mechanism is characterized by three variables x^* , C_1^+ and C_1^- , corresponding, in the implementation, to three (constant marginal) tax rates: on capital income t_I , on financial transaction τ , and on labor income t_w . This analysis yields our next lemma:

Lemma 13 *With investors and workers, in the linear-exponential case, the three policy variables : optimal consumption c_1^+ and c_1^- , optimal labor supply x^* are determined by the shadow cost of public funds λ :*

$$\frac{\lambda - 1}{\lambda} = \frac{(w - x^*)h(x^*)}{\mathcal{H}(x^*)} = \frac{u'_1(C_1^\varepsilon) - R}{(1 + \sigma\varepsilon)\mathcal{A}u''_1(C_1^\varepsilon)},$$

where λ is determined by the budget constraint of the government:

$$FR_I(\lambda) + FR_W(\lambda) = G,$$

where $FR_I(\lambda)$ and $FR_W(\lambda)$ correspond to the fiscal revenues from investors and workers when the shadow cost of public funds is λ .

We can now analyze the impact of the FTT in the complete model with labor.

Proposition 14 *When the FTT is introduced:*

- Overall welfare increases.
- The shadow cost of public funds decreases.

- *The fiscal burden on investors increases.*
- *All workers benefit.*
- *Investors benefit if and only if the slope of FR_W is not too big.*

The intuition behind the proposition, which is proven in the appendix, is simple: the FTT boosts FR_I , the fiscal revenue collected from investors, therefore it reduces the shadow cost of public funds, and the fiscal burden on workers, who all benefit. However, the impact on investors is twofold: it increases their total fiscal burden but improves the tax mix. Whether they gain or lose depends on which effect is bigger in absolute value.

7 The case where wealth is partially observable

This section considers an extension of the linear-exponential model in which total wealth is now $(y_0 + y)$, where y_0 is observable, and where the distribution of non observable wealth y conditionally on y_0 is truncated exponential with parameter $\mathcal{A}(y_0)$. We assume that $\mathcal{A}(\cdot)$ is increasing, implying a positive correlation between the two types of wealth.

The second best problem is semi-separable: the allocation that maximizes utilitarian welfare under incentive compatibility, feasibility and safety net constraints satisfies the previous formulas, only that they are now parametrized by y_0 . Like before, consumptions at date 1 do not depend on y :

$$C_1^\varepsilon(y_0, y) \equiv C_1^\varepsilon(y_0) = \mathcal{C}\left(\frac{\lambda}{\lambda - 1} \mathcal{A}(y_0)(1 + \sigma\varepsilon)\right).$$

This means that optimal taxes on income and financial transactions are still linear but their rates $t(y_0)$ and $\tau(y_0)$ are increasing in y_0 . Optimal consumption at date 2 is

$$C_2(y_0, y) = U(y_0, y) - E_\varepsilon u_1(C_1^\varepsilon(y_0)),$$

where

$$U(y_0, y) = U(y_0, y_{min}) + (y - y_{min})E_\varepsilon[(1 + \sigma\varepsilon)u_1'(C_1^\varepsilon(y_0))].$$

The only link between investors with different y_0 s is the shadow cost of public funds λ , which appears in the formulas above and is implicitly determined by the feasibility constraint:

$$E[RC_1^\varepsilon(y_0) + C_2(y_0, y)] \leq Re - G.$$

It is clear that the informational rent of investors is minimized by taking $U(y_0, y_{min})$ as small as possible, that is:

$$U(y_0, y_{min}) \equiv u_{min},$$

which relaxes the feasibility constraint while satisfying the safety net constraint.

This implies that a wealth (or inheritance) tax $T_0(y_0)$ is needed for implementing this allocation. Indeed, all investors of type (y_0, y_{min}) should have the same utility level, but the taxes they pay on financial income and financial transactions increase with y_0 . The inheritance tax T_0 compensates for this difference in tax burdens:

$$T_0(y_0) = \text{Max} E_{S, \Delta} [u_1(y_0 + y(1 + \sigma\varepsilon) - S - \varepsilon\Delta) + RS(1 - t(y_0)) - \tau(y_0)\Delta] - u_{min}.$$

BRUNO: THE WEALTH TAX COULD ALTERNATIVELY BE PAID AT T=1. POSITIVE CORRELATION PREVENTS THIS TAX TO BE COMPLETELY EXTORTIVE.

