Online Appendix for "A Simpler Theory of Optimal Capital Taxation"*

Emmanuel Saez, UC Berkeley

Stefanie Stantcheva, Harvard

July 6, 2016

A.1 Anticipated Reforms in the Generalized Model

In this section, we extend our analysis from the main text Section 4 to consider anticipated reforms.

Suppose that an anticipated reform to the capital income tax $d\tau_K$ happens at time $T > 0$. Capital and labor already start adjusting in anticipation of the reform before time T. The change in the utility of individual i is:

$$dV_i = d\tau_K \cdot \delta_i [\int_T^\infty u_{ic}(c_i(t), k_i(t)) r k^m(t) \cdot e^{-\delta_i t} - \int_T^\infty u_{ic}(c_i(t), k_i(t)) r k_i(t) \cdot e^{-\delta_i t}$$
$$- \frac{\tau_K}{1-\tau_K} \int_0^\infty u_{ic}(c_i(t), k_i(t)) r k^m(t) e_K(t) \cdot e^{-\delta_i t} dt]$$

In the steady state, $k^m(t)$ and $c_i(t)$, $k_i(t)$ are time-constant. Assume that T is large enough

*Stantcheva: Society of Fellows, Harvard University, Cambridge, MA 02138 (e-mail: sstantcheva@fas.harvard.edu); Saez: University of California, 530 Evans Hall #3880, Berkeley, CA 94720 (e-mail: saez@econ.berkeley.edu).

We thank Emmanuel Farhi and Thomas Piketty for useful discussions and comments. We acknowledge financial support from the MacArthur Foundation, and the Center for Equitable Growth at UC Berkeley.

so that the anticipatory responses only start when the economy is already in the steady state. In this case, all terms in $e_K(t)$ are zero before the steady state, so we can write:

$$dV_i = d\tau_K r k^m e^{-\delta_i T} \cdot [u_{ic}(c_i,k_i) - u_{ic}(c_i,k_i)\frac{k_i}{k^m} - \frac{\tau_K}{1-\tau_K}\delta_i u_{ic}(c_i,k_i)\int_{t<T} e_K(t)\cdot e^{-\delta_i(t-T)}dt$$
$$-\frac{\tau_K}{1-\tau_K}\delta_i u_{ic}(c_i,k_i)\int_{t\geq T} e_K(t)\cdot e^{-\delta_i(t-T)}dt]$$

We also assume here that the discount rates are the same across all agents. Use that $\int_i g_i = \int_i u_{ci}\omega_i = 1$.

$$dSWF = \int_i \omega_i dV_i$$
$$\propto 1 - \int_i g_i \frac{k_i}{k^m} - \frac{\tau_K}{1-\tau_K}\int_i g_i\delta\int_{t<T} e_K(t)\cdot e^{-\delta(t-T)}dt - \frac{\tau_K}{1-\tau_K}\delta\int_i g_i\int_{t\geq T} e_K(t)\cdot e^{-\delta(t-T)}dt]$$
$$= 1 - \int_i g_i \frac{k_i}{k^m} - \frac{\tau_K}{1-\tau_K}\delta\int_{t<T} e_K(t)\cdot e^{-\delta(t-T)}dt - \frac{\tau_K}{1-\tau_K}\delta\int_{t\geq T} e_K(t)\cdot e^{-\delta(t-T)}dt$$

Define the distributional factor $\bar{g}_K = \int_i g_i \frac{k_i}{k^m}$ and the anticipation elasticity e_K^{ante}, the post elasticity e_K^{post} and the total elasticity $\bar{e}_K$ to be:

$$e_K^{ante} = \delta\int_{t<T} e_K(t)\cdot e^{-\delta(t-T)}dt \quad , \quad e_K^{post} = \delta\int_{t\geq T} e_K(t)\cdot e^{-\delta(t-T)}dt \quad \text{and} \quad \bar{e}_K = e_K^{ante} + e_K^{post}$$

Then:

$$\tau_K = \frac{1-\bar{g}_K}{1-\bar{g}_K+\bar{e}_K}$$

Our formulas hence apply exactly (can mention it in the text), but the total elasticity $\bar{e}_K$ now contains anticipation effects as well. This formula is derived under the assumptions that T is large, the anticipation responses start happening only after the economy has already converged to its steady state, and discount rates are homogeneous across agents.

Endogenous Labor Supply with Anticipation Effects

The anticipation effects through the cross-elasticities can also start before the reform. The assumption needed is again that those anticipation effects only start once the economy has already converged to its steady state path. In this case, the formula looks as in the text with cross-elasticities.

A.1.1 Steady State and Anticipation Elasticities

We now prove two further results.

Steady state elasticities are finite with wealth in the utility.

With a general utility and wealth in the utility, the first-order condition for agent i in the steady state is:

$$u_{ki} = (\delta_i - \bar{r})u_{ci}$$

In the steady state, the budget constraint is:

$$c_i = \bar{r}k_i + z_i$$

hence the steady state can be rewritten as: $(\delta_i - \bar{r})u_{ci}(\bar{r}k_i + z_i, k_i) = u_{ki}(\bar{r}k_i + z_i, k_i)$ which is a smooth function of k_i, as long as the function $u_i(c_i, k_i)$ is smooth and concave in consumption and capital. Hence, the responses of consumption and capital to the net-of-tax return $\bar{r}$ are smooth and non-degenerate. The same argument holds with endogenous labor supply, which is chosen smoothly.

