

# Fiscal Progressivity and the Time Consistency of Monetary Policy

## Online Appendix

This Appendix contains supplementary material (Section A) and mathematical proofs of lemmas (Section B). The Arabic numbering of equations in this file refers to the main file’s equations.

### A Supplementary Material

#### A.1 Inertia in redistributive fiscal structure

The Standardized World Income Inequality Database (SWIID) provides a panel dataset of fiscal redistribution across countries: it reports comparable estimates of inequality statistics, and associated measures of absolute redistribution (the difference between the market-income and net-income Gini indices) and relative redistribution (the percentage by which market-income inequality is reduced). These variables are used to measure the redistributive structure of fiscal policy  $P_{it}$ , in country  $i$  for year  $t$ : it accounts for the effect of fiscal taxes and social transfers on the absolute and relative reduction of income inequality.<sup>31</sup>

As suggested in Figure A.1 for selected countries, there is a strong time-invariant country-specific dimension in fiscal redistribution. This intuition is confirmed in Table A.1, which reports the share of variance in absolute and relative redistribution  $Y_{it}$  explained by a simple linear country fixed-effect model:  $Y_{it} = \sum_{i=1}^N \beta_i C_i + \epsilon_{it}$ . Country dummies  $\{C_i\}$  can account for 98% of the total dispersion in relative redistribution, 97% when the sample is restricted to OECD countries. In contrast, similar regressions with market-income Gini as dependent variable yield adjusted R-squared of 85% for the whole sample and 75% when restricted to OECD countries. These results suggest that fiscal redistribution is a strong country-specific attribute, despite variations in market income inequality over time. This insight is motivating a central assumption of the modeling strategy, namely *tax-inertia* in fiscal progressivity, which is an essential component of redistributive policies.

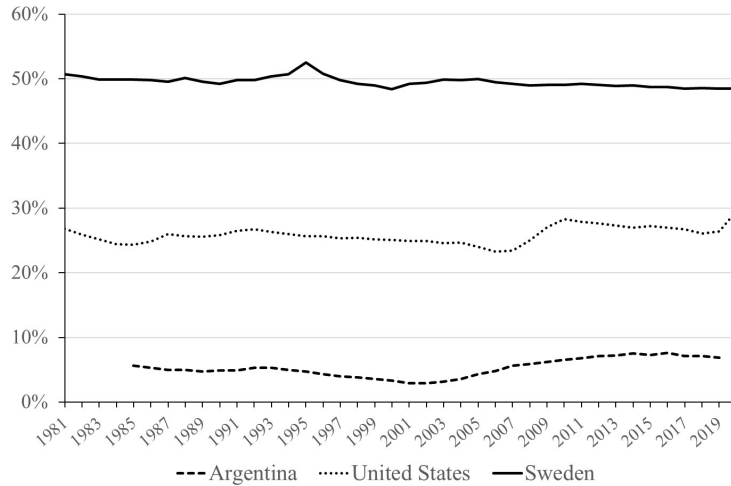
Table A.1: Adj. R2 for Country Fixed Effects Regressions

	All countries			OECD countries		
Sample years	1981 - 2021	1981 - 2001	2001 - 2021	1981 - 2021	1981 - 2001	2001 - 2021
Market income Gini	0.857	0.943	0.926	0.756	0.868	0.918
Relative redistribution	0.988	0.994	0.94	0.973	0.987	0.989
Absolute redistribution	0.974	0.987	0.991	0.938	0.970	0.978
# countries	70	70	70	38	38	38

The table reports the variance in market-income Gini and fiscal redistribution explained by country fixed effect:  $Y_{it} = \sum_{i=1}^N \beta_i C_i + \epsilon_{it}$ . Data: SWIID.

<sup>31</sup>For details about the SWIID and its construction, see Solt (2016).

Figure A.1: Inertia in Redistributive Fiscal Structures



The figure represents the variable *relative redistribution* for selected countries. It is computed as the percentage change in market- and net-income Gini coefficients. It reflects the reduction in inequality induced by labor taxes and social transfers. Data: SWIID.

## A.2 Analytical Robustness

This section discusses how results presented in Sections 3 and 4 generalize to the introduction of *ex-post* cost of inflation and richer asset structure.

### A.2.1 Ex post costs of inflation

The model includes costs of *ex ante* inflation through the effect of expected inflation on young agents’ money demand. Moreover, “unanticipated” inflation creates both costs and benefits. The benefits are at the heart of the mechanism studied in this paper: “unanticipated” inflation reduces distortionary labor taxes and thus increases output and consumption. The costs arise from the taxation of money holding.

The literature has proposed possible additional costs of inflation: price dispersion, consumption distortions or bracket creep are prominent examples. Consider the following additional term to individual agents’ value functions:

$$\mathcal{C}(z, \tilde{\pi}) = \psi(z) \frac{(1 - \tilde{\pi})^2}{2}, \tag{A.1}$$

where  $\psi(z)$  captures a possible agent specific component. Note that if  $\psi(z)$  is type-invariant, then the cost of inflation falls equally on all agents, independent of their productivity or money holding.

This additional feature does not substantively change the characterization of policy plans absent concerns for redistribution (Section 3, Proposition 1). First, the individual demand for money (7) and the relative distribution across the population (8) is not affected by the introduction of this term. Second, the policymaker is not interested in the distribution of the cost of inflation (A.1) across the population, only its average  $\Psi = \int_z \psi(z) dF(z)$ . Both under commitment or discretion, if labor taxes are raised, efficiency prescribes no

progressivity, and there is an inflation bias under policy discretion.<sup>32</sup> The relative share of inflation tax to labor income tax revenue is decreasing in  $\Psi$  though.

How do these additional costs of inflation modify the political economy analysis of Section 4? Key is to understand the influence of these additional costs (A.1) on the redistributive conflicts induced by progressive labor taxes across the population. Formally, equation (18) that characterizes bliss policies  $\lambda^p(z, \Phi_{-1})$  reads:

$$\frac{d\tilde{V}_o(z, \alpha, \Phi_{-1}, \lambda)}{d\lambda} = -\frac{\partial\tau(\cdot)}{\partial\lambda} + [\phi(z, \Phi_{-1}) + \psi(z)(1 - \tilde{\pi}(\cdot))] \frac{d\tilde{\pi}(\cdot)}{d\lambda}, \quad (\text{A.2})$$

where  $\tilde{\pi}(\cdot)$  is the inflation rate associated with a money printing rate  $\sigma$  that clears the government budget constraint given  $(\lambda, \alpha, \Phi_{-1})$ . This expression provides a decomposition of the sensitivity of value functions to labor taxes into *distortions* vs. *tax-shifting* effects, as in (19). The presence of additional costs of inflation actually mitigates the willingness to rely on the inflation tax, but it does not change the strategic conflicts induced by progressive labor taxes. For instance, if  $\psi(z)$  is proportional to relative money holding, then the decreased appetite for the inflation tax is constant across the population.<sup>33</sup> If the cost is constant across the population, then generating inflation is relatively more costly for low productivity agents, which reinforces the *tax-shifting* effect discussed in Lemma 3. Accordingly, young agents that choose progressivity behind a veil of ignorance value the reduction in the inflation bias provided by progressive labor taxes.

Overall, the presence of additional costs of inflation restrains the inflation bias, but both the *tax-shifting* effect redistributing consumption across the population and the beneficial *dynamic incentives* are still operative.

### A.2.2 Asset structure

In the model, only money is available as a store of value, whereas, arguably, households have access to greater range of assets. Whether this feature is a relevant approximation requires to assess if indeed low income households are more exposed to inflation.

Albanesi (2007) reviews evidence that supports this intuition: the cross-sectional distribution of currency holdings and transaction patterns by income level, as well as survey evidence on the perceived costs of inflation, strongly suggest that low income households are more exposed to inflation. In other terms, these elements support the view that the rich are better able to protect themselves against the effects of inflation than the poor. In particular, relatively richer agents are more likely to have access to financial instruments that hedge against inflation, while the portfolios of the poor are likely to have a larger share of cash.

This heterogeneity can be captured in the present framework with the introduction of a fixed cost to participating to asset markets and investing in real securities, such as stocks. As in Camous and Cooper (2019), such a feature sorts agents in two groups: rich agents pay the costs and invest in securities providing a hedge against inflation, while poor agents hold money and are exposed to the inflation tax. This sorting strengthens the structure of policy conflicts studied in Section 4, since inflation hurts money holders at the

<sup>32</sup>The same logic as in Lemma 1 applies. By contradiction, suppose labor income taxes are progressive. Then one can keep the total level of fiscal revenue unchanged and increase welfare by setting  $\alpha = 0$ .

<sup>33</sup>Suppose  $\psi(z) = \frac{z^2}{E(z^2)}c$ , then (A.2) rewrites:  $\frac{d\tilde{V}_o(\cdot)}{d\lambda} = -\frac{\partial\tau(\cdot)}{\partial\lambda} + \frac{z^2}{E(z^2)}[\Phi_{-1} + (1 - \tilde{\pi}(\cdot))c] \frac{d\tilde{\pi}(\cdot)}{d\lambda}$ , so that variation in inflation modifies uniformly the welfare effect of *tax-shifting* across the population.

lower end of the income distribution. The unanimity for progressive labor taxes would hold if all agents were exposed to the credibility problem of monetary policy.