8 Conclusion

The taxation of financial markets can be improved: introducing a FTT could allow to increase fiscal revenue without creating more distortions. If the government could commit not to increase the total tax burden on investors, they would all benefit from a FTT, if compensated by a proper decrease in the capital income tax. However if the government uses the FTT to increase the tax burden on investors, they may resist to it.

Our paper neglects an important aspect of the discussion about the FTT, namely the role of financial intermediaries. First, these intermediaries are likely to be opposed to the tax, because they collect fees that increase with the volume of financial transactions. Second, if the FTT is implemented, they can allow investors to escape the tax. These questions deserve further thoughts.

Finally, we stress the fact that in our model, a FTT is useful not because financial markets are imperfect but because the tax system is imperfect. This changes the traditional perspective on the FTT.

Appendix 1: Proofs of Lemma 1 and Proposition 2

Proof of Lemma 1: Define

$$U(y, \hat{y}) = E_\varepsilon [u(C_1^{SB}(\hat{y}, \varepsilon) + (y + \sigma y \varepsilon) - (\hat{y} + \sigma \hat{y} \varepsilon)) + C_2(\hat{y})].$$

We need to show that, if (??) holds and $\dot{S}^\varepsilon(y) > 0$, then the incentive compatibility condition holds, i.e., $U(y, \hat{y}) \leq U(y, y), \forall (y, \hat{y})$. To do so, we proceed in two steps.

1) First, we prove that, if $\dot{S}^\varepsilon(y) > 0$, then

$$\frac{\partial^2 U(y, y)}{\partial y \partial \hat{y}} > 0. \quad (34)$$

To do so, note that

$$\frac{\partial U(y, \hat{y})}{\partial y} = E_\varepsilon [(1 + \sigma \varepsilon) U'(C_1^{SB}(\hat{y}, \varepsilon) + (y + \sigma(y) \varepsilon) - (\hat{y} + \sigma \hat{y} \varepsilon))],$$

and correspondingly

$$\frac{\partial^2 U(y, \hat{y})}{\partial y \partial \hat{y}} = E_\varepsilon [(1 + \sigma \varepsilon) U''(C_1^{SB}(\hat{y}, \varepsilon) + (y + \sigma(y) \varepsilon) - (\hat{y} + \sigma(\hat{y}) \varepsilon)) (\dot{C}_1^{SB}(\hat{y}, \varepsilon) - (1 + \dot{\sigma}(\hat{y}) \varepsilon))]. \quad (35)$$

Now

$$S^{SB}(y, \varepsilon) = (y + \sigma(y) \varepsilon) - c^{SB}(y, \varepsilon).$$

So

$$\dot{S}^{SB}(y, \varepsilon) = (1 + \dot{\sigma}(y) \varepsilon) - \dot{c}^{SB}(y, \varepsilon).$$

Therefore if $\dot{S}^{SB}(y, \varepsilon) > 0, \forall y$, then $\dot{c}^{SB}(y, \varepsilon) - (1 + \dot{\sigma}(y) \varepsilon) < 0, \forall y$. Thus, the cross-derivative in (??) is positive, i.e., (??) holds.

2) The second step of the proof relies on (??), which states that $\frac{\partial U(y, \hat{y})}{\partial \hat{y}}$ is increasing, and the envelope condition which states that $\frac{\partial U}{\partial \hat{y}}|_{\hat{y}=y} = 0$. Together they imply that, for $\hat{y} < y, \frac{\partial U}{\partial \hat{y}} < 0$, while for $\hat{y} > y, \frac{\partial U}{\partial \hat{y}} > 0$. Thus, in the $(\hat{y}, U(y, \hat{y}))$ space, the $U(y, \hat{y})$ function is increasing for $\hat{y} < y$ and decreasing for $\hat{y} > y$. Therefore it reaches its maximum at $\hat{y} = y$, that is $U(y, \hat{y}) \leq U(y, y), \forall (y, \hat{y})$.