Anticipation elasticities are infinite with wealth in the utility and certainty, but finite with uncertainty (with or without wealth in the utility).

We can also show that the anticipation elasticities to a reform $d\tau_K$ for $t \geq T$ is infinite when there is full certainty, even with wealth in the utility. The proof is as in **?** for the Chamley-Judd model (without wealth in the utility).

With full certainty, the first-order condition of the agent with respect to capital always holds:

$$u_{ci,t} = (1 + \bar{r})/(1 + \delta_i)u_{ci,t+1} + 1/(1 + \delta_i)u_{ki,t+1}$$

Suppose we start from a situation in a well-defined steady state: $(\delta_i - \bar{r})u_{ci} = u_{ki}$ where we have perfect consumption smoothing.

The intertemporal budget constraint is:

$$\sum_{t\geq 0}\left(\frac{1}{1+r}\right)^t c_{ti} + \lim_{t\to\infty} k_{ti} = \sum_{t\geq 0}\left(\frac{1}{1+r}\right)^t z_{ti} + k_{0i}$$

Consumption smoothing implies:

$$u_{ci}(\bar{r}k_i + z_i, k_i) = \lambda$$

for the multiplier λ on the budget constraint. Hence, $k_i^\infty = \lim_{t\to\infty} k_{ti} > 0$. Given that there is perfect consumption smoothing, using the budget constraint to solve for consumption yields:

$$c = \left(1 - \frac{1}{1+r}\right)\left(\sum_{t\geq 0}\left(\frac{1}{1+r}\right)^t z_{ti} + k_{0i} - k_i^\infty\right) \tag{A1}$$

Consider what happens if the capital tax rate increases by $d\tau_K > 0$ for $t \geq T$. The present discounted value of all resources, denoted by Y_i for agent i is:

$$Y_i = k_{i0} + \sum_{t=1}^{T}\left(\frac{1}{1+r}\right)^t z_{ti} + \sum_{t\geq T}\left(\frac{1}{1+\bar{r}}\right)^t z_{ti}$$

The change in resources evaluated at $\tau_K = 0$ is:

$$dY_i = \left(\frac{1}{(1+r)}\right)^T \sum_{t\geq T} t\left(\frac{1}{(1+r)}\right)^{t-T+1} z_{ti} d\tau_K \propto \left(\frac{1}{(1+r)}\right)^T d\tau_K$$

Hence, consumption pre-reform will shift down by a factor proportional to $\left(\frac{1}{(1+r)}\right)^T d\tau_K$. From the aggregated budget constraint we have that:

$$k_t^m = (1+r)^t k_0^m - c_0^m(1 + (1+r) + (1+r)^2 + ... + (1+r)^{t-1}) + (z_{t-1}^m + .. + (1+r)^{t-1} z_0^m)$$

Therefore, the change in the aggregate capital stock is:

$$dk_t^m = -dc_0^m \left(\frac{(1+r)^{t-1} - 1}{r} \right)$$

Recall that the change in consumption (from (A1)) is proportional to $\left(\frac{1}{(1+r)}\right)^T d\tau_K$. Hence:

$$dk_t^m \propto -\left(\frac{1}{(1+r)}\right)^T \left(\frac{(1+r)^{t-1} - 1}{r} \right) d\tau_K = -(1+r)^{-T} \left(\frac{(1+r)^{t-1} - 1}{r} \right) d\tau_K$$

Hence:

$$e_{Kt} \propto k_t^m (1+r)^{-T} \left(\frac{(1+r)^{t-1} - 1}{r} \right) d\tau_K$$

Recall that the anticipation elasticity e_K^{ante} is defined as:

$$e_K^{ante} = \frac{\delta}{1+\delta} \sum_{t<T} \left(\frac{1}{1+\delta}\right)^{t-T} e_{Kt} \propto \frac{\delta}{1+\delta} \sum_{t<T} \left(\frac{1}{1+\delta}\right)^{t-T} k_t^m (1+r)^{-T} \left(\frac{(1+r)^{t-1} - 1}{r} \right) d\tau_K$$

Since we have $\delta > r$, $\lim_{T\to\infty} \left(\frac{1+\delta}{1+r}\right)^T = \infty$, which makes the sum above (to which the anticipation elasticity is proportional) converge to infinity when T goes to infinity.

A.2 Optimal Nonlinear Taxes in the Generalized Model

Consider a small reform $\delta T_K(rk)$ in which the marginal tax rate is increased by $\delta\tau_K$ in a small band $[rk^*, rk^* + d(rk^*)]$, as in the proof of Proposition 2. Let us first derive the change in revenue from the reform in any period t. We start from the steady state. Since the reform has to be budget-neutral every period, the change in transfer to the agent will depend on the change in tax revenues at each period.