Overall, an environment with a richer asset structure would deliver similar results because all nominal assets except money provide a hedge against average inflation. A diversified portfolio would then only reduce the relative exposure of rich agents to anticipated inflation in equilibrium, because it would be tilted toward these assets rather than cash.

### A.2.3 Numerical Robustness

To complement the numerical analysis, this appendix proposes an alternative set up that builds upon the numerical environment presented in Section 5. Two key features are introduced: young agents are subject to labor income taxes and the demand for money is motivated by a cash-in-advance constraint. In effect, this structure is equivalent to an overlapping heterogeneous agent cash-credit economy. This extension allows to characterize intergenerational conflicts and interpret the inflation bias not as a share of inflationary finance but as a suboptimally high level of inflation.

**Individual optimization.** The program of an agent with young age productivity  $z_y$  is characterized by the following preferences:

$$u(c_y) - g(n_y) + \beta \int_{z_h} [\alpha_1 u(c_o) + \alpha_2 u(d_o) - g(n_o)] dF(z_o|z_y), \quad (\text{A.3})$$

and a set of constraints:

$$c_y + m = z_y n_y - \lambda(z_y n_y)^{1+\alpha} + t, \quad (\text{A.4})$$

$$c_o + d_o = z_o n_o - \lambda(z_o n_o)^{1+\alpha} + m(1 - \pi) + t, \quad (\text{A.5})$$

$$d_o \leq m(1 - \pi). \quad (\text{A.6})$$

All agents work and consume when young and old. In old age, they consume two types of goods: a cash good  $d_o$ , whose consumption is subject to a cash-in-advance constraint A.6, and a regular good  $c_o$ .

**Government policy.** Each period, a planner sets the level of labor  $\lambda$  and inflation taxes  $\pi$  given progressivity  $\alpha$  to balance the following government budget constraint:

$$\int_{z_y} \lambda(z_y n_y)^{1+\alpha} dF(z_y) + \int_{z_o} \lambda(z_o n_o)^{1+\alpha} dF(z_o) + \int_{z_o} \phi(z_y, z_o) \pi dF(z_y z_o) = g + 2t \quad (\text{A.7})$$

subject to a lower bound on inflation  $\pi \geq \underline{\pi}$  and non-negativity constraint on labor taxes  $\lambda \geq 0$ . In this expression,  $\{\phi(z_y, z_o)\}$  corresponds to the distribution of money holding among old age agents as a function of their young age productivity  $z_y$ .

**Numerical solution.** Parameters are set to follow the calibration presented in Section 5.2. Equilibrium satisfies the definitions presented in Section 5.2. Policy outcomes reported in Table A.2 are characterized

under a regime of commitment and discretion, with and without progressive labor income taxes.

Table A.2: Equilibrium Statistics

	<b>Commitment</b>		<b>Discretion</b>	
	(1)	(2)	(3)	(4)
Progressivity $\alpha$	0	0.499	0	0.499
Inflation rate	-2.04%	-1.95%	2.48%	-1.98%
Lifetime welfare $\alpha$	1 (n)	1.013	0.996	1.013
Dispersion welfare	1 (n)	0.829	1.015	0.829

(n): normalization

Under commitment, with and without progressivity, a planner would select the lowest possible level of inflation  $\pi$ , effectively deflation, a result reminiscent of the Friedman rule. Under a lack of commitment instead, a planner without progressive labor taxes would choose an inflationary policy mix, to rely on predetermined money holding to service the government budget constraint. In contrast, with progressive labor income taxes, a discretionary planner would rely on the tax shifting effect to decrease consumption dispersion, hence implementing a policy mix comparable to the one under commitment. Overall, progressive labor income taxes mitigate the time inconsistency of monetary policy and neutralizes the inflation bias.

**Intergenerational conflicts.** In the presence of intergenerational heterogeneity, the conditions that allow to characterize voting outcomes as in Section 4 are not satisfied. Still, one can use the model to highlight how this additional form of heterogeneity influences individual preferences over *tax distortion* and *tax shifting* to shape redistributive conflicts, individual policy preferences and eventually economic outcomes.

When forming preferences over policy instruments, all agents in this economy face a trade-off between distortionary labor taxes, with possible differing incidence if  $\alpha > 0$ , and adjustments in the distribution of real money holding via inflation. In addition, agents belonging to different cohorts consider as well the relative exposure to labor taxes and inflation, induced by relative differences in productivity and money holding. Young agents are only exposed to labor income taxes, while old agents are exposed both to distortionary taxation and to inflationary finance. Accordingly, both cohorts have differential preferences over labor income taxes depending on individual productivity and progressivity, reflecting the intra-generational conflict highlighted in the analysis. But the intergenerational conflict determines cohort-specific preferences over inflation.

Overall, individual preferences for the monetary fiscal policy mix under inter- and intragenerational conflicts reflects three distinctive forces: the attractiveness of redistributing wealth via the inflation tax, the tax shifting effect when  $\alpha > 0$  and the relative exposure to inflation within and across generations.

## B Mathematical Proofs

### B.1 Production under Progressive Tax Plan

Consider the static production decision of an agent of type  $z$  subject to the tax plan  $\theta = (\alpha, \lambda)$ .

$$\max_y y - \tau(y, \alpha, \lambda) - \frac{(y/z)^2}{2}. \quad (\text{B.1})$$

Under the isoelastic tax plan  $\tau(y, \alpha, \lambda) = \lambda y^{1+\alpha}$ , the first order condition characterizing  $y(z, \alpha, \lambda)$  is given by:

$$1 - \lambda(1 + \alpha)y^\alpha - \frac{y}{z^2} = 0. \quad (\text{B.2})$$

The derivatives of the production function with respect to the parameters are given by:

$$\frac{dy(\cdot)}{d\lambda} = \frac{-(1 + \alpha)y^\alpha}{\lambda(1 + \alpha)\alpha y^{\alpha-1} + 1/z^2} < 0, \quad (\text{B.3})$$

$$\frac{dy(\cdot)}{d\alpha} = -\frac{\lambda y^\alpha (1 + (1 + \alpha) \log(y))}{\lambda(1 + \alpha)\alpha y^{\alpha-1} + 1/z^2}, \quad (\text{B.4})$$

$$\frac{dy(\cdot)}{dz} = \frac{2y/z^3}{\lambda(1 + \alpha)\alpha y^{\alpha-1} + 1/z^2} > 0. \quad (\text{B.5})$$

$y(\cdot)$  is decreasing in  $\lambda$  and strictly positive. When  $\alpha > 0$ ,  $y(\cdot)$  tends to 0 as  $\lambda$  goes to  $+\infty$ . Consider the second derivative of  $y(\cdot)$  with respect to  $\lambda$ . Using (B.2), rewrite (B.3) as:

$$\frac{dy(\cdot)}{d\lambda} = -(z^2)^{1+\alpha} (1 + \alpha) \frac{(y/z^2)^{1+\alpha}}{\alpha + (1 - \alpha)y/z^2}. \quad (\text{B.6})$$

The second derivative writes then:

$$\frac{d^2 y(\cdot)}{d\lambda^2} = -(z^2)^{1+\alpha} (1 + \alpha) G'(y/z^2) \frac{dy(\cdot)}{d\lambda}, \quad (\text{B.7})$$

with

$$G(X) = \frac{X^{1+\alpha}}{\alpha + (1 - \alpha)X} \quad G'(X) = \alpha X^\alpha \frac{1 + X + \alpha(1 - X)}{[\alpha + (1 - \alpha)X]^2} > 0 \quad \forall X \in [0, 1]. \quad (\text{B.8})$$

Overall, the labor supply function  $y(z, \alpha, \lambda)$  is strictly decreasing and convex for all  $\lambda \geq 0$  and  $\alpha > 0$ .

### B.2 Tax Function under Progressive Tax Plan

This section establishes the following properties of tax functions. For any  $\alpha \geq 0$ , the tax function  $t(z, \alpha, \lambda)$  is single peaked at  $\bar{\lambda}(z, \alpha) = \frac{1}{2(1+\alpha)(\frac{z}{2})^\alpha}$ , with the following properties:

- Strictly concave on the upward sloping part of the Laffer curve.

- Strictly increasing in productivity  $z$ :  $\frac{dt(\cdot)}{dz} > 0$ .

Similarly, the aggregate tax function  $T(\alpha, \lambda)$  is single-peaked and strictly concave on the upward sloping part of the Laffer curve.