QED

Proof of Proposition 2: Note that

$$\frac{\partial^2 \varphi(c, R, A)}{\partial c \partial A} = -u_1''(c) > 0,$$

which implies

$$\frac{\partial \mathcal{C}}{\partial A} > 0.$$

Moreover, since $u_1''' > 0$, if $A > 0$ then φ is concave in c . Hence,

$$\mathcal{C}(R, A) = c \iff u_1'(c) = R + Au_1''(c). \quad (36)$$

Combining (??) and (??) yields Proposition 2.

QED

Appendix 2: Characterization of the Second Best

Optimality conditions: For the variation calculus problem, there are two optimality conditions. The Euler equation:

$$\frac{d}{dy} \left(\frac{\partial \mathcal{L}}{\partial \dot{U}} \right) = \frac{\partial \mathcal{L}}{\partial U}, \quad (37)$$

and the transversality condition:

$$\frac{\partial \mathcal{L}}{\partial \dot{U}} \Big|_{y_{\max}} = 0. \quad (38)$$

For the pointwise optimization problem, we assume (and will check later) that $\alpha(y) + \gamma(y) > 0$. Since u_2 is concave, \mathcal{L} is concave in $c_2^\varepsilon(y)$. Since $u_1'' < 0 < u_1'''$, \mathcal{L} is concave in $c_1^\varepsilon(y)$. Hence the solution is characterised by the first order conditions with respect to $c_1^\varepsilon(y)$ and to $c_2^\varepsilon(y)$.

Variation calculus problem: Inspecting the Lagrangian, we see that

$$\frac{\partial \mathcal{L}}{\partial \dot{U}} = \beta(y)f(y), \quad (39)$$

and

$$\frac{\partial \mathcal{L}}{\partial U} = -\gamma(y)f(y). \quad (40)$$

Integrating the Euler equation (??) we have

$$\left[\frac{\partial \mathcal{L}}{\partial \dot{U}} \right]_y^{y_{\max}} = \int_y^{y_{\max}} \frac{\partial \mathcal{L}}{\partial U(s)} ds.$$

Because of the transversality condition (??) this is

$$\frac{\partial \mathcal{L}}{\partial \dot{U}(y)} = - \int_y^{y_{\max}} \frac{\partial \mathcal{L}}{\partial U(s)} ds. \quad (41)$$

Substituting (??) and (??) into (??), we have

$$\beta(y)f(y) = \int_y^{y_{\max}} \gamma(s)f(s)ds. \quad (42)$$

Combining the optimality conditions of the pointwise maximisation and variation calculus problems: (??) implies $c_2^\varepsilon(y)$ is a function of y (and does not depend on ε). Denote that function by $c_2(y)$. (??) can be written as

$$\alpha(y) + \gamma(y) = \frac{\lambda}{u_2'(c_2(y))}.$$

That is

$$\gamma(y) = \frac{\lambda}{u'_2(c_2(y))} - \alpha(y). \quad (43)$$

Substituting, (??) rewrites as

$$\beta(y)f(y) = \int_y^{y_{\max}} \left(\frac{\lambda}{u'_2(c_2(s))} - \alpha(s) \right) f(s) ds.$$

Finding the solution: The above analysis implies we can write the first order condition with respect to $c_1^\varepsilon(y)$ as

$$c_1^\varepsilon(y) = \mathcal{C}(Ru'_2(c_2(y)), u'_2(c_2(y)) \frac{\beta(y)}{\lambda} (1 + \varepsilon\dot{\sigma}(y))). \quad (44)$$

This gives $c_1^\varepsilon(y)$ as a function of $c_2(y)$ and $\beta(y)$. To solve for $c_2(y)$ we differentiate (??)

$$U(y) = E_{\varepsilon,y} [u_1(c_1^\varepsilon(y))] + u_2(c_2(y)),$$

and equate it to (??)

$$\dot{U}(y) = E_\varepsilon [(1 + \varepsilon\dot{\sigma}(y))u'(c_1^\varepsilon(y))].$$

This gives a first-order differential equation that $c_2(y)$ has to satisfy.

$$u'_2(c_2(y))c'_2(y) = E_\varepsilon [(1 + \varepsilon\dot{\sigma}(y) - c_1^{\varepsilon'}(y))u'(c_1^\varepsilon(y))]. \quad (45)$$