**Single-Peaked Laffer Curve.** By definition,  $t(z, \alpha, \lambda) = \tau(y(z, \alpha, \lambda), \alpha, \lambda)$ . Taking the total derivative of  $t(\cdot)$  with respect to  $\lambda$ :

$$\frac{dt(z, \alpha, \lambda)}{d\lambda} = \frac{\partial \tau(\cdot)}{\partial y} \frac{dy(\cdot)}{d\lambda} + \frac{\partial \tau(\cdot)}{\partial \lambda} = \lambda(1 + \alpha)y(\cdot)^\alpha \frac{dy(\cdot)}{d\lambda} + y(\cdot)^{1+\alpha}. \quad (\text{B.9})$$

Using (B.3), we can rewrite (B.9) as:

$$\frac{dt(\cdot)}{d\lambda} = -\frac{1}{1 + \alpha} \frac{dy(\cdot)}{d\lambda} \left[ \frac{2y(\cdot)}{z^2} - 1 \right]. \quad (\text{B.10})$$

Hence we get  $\frac{dt(\cdot)}{d\lambda} \geq 0$  if and only if  $y(z, \alpha, \lambda) \leq \frac{z^2}{2}$ , i.e., using (B.2), if and only if:

$$0 \leq \lambda \leq \bar{\lambda}(z, \alpha) = \frac{1}{2(1 + \alpha)(z^2/2)^\alpha}. \quad (\text{B.11})$$

**Strict Concavity on the Upward Sloping Part of the Laffer Curve.** Take the second derivative of the tax function w.r.t.  $\lambda$ :

$$\frac{d^2 t(\cdot)}{d\lambda^2} = -\frac{1}{1 + \alpha} \left[ \frac{d^2 y(\cdot)}{d\lambda^2} \left( \frac{2y(\cdot)}{z^2} - 1 \right) + \frac{2}{z^2} \left( \frac{dy(\cdot)}{d\lambda} \right)^2 \right]. \quad (\text{B.12})$$

Rewrite (B.3) as:

$$\frac{dy(\cdot)}{d\lambda} = \frac{-(1 + \alpha)y(\cdot)^{1+\alpha}}{\alpha + (1 - \alpha)y(\cdot)/z^2}, \quad (\text{B.13})$$

and get:

$$\frac{d^2 y(\cdot)}{d\lambda^2} = \frac{\alpha}{y(\cdot)D(\cdot)} [1 + D(\cdot)] \left( \frac{dy(\cdot)}{d\lambda} \right)^2, \quad (\text{B.14})$$

where  $D(\cdot)$  is the denominator of (B.13):  $D(\cdot) = \alpha + (1 - \alpha)\frac{y(\cdot)}{z^2}$ . Using (B.14), we can rewrite (B.12) as:

$$\frac{d^2 t(\cdot)}{d\lambda^2} = -\frac{1}{1 + \alpha} \frac{1}{D(\cdot)y(\cdot)} \left( \frac{dy(\cdot)}{d\lambda} \right)^2 \left[ \alpha(1 + D(\cdot)) \left( \frac{2y(\cdot)}{z^2} - 1 \right) + \frac{2y(\cdot)}{z^2} D(\cdot) \right]. \quad (\text{B.15})$$

The term into brackets is critical for the sign of  $\frac{d^2 t(\cdot)}{d\lambda^2}$ . Posing  $X = \frac{y}{z^2}$ , we can rewrite the term into brackets as a polynomial  $P(X)$ , where the range of interest is  $X \in [0, 1]$ :

$$P(X) = \alpha(1 + D(\cdot))(2X - 1) + 2XD(\cdot), \quad (\text{B.16})$$

with  $D(\cdot) = \alpha + (1 - \alpha)X$ . Further computations lead to:

$$P(X) = 2(1 - \alpha^2)X^2 + 3\alpha(1 + \alpha)X - \alpha(1 + \alpha). \quad (\text{B.17})$$

We verify:

$$P(0) = -\alpha(1 + \alpha) < 0 \quad P(1/2) = \frac{1 + \alpha}{2} > 0 \quad P(1) = 2(1 + \alpha) > 0. \quad (\text{B.18})$$

Hence, there is a unique  $\hat{X} \in (0, 1/2)$  such that  $P(X) > 0$  if and only if  $X \in [\hat{X}, 1]$ , i.e., there is a unique  $\hat{\lambda}(z, \alpha) > \bar{\lambda}(z, \alpha)$  such that:

$$\frac{d^2 t(\cdot)}{d\lambda^2} \leq 0 \iff \lambda \leq \hat{\lambda}(z, \alpha). \quad (\text{B.19})$$

Further note:

$$\frac{dX}{dz} = \frac{dy(\cdot)/z^2}{dz} = -\lambda(1 + \alpha)\alpha y(\cdot)^{\alpha-1} \frac{dy(\cdot)}{dz} < 0, \quad (\text{B.20})$$

which imply  $\frac{d\hat{\lambda}(z, \alpha)}{dz} < 0$ , i.e., the upper bound of strict concavity of the individual Laffer curves are (inversely) ordered by productivity  $z$ .

**Ordering of the Tax Functions by Productivity.** We are to left verify the tax functions are ordered by productivity type  $z$ . Using (B.5):

$$\frac{dt(z, \alpha, \lambda)}{dz} = \frac{\partial \tau(\cdot)}{\partial y} \frac{dy(\cdot)}{dz} > 0. \quad (\text{B.21})$$

**Aggregate Tax Function.** The argument considers partitions of the productivity space  $[z_l, z_h]$  and shows the properties of the individual tax functions  $t(z, \alpha, \lambda)$  are preserved when these functions are sequentially added.

Previous elements established that for all  $z$ ,  $t(z, \alpha, \lambda)$  is single-peaked and strictly concave for all  $\lambda \in [0, \hat{\lambda}(z, \alpha)]$ , with  $\hat{\lambda}(z, \alpha) > \bar{\lambda}(z, \alpha)$ . Moreover both  $\bar{\lambda}(z, \alpha)$  and  $\hat{\lambda}(z, \alpha)$  are decreasing in  $z$ .

Consider  $F(z^0, \alpha, \lambda) = f(z_l)t(z_l, \alpha, \lambda)$ . This function satisfies the same properties as  $t(z_l, \alpha, \lambda)$ . Note  $\bar{\lambda}(z^0, \alpha)$  the value of  $\lambda$  that maximizes  $F(z^0, \alpha, \lambda)$ . Naturally,  $\bar{\lambda}(z^0, \alpha) = \bar{\lambda}(z_l, \alpha)$ .

There is a productivity level  $z^1 \in (z^0, z_h]$  such that  $\hat{\lambda}(z^1, \alpha) = \bar{\lambda}(z^0, \alpha)$  and for all  $z \in [z^0, z^1]$ ,  $f(z)t(z, \alpha, \lambda)$  is strictly concave and single-peaked, for all  $\lambda \in [0, \hat{\lambda}(z^1, \alpha)]$ . Accordingly,

$$F(z^1, \alpha, \lambda) = \int_{z^0}^{z^1} f(z)t(z, \alpha, \lambda)dF(z) + F(z^0, \alpha, \lambda) \quad (\text{B.22})$$

is also single-peaked, reached at  $\lambda = \bar{\lambda}(z^1, \alpha) < \bar{\lambda}(z^0, \alpha)$  and strictly concave over  $[0, \bar{\lambda}(z^1, \alpha)]$ .

Similarly, there is a productivity level  $z^2 \in (z^1, z_h]$  such that  $\hat{\lambda}(z^2, \alpha) = \bar{\lambda}(z^1, \alpha)$ , and by the same

argument,

$$F(z^2, \alpha, \lambda) = \int_{z^1}^{z^2} f(z)t(z, \alpha, \lambda)dF(z) + F(z^1, \alpha, \lambda) \quad (\text{B.23})$$

is also single-peaked, reached at  $\lambda = \bar{\lambda}(z^2, \alpha) < \bar{\lambda}(z^1, \alpha)$ , and strictly concave over  $[0, \bar{\lambda}(z^2, \alpha)]$ .

Eventually, after  $n$  iterations,

$$F(z^n, \alpha, \lambda) = \int_{z^{n-1}}^{z^n} f(z)t(z, \alpha, \lambda)dF(z) + F(z^{n-1}, \alpha, \lambda) = \int_{z_l}^{z_h} t(z, \alpha, \lambda)dF(z) = T(\alpha, \lambda), \quad (\text{B.24})$$

reaches a global maximum for  $\lambda = \bar{\lambda}(\alpha) \in (\bar{\lambda}(z_h, \alpha), \bar{\lambda}(z_l, \alpha))$ , and is strictly concave on its upward sloping part.

### B.3 Progressivity and Productive Efficiency (Lemma 1)

Consider first the case of a representative economy, and the following static program:

$$\max_{\alpha, \lambda} W(z, \alpha, \lambda) = y(z, \alpha, \lambda) - t(z, \alpha, \lambda) - \frac{(y(z, \alpha, \lambda)/z)^2}{2}, \quad (\text{B.25})$$

$$\text{s.t. } t(z, \alpha, \lambda) = \tau(y(z, \alpha, \lambda), \alpha, \lambda) = g, \quad (\text{B.26})$$

with non negativity constraints  $\alpha \geq 0, \lambda \geq 0$ .

The budget constraint (B.26) implicitly defines  $\lambda(\alpha)$ , the level of labor taxes required to finance  $g$  given progressivity  $\alpha$ . Accordingly, the problem rewrites:

$$\max_{\alpha} \tilde{W}(z, \alpha) = \max_{\alpha} W(z, \alpha, \lambda(\alpha)). \quad (\text{B.27})$$

Using the implementability condition (B.2), we can compute:

$$\frac{d\tilde{W}(\cdot)}{d\alpha} = -\frac{\partial\tau(\cdot)}{\partial\alpha} - \frac{\partial\tau(\cdot)}{\partial\lambda} \frac{d\lambda(\cdot)}{d\alpha}, \quad (\text{B.28})$$

and totally differentiating (B.26) with respect to  $\lambda$  and  $\alpha$ :

$$\frac{d\lambda(\cdot)}{d\alpha} = -\frac{d\tau(\cdot)/d\alpha}{d\tau(\cdot)/d\lambda}, \quad \text{with } \frac{d\tau(\cdot)}{dx} = \frac{\partial\tau(\cdot)}{\partial y} \frac{dy(\cdot)}{dx} + \frac{\partial\tau(\cdot)}{\partial x}, \quad \text{for } x \in \{\alpha, \lambda\}. \quad (\text{B.29})$$

These expressions allow to rewrite (B.28) as:

$$\frac{d\tilde{W}(\cdot)}{d\alpha} = \frac{\partial\tau(\cdot)/\partial y}{d\tau(\cdot)/d\lambda} \frac{\lambda y(\cdot)^{1+\alpha}}{1+\alpha} \frac{dy(\cdot)}{d\lambda}. \quad (\text{B.30})$$

Because  $d\tau(\cdot)/d\lambda > 0$  on the upward sloping part of the Laffer curve, it gives the desired result, i.e.,  $\tilde{W}(\alpha)$  is strictly decreasing in  $\alpha$  and therefore is maximum for  $\alpha = 0$ .

Next, consider a similar problem, where agents differ in productivity  $z \sim [z_l, z_h]$ . Let's note the optimal

tax plan  $(\alpha^*, \lambda^*)$ . This plan induces a distribution  $\{g_z\}$  of the tax burden  $g$  across the population. Formally, for all  $z$ ,  $t(z, \alpha^*, \lambda^*) = g_z$ . Assume for now type-specific flat rates are feasible. Because all agents dislike progressivity, they would unanimously favor a type-specific tax scheme  $\{\lambda_z\}$  that replicates the distribution of the tax burden  $\{g_z\}$ , but with no progressivity, i.e.,  $\lambda_z y(z, 0, \lambda_z) = g_z$ . So does the benevolent policymaker. Now, within this class of tax schemes, the policymaker prefers one that implements a flat tax rate across the population. Indeed, efficiency requires to equalize labor supply elasticities across the population.<sup>34</sup> Agent  $z$  labor supply elasticity to tax rate  $\lambda_z$  writes  $\epsilon(z, \lambda_z) = \frac{\lambda_z}{n(\cdot)} \frac{dn(\cdot)}{d\lambda_z} = -\frac{\lambda_z}{1-\lambda_z}$ . Accordingly, for all  $z' \neq z$ ,  $\epsilon(z', \lambda_{z'}) = \epsilon(z, \lambda_z)$  if and only if  $\lambda_z = \lambda_{z'}$ .

Overall, within the class of isoelastic tax scheme  $(\alpha, \lambda)$ , a benevolent policymaker implements one with no progressivity, i.e.,  $\alpha = 0$ .

## B.4 Stationary Equilibria (Lemma 2)

One needs to establish the existence of a time invariant level of aggregate real money holding  $\Phi$ , that reflects a stationary demand for money and inflation expectations. The aggregate demand for money satisfies:

$$\Phi(\mathcal{S}_{-1}) = \int_z z^2 \tilde{\pi}(\mathcal{S}) = E(z^2) \frac{\Phi(\mathcal{S})}{\Phi(\mathcal{S}_{-1})} \frac{1}{1 + \sigma(\mathcal{S})} dF(z), \quad (\text{B.31})$$

where the second equality comes from the expression for inflation

$$\tilde{\pi}(\mathcal{S}) = \frac{\Phi(\mathcal{S})}{\Phi(\mathcal{S}_{-1})} \frac{1}{1 + \sigma(\mathcal{S})}, \quad (\text{B.32})$$

and the money market clearing condition. Let's guess the stationary rate of inflation is associated with a stationary money growth rate  $\sigma$ . In this case, there is a level  $\Phi$  of aggregate money demand that satisfies  $\Phi(\alpha, \Phi) = \Phi$ . Formally, using (B.31):

$$\Phi = \frac{E(z^2)}{1 + \sigma}. \quad (\text{B.33})$$

We now verify there is a constant rate of money creation  $\sigma$  that sustains  $\Phi$ . Using the expression for the time invariant rate of inverse inflation  $\tilde{\pi}$ , one gets  $\sigma = \frac{1}{\tilde{\pi}} - 1$ . Real money demand is then  $m(z, \tilde{\pi}) = z^2 \tilde{\pi}$ . Because money market clears, the good market clears. Finally, production decision of old agents reflect intratemporal choices driven by the time invariant policy parameters  $(\alpha, \lambda)$ , see equation (B.2).

<sup>34</sup>One could solve  $\max_{\{\lambda_z\}} \int_z W(z, 0, \lambda_z) dF(z)$  subject to  $\int_z t(z, 0, \lambda_z) dF(z) = g$ .

## B.5 Individual bliss policies - Stage 2 (Lemma 3)

### B.5.1 Single-Peaked Preferences

This section shows the value function of old agents is single peaked for any  $\lambda \geq 0$ , any  $\alpha \geq 0$  and any  $z \in [z_l, z_h]$ . Formally, the shape of this function is given by:

$$\frac{d\tilde{V}_o(z, \alpha, \Phi_{-1}, \lambda)}{d\lambda} = -\frac{\partial\tau(\cdot)}{\partial\lambda} + \phi(z, \Phi_{-1})\frac{d\tilde{\pi}(\cdot)}{d\lambda}. \quad (\text{B.34})$$

Totally differentiating the government budget constraint w.r.t.  $\tilde{\pi}$  and  $\lambda$ , one gets:

$$\frac{dT(\alpha, \lambda)}{d\lambda}d\lambda + \left[\frac{d\sigma}{d\tilde{\pi}}\tilde{\pi}\Phi_{-1} + \sigma\Phi_{-1}\right]d\tilde{\pi} = 0. \quad (\text{B.35})$$

From (B.32), one gets  $\frac{d\sigma}{d\tilde{\pi}} = -\frac{1+\sigma}{\tilde{\pi}}$ , which then gives  $\frac{d\tilde{\pi}}{d\lambda} = \frac{dT(\cdot)/d\lambda}{\Phi_{-1}}$ . (B.34) rewrites:

$$\frac{1}{z^2}\frac{d\tilde{V}_o(\cdot)}{d\lambda} = -\frac{1}{z^2}\frac{\partial\tau(\cdot)}{\partial\lambda} + \frac{1}{E(z^2)}\int_z \frac{dt(z, \alpha, \lambda)}{d\lambda}dF(z), \quad (\text{B.36})$$

where for all  $z$ , the derivative of the tax function  $t(z, \cdot)$  is the sum of the *behavioral*  $\frac{\partial\tau(z, \cdot)}{\partial y_o} \frac{dy_o(z, \cdot)}{d\lambda}$  and the *mechanical*  $\frac{\partial\tau(z, \cdot)}{\partial\lambda}$  response terms.

**With no progressivity.** With  $\alpha = 0$ , then  $y_o(z, 0, \lambda) = z^2(1 - \lambda)$  and  $T(0, \lambda) = E(z^2)(1 - \lambda)\lambda$ . Therefore (B.36) rewrites:

$$\frac{d\tilde{V}_o(z, \alpha, \Phi_{-1}, \lambda)}{d\lambda} = -z^2\lambda \leq 0. \quad (\text{B.37})$$

Thus, for all  $z$ ,  $\tilde{V}_o(\cdot)$  is decreasing in  $\lambda$ . All agents have single-peaked preferences and the peak is reached for  $\lambda^P(z, 0) = 0$ .

**With progressivity -  $\alpha > 0$ .** First, let's establish the single-peaked property of old agents value functions for specific distributions  $F(\cdot)$  and then use an aggregation approach to generalize the result to any probability distribution function.

Note that for all  $\lambda \geq \bar{\lambda}(\alpha)$  (downward sloping part of the aggregate Laffer curve),  $\frac{d\tilde{V}_o(\cdot)}{d\lambda}$  is strictly negative (see (B.36)). Hence we are interested in the behavior of individual preferences over labor taxation on the upward sloping part of the aggregate Laffer curve, i.e., for  $\lambda \in [0, \bar{\lambda}(\alpha)]$ .

Two intermediate results prove useful in the following developments:

i. Behavioral response. Consider the term  $G(z, \lambda) = \frac{1}{z^2}\frac{\partial\tau(\cdot)}{\partial\lambda}$ , then:

$$\frac{dG(z, \lambda)}{dz} > 0 \quad \frac{d^2G(z, \lambda)}{dzd\lambda} < 0 \quad \forall \lambda \in [0, \bar{\lambda}(z, \alpha)]. \quad (\text{IR.1})$$

ii. Mechanical response. Consider the term  $H(z, \lambda) = \frac{1}{z^2}\frac{\partial\tau(\cdot)}{\partial y_o} \frac{dy_o(\cdot)}{d\lambda}$ . For any  $\lambda \geq 0$ , it is negative, initially

decreasing, with at most one critical point.