To explicitly write the differential equation, the terms $c_1^{\varepsilon'}(y)$ and $u'(c_1^\varepsilon(y))$ on the right hand side of should be expressed as function of $c_2(y)$ using (??). The first order differential equation determines $c_2(y)$ up to a constant that is determined by the resource constraint

$$eR - G = E_{\varepsilon,y} \left[RC(Ru'_2(c_2(y)), u'_2(c_2(y)) \frac{\beta(y)}{\lambda} (1 + \varepsilon\dot{\sigma}(y))) + c_2(y) \right].$$

Appendix 3: Proofs of Propositions 6 and 7 and Lemma 2

Proof of Proposition 5: When $u_{2'} = 1$, $\dot{\sigma} = 1$, and $\alpha(y) = 1$, (??) simplifies to

$$\gamma(y) = \lambda - 1$$

while (??) simplifies to

$$\beta(y)f(y) = (\lambda - 1) \int_y^{y^{\max}} \gamma(s)f(s)ds = (\lambda - 1)(1 - F(y)),$$

which yields

$$\beta(y) = (\lambda - 1) \frac{1 - F(y)}{f(y)} > 0.$$

QED

Proof of Proposition 6: Substituting $u_1'' = -1$ and (??) into (??) we obtain the marginal financial transaction tax

$$\tau'(\Delta(y)) = \frac{\lambda - 1}{\lambda} \frac{1 - F(y)}{f(y)}.$$

Similarly, (??) yields the marginal tax on capital income

$$T'(RS(y)) = \frac{\lambda - 1}{\lambda R} \frac{1 - F(y)}{f(y)}.$$

By definition

$$U(y) = E_\varepsilon [u_1(c_1^\varepsilon(y)) + c_2(y)].$$

The resource constraint at time 2 is

$$E_y [c_2(y)] + G = RE_{\varepsilon,y} [y - c_1^\varepsilon(y)].$$

Substituting this constraint into the definition of U

$$U(y) = E_\varepsilon [u_1(c_1^\varepsilon(y)) + R(y - c_1^\varepsilon(y)) - G].$$

Now

$$U(y) = u_{\min} + \int_{y_{\min}}^y \dot{U}(z)d(z).$$

Equating the two and taking expectations

$$E_{\varepsilon y} \left[u_1(c_1^\varepsilon(y)) + R(y - c_1^\varepsilon(y)) - \int_{y_{\min}}^y \dot{U}(z)d(z) \right] = u_{\min} + G.$$

After some manipulations, one obtains

$$E_{\varepsilon,y} \left[u_1(c_1^\varepsilon(y)) + R(e - c_1^\varepsilon(y)) - \frac{1 - F(y)}{f(y)}(1 + \sigma\varepsilon)u_1'(c_1^\varepsilon(y)) \right] = u_{\min} + G,$$

which yields (??).

QED

Proof of Lemma 2: In the quasi linear case, time 2 consumption can be written as

$$E_{y,\varepsilon}c_2^\varepsilon(y) = U(y) - u_1E_{y,\varepsilon}u_1(c_1^\varepsilon(y)), c(x) = V(x) + x\ell(x).$$

So the feasibility constraint, (??), rewrites

$$E_{y,\varepsilon} [U(y) - u_1(c_1^\varepsilon(y)) + Rc_1^\varepsilon(y)] + E_x [V(x) - (w - x)\ell(x)] \leq Re - G. \quad (46)$$

Now

$$\begin{aligned} E_y U(y) &= - \int_{y_{\min}}^{y_{\max}} U(y) d(1 - F(y)) \\ &= - [U(y)(1 - F(y))]_{y_{\min}}^{y_{\max}} + \int_{y_{\min}}^{y_{\max}} \dot{U}(y)(1 - F(y)) dy \\ &= u_{\min} + E_{\varepsilon,y} [(1 + \sigma y \varepsilon) u_1'(c_1^\varepsilon(y)) \mathcal{A}]. \end{aligned} \quad (47)$$

where the second equality stems from substituting (??), while the last equality stems from the binding safety net constraint at y_{\min} and the the exponential distribution assumption. Similarly,

$$\begin{aligned} E_x V(x) &= \int_{x_{\min}}^{x_{\max}} V(x) dH(x) \\ &= [V(x)H(x)]_{x_{\min}}^{x_{\max}} - \int_{x_{\min}}^{x_{\max}} \dot{V}(x)H(x) \\ &= u_{\min} + \int_{x_{\min}}^{x_{\max}} \ell(x)H(x) \\ &= u_{\min} + E_x \left[\ell(x) \frac{H(x)}{h(x)} \right], \end{aligned} \quad (48)$$

where the second equality stems from substituting (??), while the last equality stems from the binding safety net constraint at y_{\min} . Substituting (??) and (??) in (??) and (??), we obtain Lemma 2.

QED

References

- Atkinson, A. B. and J.E. Stiglitz, 1976, "The Design of Tax Structure: Direct Versus Indirect Taxation." *Journal of Public Economics*, 6(1-2): 55-75.
- Banks, J. and P.Diamond , 2011, "The Base for Direct Taxation, "IFS, London.
- Boadway, R. and P.Pestieau, 2018, "The tenuous case for an annual wealth tax "IEB, Barcelona.
- Chamley, C. (1986). "Optimal taxation of capital income in general equilibrium with infinite lives." *Econometrica* 54, 607-622. Chone, P. and G. Laroque, 2011, "Optimal taxation in the extensive model." *Journal of Economic Theory* 146,2.
- Claessens,S., M. Keen, and C., 2010, "Financial Sector Taxation. The IMF's Report to the G-20, "IMF, Washington D.C.
- Colliard, J.E, and P. Hoffman, 2017, "Financial Transaction Taxes, Market Composition, and Liquidity", *Journal of Finance*, 72, 2685-2716.
- Cremer, H. and F. Gahvari, 1995. "Uncertainty, Optimal Taxation and the Direct Versus Indirect Tax Controversy." *Economic Journal*, 105, 1165-1179.
- Cremer, H. and P. Pestieau, 2003. "Wealth Transfer Taxation: A Survey." CESifo working paper 1061.
- Cremer, H., P. Pestieau, and J.C. Rochet. 2003. "Capital Income Taxation when Inherited Wealth is Not Observable." *Journal of Public Economics*, 87(11): 2475-90.
- Davila, E., 2017, "Optimal Financial Transaction Taxes", Working paper, N.Y.U.
- Diamond, P., 1998, "Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates," *American Economic Review*, 88, 83-95.
- Dow, J., and R. Rahi, 2000, "Should Speculators be Taxed?", *Journal of Business*, 73 89-107.
- Farhi, E., and I. Werning. 2010. "Progressive Estate Taxation." *Quarterly Journal of Economics* 125 (2): 635-673.
- Golosov, M., N. Kocherlakota and A. Tsyvinski, 2003. "Optimal Indirect and Capital Taxation." *Review of Economic Studies*, 70(3): 569-87.
- Golosov, M., A. Tsyvinski and I.Werning, 2007 "New Dynamic Public Finance: A User's Guide "NBER 1181.
- Judd, K.L. 1985. "Redistributive Taxation in a Simple Perfect Foresight Model." *Journal of Public Economics*, 28(1): 59-83.
- Mirrlees, J, 1971, "An Exploration in the Theory of Optimum Income Taxation." *Review of Economic Studies*. 38 175-208.
- Piketty, T. and E. Saez, 2013 "A Theory of Optimal Inheritance Taxation", *Econometrica*, Vol. 81, No. 5, September 2013, 1851–1886).
- Saez, E. 2002c. "Optimal Income Transfer Programs: Intensive Versus Extensive Labour Supply Responses." *Quarterly Journal of Economics*, 117(3): 1039-73.

Smith, M., O.Zdar and E.Zwick (2020) "Top wealth in the United States: new estimates and implications for taxing the rich. "Princeton University.

Sorensen, P. (2007) "The theory of optimal taxation: what is the policy relevance? " *International Tax Public Finance* (2007) 14: 383–406.

Stiglitz, J.E. 1989. "Using tax policy to curb speculative short-term trading." *Journal of Financial Services Research*, 3(2): 101-115.

Subrahmanyam, A, 1998, "Transaction Taxes and Financial Market Equilibrium." *Journal of Business* 71, 81–118.

Summers, L.H., and V.P. Summers. 1989. "When financial markets work too well: a cautious case for a securities transactions tax." *Journal of financial services research*, 3 261–286.

Tobin, J., 1978, "A Proposal for International Monetary Reform," *Eastern Economic Journal*, 4 153-159.