(IR.2)

**Proof of (IR.1)** Let  $G(z, \lambda) = \frac{1}{z^2} \frac{\partial \tau(\cdot)}{\partial \lambda} = \frac{y_o(\cdot)^{1+\alpha}}{z^2}$ . The first derivative w.r.t.  $z$  writes:

$$\begin{aligned} \frac{dG(z, \lambda)}{dz} &= \frac{1}{z^4} \left[ (1 + \alpha) y_o(\cdot)^\alpha \frac{dy_o(\cdot)}{dz} z^2 - 2z y_o(\cdot)^{1+\alpha} \right] \\ &= \frac{2\alpha}{z^3} \frac{y_o(\cdot)^\alpha}{\lambda(1 + \alpha)\alpha y_o(\cdot)^{\alpha-1} + 1/z^2} \left[ \frac{2y_o(\cdot)}{z^2} - 1 \right], \end{aligned} \quad (\text{B.38})$$

which is positive on the upward sloping part of the Laffer curve. Further, we can rewrite this expression as:

$$\frac{dG(z, \lambda)}{dz} = \frac{2\alpha z^{2(1+\alpha)}}{z^3} \frac{(y_o(\cdot)/z^2)^{1+\alpha}}{\alpha(1 - y_o(\cdot)/z^2) + y_o(\cdot)/z^2} \left[ \frac{2y_o(\cdot)}{z^2} - 1 \right]. \quad (\text{B.39})$$

Note  $Q(X) = \frac{X^{1+\alpha}}{\alpha+(1-\alpha)X} (2X - 1)$  so that the cross-second derivative of  $G(\cdot)$  writes:

$$\frac{d^2 G(z, \lambda)}{dz d\lambda} = \frac{2\alpha z^{2(1+\alpha)}}{z^3} Q'(y_o(\cdot)/z^2) \frac{dy_o(\cdot)}{d\lambda} \frac{1}{z^2}. \quad (\text{B.40})$$

The sign of  $Q'(\cdot)$  is critical for the sign of this expression. Formally, with  $Q(X) = N(X)/D(X)$ , it writes:

$$Q'(X) = \frac{N'(X)D(X) - D'(X)N(X)}{D(X)^2} \quad (\text{B.41})$$

$$N'(X) = X^\alpha [2(2 + \alpha)X - (1 + \alpha)] \quad D'(X) = (1 - \alpha). \quad (\text{B.42})$$

Reorganizing the numerator of  $Q'(X)$ :

$$Q'(X) = \frac{X^\alpha(1 + \alpha)}{D(X)^2} [-\alpha + 3\alpha X + 2(1 - \alpha)X^2]. \quad (\text{B.43})$$

Now consider  $P(X) = -\alpha + 3\alpha X + 2(1 - \alpha)X^2$  and verify that  $P(0) = -\alpha < 0$ ,  $P(1/2) = 1/2 > 0$  and  $P(1) = 2 > 0$ , so that for all  $X \in [1/2, 1]$ ,  $P(X) > 0$ .

Hence, for all  $\lambda \in [0, \bar{\lambda}(z, \alpha)]$ , as  $y_2(\cdot)/z^2 \in [1/2, 1]$ , we have  $\frac{d^2 G(z, \lambda)}{dz d\lambda} < 0$ .

**Proof of (IR.2)** Consider now  $H(z, \lambda) = \frac{1}{z^2} \frac{\partial \tau(\cdot)}{\partial y_o} \frac{dy_o(\cdot)}{d\lambda}$ . It is unambiguously negative. Rewrite  $H(\cdot)$  as:

$$\begin{aligned} H(z, \lambda) &= \frac{1}{z^2} \frac{\partial \tau(\cdot)}{\partial y_o} \frac{dy_o(\cdot)}{d\lambda} \\ &= -\frac{z^{2(1+\alpha)}}{z^2} \left[ 1 - \frac{y_o(\cdot)}{z^2} \right] \frac{(1 + \alpha)(y_o(\cdot)/z^2)^{1+\alpha}}{\alpha + (1 - \alpha)y_o(\cdot)/z^2} \\ &= -z^{2\alpha}(1 + \alpha)Q(y_o(\cdot)/z^2), \end{aligned} \quad (\text{B.44})$$

with  $Q(X) = (1 - X) \frac{X^{1+\alpha}}{\alpha+(1-\alpha)X}$ . The first derivative of  $H(\cdot)$  w.r.t.  $\lambda$  writes then:

$$\frac{dH(z, \lambda)}{d\lambda} = -z^{2\alpha}(1 + \alpha)Q'(y_o(\cdot)/z^2) \frac{dy_o(\cdot)}{d\lambda} \frac{1}{z^2}. \quad (\text{B.45})$$

The sign of  $Q'(\cdot)$  is critical for the sign of  $\frac{dH(z,\lambda)}{d\lambda}$ . Formally, with  $Q(X) = N(X)/D(X)$ , we can derive:

$$Q'(X) = \frac{N'(X)D(X) - D'(X)N(X)}{D(X)^2} \quad (\text{B.46})$$

$$N'(X) = X\alpha[1 + \alpha - (2 + \alpha)X] \quad D'(X) = (1 - \alpha). \quad (\text{B.47})$$

Reorganizing the numerator of  $Q'(X)$ :

$$Q'(X) = \frac{X^\alpha}{D(X)^2} [(1 + \alpha)\alpha - X\alpha(2\alpha + 1) - X^2(1 - \alpha)(1 + \alpha)]. \quad (\text{B.48})$$

Now consider  $P(X) = (1 + \alpha)\alpha - X\alpha(2\alpha + 1) - X^2(1 - \alpha)(1 + \alpha)$  and verify  $P(0) = (1 + \alpha)\alpha > 0$  and  $P(1) = -1 < 0$ . Since  $X = y_o(\cdot)/z^2$  and  $\frac{dy_o(\cdot)}{d\lambda}$ , we can conclude that  $Q'(y_o(\cdot)/z^2) = -1 < 0$  when  $\lambda = 0$  and over  $\lambda \geq 0$ ,  $Q'(y_o(\cdot)/z^2) = 0$  has a unique solution.

Overall,  $H(z, \lambda)$  is negative, initially increasing and has a unique critical point in  $\lambda$  over  $[0, +\infty]$ .

**Degenerate Distribution** Consider the preferences over the tax mix of an agent of type  $z$  when the probability distribution of productivity is a degenerate distribution in  $h$ .<sup>35</sup> For this special case, note  $\tilde{V}_o(\cdot)$  the value function of an agent of type  $z$  and rewrite (B.36) as:

$$\frac{1}{z^2} \frac{d\tilde{V}_o(\cdot)}{d\lambda} = \frac{1}{h^2} \frac{\partial \tau(h, \cdot)}{\partial y_o} \frac{dy_o(h, \cdot)}{d\lambda} + \frac{1}{h^2} \frac{\partial \tau(h, \cdot)}{\partial \lambda} - \frac{1}{z^2} \frac{\partial \tau(z, \cdot)}{\partial \lambda}. \quad (\text{B.49})$$

The first term is 0 for  $\lambda = 0$ , and strictly negative for all  $\lambda > 0$ . Let's consider the following cases:

i If  $z = h$ : second and third terms in (B.49) cancel out. Unambiguously, for all  $\lambda \geq 0$ :

$$\frac{d\tilde{V}_o(\cdot)}{d\lambda} \leq 0, \quad (\text{B.50})$$

where the inequality is binding if and only if  $\lambda = 0$ . Accordingly, the value function is single peaked, where the maximum is reached for  $\lambda = 0$ .

ii If  $z > h$ : the sum of second and third terms in (B.49) is strictly negative.

Indeed, using (IR.1),  $\frac{d}{dz} \left[ \frac{1}{z^2} \frac{\partial \tau(\cdot)}{\partial \lambda} \right] > 0$  on the upward sloping part of the Laffer curve. Unambiguously, for all  $\lambda > 0$ :

$$\frac{d\tilde{V}_o(\cdot)}{d\lambda} < 0. \quad (\text{B.51})$$

Accordingly, the value function is single peaked, where the maximum is reached for  $\lambda = 0$ .

iii If  $z < h$ : using (IR.1), the sum of second and third terms in (B.49) is strictly positive.

---

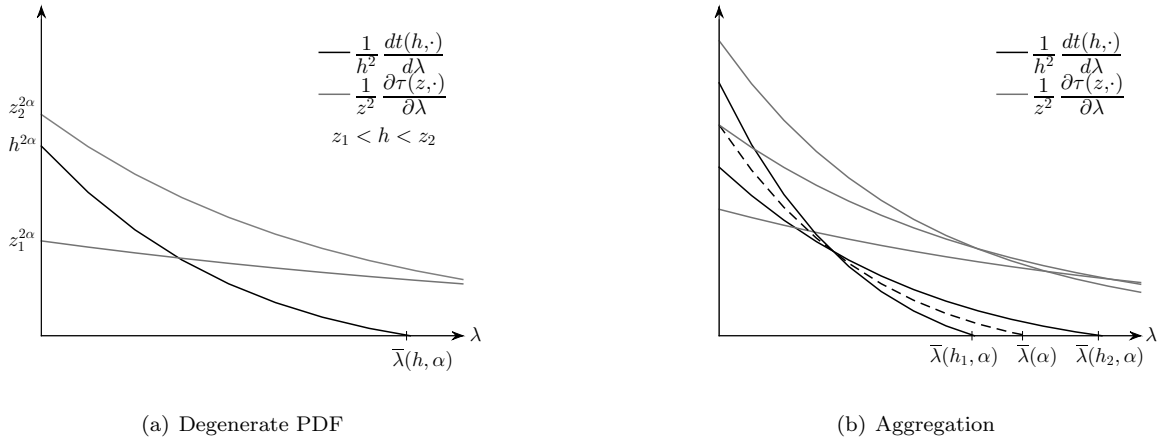
<sup>35</sup>Specifically, the population is composed of a mass 1 of agents of productivity  $h$  and mass 0 of agent of productivity  $z$ .

Accordingly,  $\left. \frac{d\tilde{V}_o(\cdot)}{d\lambda} \right|_{\lambda=0} > 0$ . Since for all  $\lambda > \bar{\lambda}(h, \alpha)$   $\left. \frac{d\tilde{V}_o(\cdot)}{d\lambda} \right|_{\lambda=0} < 0$ , we can conclude  $\tilde{V}_o(\cdot)$  has a critical point in  $[0, \bar{\lambda}(h, \alpha)]$  that characterizes a global maximum. To confirm the single-peaked property, we show this critical point is unique. Let  $\lambda^*(z, h, \alpha)$  be a solution to  $\frac{d\tilde{V}_o(\cdot)}{d\lambda} = 0$ . It satisfies:

$$-\frac{1}{h^2} \frac{\partial \tau(h, \cdot)}{\partial y_o} \frac{dy_o(h, \cdot)}{d\lambda} = \frac{1}{h^2} \frac{\partial \tau(h, \cdot)}{\partial \lambda} - \frac{1}{z^2} \frac{\partial \tau(z, \cdot)}{\partial \lambda}. \quad (\text{B.52})$$

Over  $[0, \bar{\lambda}(h, \alpha)]$ , the right-hand side is positive and decreasing. By **IR.2**, The left-hand side is 0 for  $\lambda = 0$ , positive, initially increasing, has at most one critical point, and is strictly superior to the right-hand side for  $\lambda = \bar{\lambda}(h, \alpha)$ . Accordingly,  $\lambda^*(z, h, \alpha)$  is the unique critical point of (B.49) for  $\lambda \geq 0$ . It lies on the upward sloping part of the left-hand side of (B.52). It characterizes a global maximum. The value function is single-peaked.

Figure B.2: Single-Peaked Preferences



This figure represents graphically each term of the derivative w.r.t.  $\lambda$  of the value function, see (B.36). The left panel considers the cases of degenerate PDF, where agent  $z$  expresses its preferences when the population is made of agents of type  $h$ . The value function has exactly one critical point which characterizes a global maximum, if and only if  $z < h$ . Otherwise, the value function reaches a maximum for  $\lambda = 0$ . The right panel considers the aggregation process that leads to the single-peaked preference result for any PDF. Especially, the dashed line represents  $\frac{1}{E(h^2)} \int_h \frac{dt(\cdot)}{d\lambda} dF(h)$ , which as shown in Lemma ??, has the same properties as any degenerate PDF.

Figure 3(a) summarizes these findings by representing the first derivative of the Laffer curve in  $h$ , i.e., first two terms in (B.49), and the individual cost of taxation to agent  $z$ , i.e., third term in (B.49).

**Aggregation** The generalization of the single-peaked property of value functions relies on two elements. First, for any probability function  $F(\cdot)$ , (B.36) can be written as a weighted sum of the functions (B.49). Formally,

$$\begin{aligned} \frac{1}{z^2} \frac{d\tilde{V}_o(\cdot)}{d\lambda} &= \frac{1}{E(h^2)} \int_h \frac{1}{z^2} \frac{d\tilde{V}_o(\cdot)}{d\lambda} h^2 f(h) dh \\ &= -\frac{1}{z^2} \frac{\partial \tau(z, \cdot)}{\partial \lambda} + \frac{1}{E(h^2)} \frac{dT(\alpha, \lambda)}{d\lambda}. \end{aligned} \quad (\text{B.53})$$

Second, as shown in Online Appendix 2, the properties of individual tax functions  $t(h, \alpha, \lambda)$  carry to the aggregate tax function  $\int_h t(h, \alpha, \lambda) dF(h)$  for any  $F(\cdot)$ . Accordingly, the single-peaked properties of  $\tilde{V}_o(\cdot)$  is also preserved under additivity.<sup>36</sup> Figure 3(b) presents a graphical argument to make this point clear, relying on the additive properties of individual tax functions.

Overall, old agents' value function is single peaked for any  $z$ , any  $\alpha$  and any  $F(\cdot)$ . A necessary and sufficient condition for the peak to be non 0, i.e. to be reached at  $\lambda > 0$ , is:

$$\left. \frac{d\tilde{V}_o(\cdot)}{d\lambda} \right|_{\lambda=0} > 0. \quad (\text{B.54})$$

### B.5.2 Policy Conflicts under Progressivity

When  $\alpha = 0$ , then  $y_o(z, 0, \lambda) = z^2(1-\lambda)$  and  $T(0, \lambda) = E(z^2)(1-\lambda)\lambda$ . Therefore, (19) rewrites  $\frac{d\tilde{V}_o(z, \alpha, M, \lambda)}{d\lambda} = -z^2\lambda \leq 0$ . Thus, for all  $z$ ,  $\tilde{V}_o(\cdot)$  is decreasing in  $\lambda$ . All agents have single-peaked preferences with  $\lambda^p(z, 0) = 0$ .

Consider individual policy choices under progressive labor taxes, i.e.,  $\alpha > 0$ . By Lemma 3, the value function  $\tilde{V}_o(\cdot)$  is single-peaked and downward sloping  $\forall \lambda \geq \bar{\lambda}(\alpha)$ . If there is an interior global maximum  $\lambda^p(z, \alpha) > 0$ , then it is unique and it satisfies the following conditions:

$$\frac{d\tilde{V}_o(\cdot)}{d\lambda} = -\frac{\partial\tau(\cdot)}{\partial\lambda} + \frac{z^2}{E(z^2)} \frac{dT(\alpha, \lambda)}{d\lambda} = 0 \quad \left. \frac{d^2\tilde{V}_o(\cdot)}{d\lambda^2} \right|_{\lambda=\lambda^p(z, \alpha)} < 0. \quad (\text{B.55})$$

Therefore, a necessary and sufficient condition for existence of an interior global maximum is:

$$\left. \frac{d\tilde{V}_o(\cdot)}{d\lambda} \right|_{\lambda=0} \geq 0. \quad (\text{B.56})$$

This condition induces the cut-off  $\bar{z}(\alpha)$ , such that  $\lambda^p(z, \alpha) > 0$  if and only if  $z < \bar{z}(\alpha)$ . If  $z \geq \bar{z}(\alpha)$ , then  $\lambda^p(z, \alpha) = 0$ . Formally, solving (B.56) when the inequality holds with equality,  $\bar{z}(\alpha)$  is defined by:

$$\bar{z}^{2\alpha} = \frac{E(z^{2(1+\alpha)})}{E(z^2)}. \quad (\text{B.57})$$

To verify the ordering of bliss point policy choice by productivity type, I derive the following comparative statics for all  $z < \bar{z}(\alpha)$ . Totally differentiating (B.55) with respect to  $\lambda$  and  $z$ :

$$\frac{d\lambda^p(z, \alpha)}{dz} = -\frac{\left. \frac{d^2\tilde{V}_o(\cdot)}{d\lambda dz} \right|_{\lambda=\lambda^p(z, \alpha)}}{\left. \frac{d^2\tilde{V}_o(\cdot)}{d\lambda^2} \right|_{\lambda=\lambda^p(z, \alpha)}}. \quad (\text{B.58})$$

The denominator is negative because  $\lambda^p(\alpha, z)$  is a global maximum.

Next, I show the numerator is negative if and only if the marginal rate of substitution (MRS) between inflation and labor taxes is decreasing in  $z$ . The MRS for an agent of type  $z$  is defined as:

$$\text{MRS}(z) = -\frac{dV_o(\cdot)/d\lambda}{dV_o(\cdot)/d\pi} = -\frac{\partial\tau(\cdot)/\partial\lambda}{\phi(z, \Phi_{-1})} = -\frac{E(z^2)}{\Phi_{-1}} \frac{\partial\tau(\cdot)/\partial\lambda}{z^2}, \quad (\text{B.59})$$

<sup>36</sup>Importantly, the multiplying or *weighting* terms in (B.53) are all positive and do not modify the variations of the functions considered.

and its derivative w.r.t.  $z$ :

$$\frac{d\text{MRS}(z)}{dz} = \frac{E(z^2)}{z^2\Phi_{-1}} \left[ -\frac{d\partial\tau(\cdot)/\partial\lambda}{dz} + \frac{2}{z} \frac{\partial\tau(\cdot)}{\partial\lambda} \right]. \quad (\text{B.60})$$

Now, taking the derivative of  $\frac{d\tilde{V}_o(\cdot)}{d\lambda}$  w.r.t.  $z$ :

$$\frac{d^2\tilde{V}_o(\cdot)}{d\lambda dz} = -\frac{d\partial\tau(\cdot)/\partial\lambda}{dz} + \frac{2z}{E(z^2)} \frac{dT(\alpha, \lambda)}{d\lambda}, \quad (\text{B.61})$$

and evaluating this expression in  $\lambda^p(z, \alpha)$ , using (B.55):

$$\left. \frac{d^2\tilde{V}_o(\cdot)}{d\lambda dz} \right|_{\lambda=\lambda^p(z, \alpha)} = -\frac{d\partial\tau(\cdot)/\partial\lambda}{dz} + \frac{2}{z} \frac{\partial\tau(\cdot)}{\partial\lambda} = \frac{z^2\Phi_{-1}}{E(z^2)} \frac{d\text{MRS}(z)}{dz}. \quad (\text{B.62})$$

Overall, we get:

$$\left. \frac{d^2\tilde{V}_o(\cdot)}{d\lambda dz} \right|_{\lambda=\lambda^p(z, \alpha)} < 0 \Leftrightarrow \left. \frac{d\text{MRS}(z)}{dz} \right|_{\lambda=\lambda^p(z, \alpha)} < 0. \quad (\text{B.63})$$

Next I show the derivative of the MRS w.r.t.  $z$  is indeed negative whenever agent  $z$  selects a value of  $\lambda$  on the upward sloping part of its Laffer curve, i.e., for all  $\lambda \leq \bar{\lambda}(z, \alpha)$ . From (B.60):

$$\begin{aligned} \frac{d\text{MRS}(z)}{dz} &= -\frac{E(z^2)}{\Phi_{-1}} \frac{1}{z^4} \left[ (1+\alpha)y_o(\cdot)^\alpha \frac{dy_o(\cdot)}{dz} z^2 - 2zy_o(\cdot)^{1+\alpha} \right] \\ &= -\frac{y_o(\cdot)^\alpha}{z^3} \frac{E(z^2)}{\Phi_{-1}} \frac{2\alpha}{\lambda\alpha(1+\alpha)y_o(\cdot)^{\alpha-1} + 1/z^2} \left[ \frac{2y_o(\cdot)}{z^2} - 1 \right], \end{aligned} \quad (\text{B.64})$$

which is negative as long as  $y_o(\cdot) \leq \frac{z^2}{2}$ , i.e., as long as  $\lambda \leq \bar{\lambda}(z, \alpha)$ .

Finally, I show  $\lambda^p(z, \alpha)$  is necessarily on the upward sloping part of agent  $z$  Laffer curve.<sup>37</sup>

$$\begin{aligned} \lambda^p(z, \alpha) < \bar{\lambda}(z, \alpha) &\Leftrightarrow \left. \frac{d\tilde{V}_o}{d\lambda} \right|_{\lambda=\bar{\lambda}(z, \alpha)} < 0 \\ &\Leftrightarrow \left. \frac{1}{z^2} \frac{\partial\tau(\cdot)}{\partial\lambda} \right|_{\lambda=\bar{\lambda}(z, \alpha)} > \left. \frac{1}{E(z^2)} \frac{dT(\alpha, \lambda)}{d\lambda} \right|_{\lambda=\bar{\lambda}(z, \alpha)} \\ &\Leftrightarrow \frac{1}{2} \left( \frac{z^2}{2} \right)^\alpha > \frac{1}{E(h^2)} \left[ \int_{z_l}^{z_h} \left. \frac{dt(h)}{d\lambda} \right|_{\lambda=\bar{\lambda}(z, \alpha)} dF(h) \right]. \end{aligned} \quad (\text{B.65})$$

Accordingly, if for all  $h \in [z_l, z_h]$ ,  $\left. \frac{dt(h)}{d\lambda} \right|_{\lambda=\bar{\lambda}(z, \alpha)} < \frac{h^2}{2} \left( \frac{z^2}{2} \right)^\alpha$ , then we have the desired result. Using  $\bar{\lambda}(z, \alpha) = \frac{1}{2(1+\alpha)\left(\frac{z^2}{2}\right)^\alpha}$ , we can verify that  $y_o(h, \cdot)$  evaluated in  $\bar{\lambda}(z, \alpha)$  is defined by:

$$1 - \frac{y_o^\alpha}{2\left(\frac{z^2}{2}\right)^\alpha} - \frac{y_o}{h^2} = 0. \quad (\text{B.66})$$

<sup>37</sup>Intuitively, if  $\lambda^p(z, \alpha) > \bar{\lambda}(z, \alpha)$  and because  $\bar{\lambda}(z, \alpha)$  is decreasing in  $z$ , all higher-productivity agents are taxed at a level on the downward sloping part of their Laffer curve. Agent  $z$  could then increase the total tax bill on higher-productivity agents by reducing the level of labor taxes  $\lambda$ .

The inequality (B.65) is then satisfied if and only if, for  $\lambda = \bar{\lambda}(z, \alpha)$ :

$$2\left[1 - \frac{y_o(h, \cdot)}{h^2}\right] \left[\frac{2y_o(h, \cdot)}{h^2} - 1\right] < \frac{\alpha}{2} \frac{h^2}{y_o(h, \cdot)} + \frac{1 - \alpha}{2}. \quad (\text{B.67})$$

Let  $X = \frac{y_o(h, \cdot)}{h^2} \in [0, 1]$ . The last expression rewrites then:

$$2[1 - X][2X - 1] < \frac{\alpha}{2X} + \frac{1 - \alpha}{2} \quad (\text{B.68})$$

The right-hand side is bigger than  $1/2$  for all  $X \in [0, 1]$ , whereas the left-hand side reaches a maximum value of  $1/4$ .

Altogether, we have the desired result:  $\forall z \in [z_l, \bar{z}(\alpha)]$ ,  $\frac{d\lambda^p(z, \alpha)}{dz} < 0$ . Finally, note that for all  $\alpha > 0$ ,

$$\lim_{z \rightarrow 0} \frac{1}{z^2} \frac{\partial \tau(\cdot)}{\partial \lambda} \Big|_{\lambda=0} = \lim_{z \rightarrow 0} z^{2\alpha} = 0,$$

which induces  $\lim_{z \rightarrow 0} \frac{1}{E(z^2)} dT(\alpha, \lambda) \Big|_{\lambda=\lambda^p(z, \alpha)} = 0$  and therefore  $\lim_{z \rightarrow 0} \lambda^p(z, \alpha) = \bar{\lambda}(\alpha)$ . This property outlines the clear dominance of the *tax-shifting* effect for low values of  $z$ .

## B.6 Aggregate Labor Tax Income as a Function of Progressivity (Lemma 4)

### B.6.1 Proof of the Lemma

The tax function induced by the outcome of the vote writes:

$$T(\alpha, \lambda^p(\alpha, \Phi_{-1})) = \int_z t(z, \alpha, \lambda^p(\alpha, \Phi_{-1})) dF(z). \quad (\text{B.69})$$

First, from Lemma 3,  $T(0, \lambda^p(0, \Phi_{-1})) = 0$ . Second, as shown in B.6.2, the derivative when  $\alpha = 0$  is strictly positive:

$$\frac{dT(\alpha, \lambda^p(\cdot))}{d\alpha} \Big|_{\alpha=0} > 0. \quad (\text{B.70})$$

Further, as for any  $\lambda > 0$  and  $\alpha > 0$ ,  $y(z, \alpha, \lambda) > 0$  and  $t(z, \alpha, \lambda) > 0$ , we have that  $T(\alpha, \lambda^p(\alpha, \Phi_{-1})) > 0$ . Finally, consider the level of labor taxes when progressivity gets to infinity. Rewrite the tax function as:

$$t(z, \alpha, \lambda) = \lambda e^{(1+\alpha) \log(y(z, \alpha, \lambda))}. \quad (\text{B.71})$$

From this expression, we have  $\lim_{\alpha \rightarrow +\infty} t(z, \alpha, \lambda) = 0$ . A fortiori,  $\lim_{\alpha \rightarrow +\infty} T(\alpha, \lambda^p(\alpha, \Phi_{-1})) = 0$ . Given these properties, the tax function  $T(\alpha, \lambda^p(\alpha, \Phi_{-1}))$  has a global maximum.

### B.6.2 Proof of partial result (B.70)

This section derives the sign of the derivative (B.70). Formally, the total derivative of the tax function is given by:

$$\frac{dT(\alpha, \lambda^p(\cdot))}{d\alpha} = \int_z \frac{\partial \tau(\cdot)}{\partial y_o} \frac{dy_o(\cdot)}{d\alpha} + \frac{\partial \tau(\cdot)}{\partial \alpha} + \frac{\partial \tau(\cdot)}{\partial \lambda} \frac{d\lambda^p(\cdot)}{d\alpha} dF(z). \quad (\text{B.72})$$

From Lemma 3,  $\lambda^p(0, \Phi_{-1}) = 0$ . Therefore, we can easily verify  $\left. \frac{\partial \tau(\cdot)}{\partial y_o} \right|_{\alpha=0} = 0$ ,  $\left. \frac{dy_o(\cdot)}{d\alpha} \right|_{\alpha=0} = 0$  and  $\left. \frac{\partial \tau(\cdot)}{\partial \alpha} \right|_{\alpha=0} = 0$ . Since  $\frac{\partial \tau(\cdot)}{\partial \lambda} > 0$ , (B.70) holds if and only if:

$$\left. \frac{d\lambda^p(\cdot)}{d\alpha} \right|_{\alpha=0} > 0, \quad (\text{B.73})$$

where  $\lambda^p(\cdot) = \lambda^p(\alpha, \Phi_{-1})$  is given by the bliss policy choice of the median-productivity agent. It is the solution to:

$$\frac{d\tilde{V}_o(z_m, \alpha, \Phi_{-1}, \lambda)}{d\lambda} = 0. \quad (\text{B.74})$$

Totally differentiating this expression with respect to  $\lambda$  and  $\alpha$  gives:

$$\frac{d\lambda^p(\cdot)}{d\alpha} = - \frac{d^2 \tilde{V}_o(\cdot)/d\lambda d\alpha}{d^2 \tilde{V}_o(\cdot)/d\lambda^2} \Big|_{\lambda=\lambda^p(\cdot)}. \quad (\text{B.75})$$

The denominator is negative because the value function is strictly-quasi concave. The numerator is given by:

$$\frac{d^2 \tilde{V}_o(z_m, \alpha, \Phi_{-1}, \lambda)}{d\lambda d\alpha} = - \frac{d\partial \tau(\cdot)/\partial \lambda}{d\alpha} + \frac{z_m^2}{E(z^2)} \frac{d^2 T(\alpha, \lambda)}{d\lambda d\alpha}. \quad (\text{B.76})$$

Consider the second term in this expression:

$$\frac{d^2 T(\alpha, \lambda)}{d\lambda d\alpha} = \int_z \frac{d}{d\alpha} \left[ \frac{\partial \tau(\cdot)}{\partial y_o} \frac{dy_o(\cdot)}{d\lambda} \right] + \frac{d\partial \tau(\cdot)/\partial \lambda}{d\alpha} dF(z). \quad (\text{B.77})$$

As  $\frac{\partial \tau(\cdot)}{\partial y_o} \frac{dy_o(\cdot)}{d\lambda} = -\lambda \frac{(1+\alpha)^2 y_o(\cdot)^{2\alpha}}{\lambda(1+\alpha)\alpha y_o(\cdot)^{\alpha-1} + 1/z^2}$ , we can easily show using  $\lambda^p(0, \Phi_{-1}) = 0$ :

$$\int_z \frac{d}{d\alpha} \left[ \frac{\partial \tau(\cdot)}{\partial y_o} \frac{dy_o(\cdot)}{d\lambda} \right] \Big|_{\alpha=0} dF(z) = 0. \quad (\text{B.78})$$

Further:

$$\frac{d\partial \tau(\cdot)/\partial \lambda}{d\alpha} = \left[ \log(y_o(\cdot)) + \frac{1+\alpha}{y_o(\cdot)} \frac{dy_o(\cdot)}{d\alpha} \right] y_o(\cdot)^{1+\alpha} \Rightarrow \left. \frac{d\partial \tau(\cdot)/\partial \lambda}{d\alpha} \right|_{\alpha=0} = z^2 \log(z^2). \quad (\text{B.79})$$

Rewrite (B.76) then as:

$$\left. \frac{d^2 \tilde{V}_o(z_m, \cdot)}{d\lambda d\alpha} \right|_{\alpha=0} = -z_m^2 \log(z_m^2) + \frac{z_m^2}{E(z^2)} \int_z z^2 \log(z^2) dF(z). \quad (\text{B.80})$$

Using the formula for the co-variance,<sup>38</sup> and because  $z^2$  and  $\log(z^2)$  are both increasing of  $z$ :

$$\left. \frac{d^2 \tilde{V}_o(z_m, \cdot)}{d\lambda d\alpha} \right|_{\alpha=0} > z_m^2 [-\log(z_m^2) + \log E(z^2)]. \quad (\text{B.81})$$

Since  $\log E(z^2) > \log E(z)^2$ , and using Assumption 1, we finally get:

$$\left. \frac{d^2 \tilde{V}_o(z_m, \cdot)}{d\lambda d\alpha} \right|_{\alpha=0} > 0, \quad (\text{B.82})$$

so that (B.73) holds and *a fortiori* (B.70).

## B.7 Individual Preferences for Progressivity (Lemma 5)

To establish this result, one needs to show the derivative of young agents' value function  $\tilde{V}_y(z, \alpha)$  when  $\alpha = 0$  is strictly positive for any  $z$ .  $\lambda^p(\alpha)$  and  $\tilde{\pi}^p(\alpha)$  satisfy the government budget constraint:<sup>39</sup>

$$T(\alpha, \lambda^p(\alpha)) + E(z^2) \tilde{\pi}^p(\alpha) (1 - \tilde{\pi}^p(\alpha)) = g. \quad (\text{B.83})$$

Importantly, from the analysis in Section 4.2, we have that whenever  $\alpha = 0$ ,  $\lambda^p(0) = 0$  and all the public good is financed with seigniorage only. Therefore:

$$\frac{\partial \tau(\cdot)}{\partial y_2} = \lambda(1 + \alpha) y_2(\cdot)^\alpha \quad \Rightarrow \quad \left. \frac{\partial \tau(\cdot)}{\partial y_2} \right|_{\alpha=0, \lambda=\lambda^p(0)} = 0, \quad (\text{B.84})$$

$$\frac{\partial \tau(\cdot)}{\partial \alpha} = \lambda \log(y_2(\cdot)) y_2(\cdot)^{1+\alpha} \quad \Rightarrow \quad \left. \frac{\partial \tau(\cdot)}{\partial \alpha} \right|_{\alpha=0, \lambda=\lambda^p(0)} = 0, \quad (\text{B.85})$$

$$\frac{\partial \tau(\cdot)}{\partial \lambda} = y_2(\cdot)^{1+\alpha} \quad \Rightarrow \quad \left. \frac{\partial \tau(\cdot)}{\partial \lambda} \right|_{\alpha=0, \lambda=\lambda^p(0)} = z^2. \quad (\text{B.86})$$

Totally differentiating the government budget constraint (B.83) with respect to  $\alpha$ :

$$\frac{dT(\alpha, \lambda^p(\alpha))}{d\alpha} + E(z^2) (1 - 2\tilde{\pi}^p(\alpha)) \frac{d\tilde{\pi}^p(\alpha)}{d\alpha} = 0. \quad (\text{B.87})$$

The first term writes:

$$\frac{dT(\alpha, \lambda^p(\alpha))}{d\alpha} = \int_z \frac{\partial \tau(\cdot)}{\partial y_2} \frac{dy_2(\cdot)}{d\alpha} + \frac{\partial \tau(\cdot)}{\partial \alpha} + \frac{\partial \tau(\cdot)}{\partial \lambda} \frac{d\lambda^p(\cdot)}{d\alpha} dF(z). \quad (\text{B.88})$$

<sup>38</sup> $\text{cov}(X, Y) = E(XY) - E(X)E(Y)$ .

<sup>39</sup>Recall from Lemma 3 that the outcome of the vote  $\lambda^p(\cdot)$  is independent of the aggregate seigniorage tax base  $\Phi_{-1}$ . In (B.83), the dynamic between inflation and seigniorage tax base is captured by the quadratic term in  $\tilde{\pi}^p(\cdot)$ .

Using (B.84), (B.85) and (B.86), we can evaluate (B.87) in  $\alpha = 0$  (and  $\lambda = \lambda^p(0)$ ) and get:

$$\frac{d\lambda^p(\cdot)}{d\alpha}\Big|_{\alpha=0} + (1 - 2\tilde{\pi}^p(0))\frac{d\tilde{\pi}^p(\cdot)}{d\alpha}\Big|_{\alpha=0} = 0. \quad (\text{B.89})$$

As  $\frac{d\lambda^p(\cdot)}{d\alpha}\Big|_{\alpha=0} > 0$  from (B.73), and  $\frac{1}{2} < \tilde{\pi}^p(\cdot) < 1$ , we have  $\frac{d\tilde{\pi}^p(\cdot)}{d\alpha}\Big|_{\alpha=0} > 0$ . Substituting this last expression into  $\frac{d\tilde{V}_y(z,\alpha)}{d\alpha}$ :

$$\begin{aligned} \frac{d\tilde{V}_y(\alpha, z)}{d\alpha}\Big|_{\alpha=0} &= -z^2\frac{d\lambda^p(\cdot)}{d\alpha}\Big|_{\alpha=0} + z^2\tilde{\pi}^p(0)\frac{d\tilde{\pi}^p(\cdot)}{d\alpha}\Big|_{\alpha=0} \\ &= z^2(1 - \tilde{\pi}^p(0))\frac{d\tilde{\pi}^p(\cdot)}{d\alpha}\Big|_{\alpha=0} > 0, \end{aligned} \quad (\text{B.90})$$

where the last inequality uses  $\frac{d\tilde{\pi}^p(\cdot)}{d\alpha}\Big|_{\alpha=0} > 0$  and  $\tilde{\pi}^p(0) < 1$ .

## References

ALBANESI, S. (2007): "Inflation and inequality," Journal of Monetary Economics, 54(4), 1088-1114.

CAMOUS, A. AND R. COOPER (2019): "Whatever It Takes Is All You Need: Monetary Policy and Debt Fragility," American Economic Journal: Macroeconomics, 11(4), 38-81.

SOLT, F. (2016): "The Standardized World income Inequality Database," Social Science Quarterly, 97(5), 1267-1281