First, for any capital income rk above capital income rk^*, additional revenue equal to $d(rk^*)\delta\tau_K$ is raised. The total additional tax collected is $(1 - H_K(rk^*))d(rk^*)\delta\tau_K$. Second, for taxpayers in the small band $[rk^*, rk^* + d(rk^*)]$, the change in marginal tax rates generates changes in capital income through two channels. The first channel is a pure substitution effect due to the change $\delta\tau_K$ in marginal tax rates. The second channel is through the shift in capital income along the nonlinear tax schedule, which leads to an additional change in marginal tax rates equal to $dT_i' = T_K''(rk_i)\delta(rk_i, t)$. Let $e_K^c(rk, t)$ be the elasticity in period t at capital income level rk to a small change in the marginal tax rate that i) is unanticipated and occurs at time 0 and ii) lasts for all periods $t \geq 0$. $e_K^c(rk, t)$ is thus a policy elasticity. Formally, $e_K^c(rk, t) = \frac{dk}{d(1-T_K'(rk))}\frac{(1-T_K'(rk))}{k}$ where dk is the total change in capital for the reform described.

Hence the total change in capital income from tax payers in the small band is:

$$\delta(rk, t) = -e_K^c(rk^*, t)rk^* \frac{\delta\tau_K + T_K''(rk^*)\delta(rk, t)}{1 - T_K'(rk^*)}$$

Rearranging this yields:

$$\delta(rk, t) = -e_K^c(t)rk^* \frac{\delta\tau_K}{1 - T_K'(rk^*) + e_K^c(rk^*, t)rk^* T_K''(rk^*)}$$

The total behavioral effect in the small band is hence:

$$-e_K^c(rk^*, t)rk^* \frac{T_K'(rk^*)}{1 - T_K'(rk^*) + e_K^c(rk^*, t)rk^* T_K''(rk^*)} h_K(rk^*, t)\delta\tau_K d(rk^*)$$

Because we start from the steady state, the density is constant across time and $h_K(rk^*, t) = h_K(rk^*), \quad \forall t$.

Third, for taxpayers above rk^*, there is a change in the average tax liability. Let $\eta(rk, t)$ be the elasticity of capital income in period t at capital income level rk for a small change in virtual income that i) is unanticipated and occurs at time 0 and ii) lasts for all periods $t \geq 0$. $\eta(rk, t)$ is thus also policy elasticity.

There are again two channels: the direct impact of the tax change, equal to $\eta(rk, t)\delta\tau_K d(rk^*)$ and the indirect effect due to the move along the nonlinear tax schedule, which increases marginal tax rates by $dT_i' = T_K''(rk)\delta(rk, t)$. The total effect is hence:

$$\delta(rk, t) = \eta(rk, t) \frac{\delta\tau_K d(rk^*)}{1 - T_K'(rk) + rk e_K^c(rk, t) T_K''(rk)}$$

Integrating over all taxpayers with incomes above the small band, the total tax revenue raised through this third component is:

$$\delta\tau_K d(rk^*)\int_{rk^*}^{\infty} -\eta(s,t)\frac{T_K'(s)}{1-T_K'(s)+se_K^c(s,t)T_K''(s)}h_K(s)d(s)$$

The total change in revenue $dG(t)$ is:

$$d(rk^*)\delta\tau_K\cdot[(1-H_K(rk^*))-e_K^c(rk^*,t)rk^*\frac{T_K'(rk^*)}{1-T_K'(rk^*)+rk^*e_K^c(rk^*,t)T_K''(rk^*)}h_K(rk^*) \quad \text{(A2)}$$
$$+\int_{rk^*}^{\infty} -\eta(s,t)\frac{T_K'(s)}{1-T_K'(s)+se_K^c(s,t)T_K''(s)}h_K(s)d(s)]$$

For agents below the small band, there is no change in the tax paid, but they benefit from the lump-sum rebate in revenue dG. For them $dT_i(t) = dG(t)$. Hence, the welfare impact for agents with $rk_i \leq rk^*$ is $\int_{i:rk_i<rk^*} g_i dG(t)di$. On the other hand, above the small band, agents receive the lump-sum increase in revenue $dG(t)$ but also pay an extra tax $\delta\tau_K d(rk^*)$, so that for them $dT_i(t) = (-\delta\tau_K d(rk^*) + dG(t))$. Hence, the welfare effect on agents above the small band is: $\int_{i:rk_i\geq rk^*} g_i(-\delta\tau_K d(rk^*) + dG(t))$. Thus the total change in welfare is:

$$\int_i \delta_i g_i \int_t dG(t)e^{-\delta_i t} - \int_{i:rk_i\geq rk^*} \delta_i g_i \int_t \delta\tau_K d(rk^*)e^{-\delta_i t}$$

Welfare weights g_i do not depend on time is because we start from a steady state (even if they are standard social welfare weights with $g_i = \omega_i u_{ci}$). We normalize $\int_i g_i = 1$.

Substituting for the change in revenue from (A2), the change in welfare is:

$$d(rk^*)\delta\tau_K\cdot[(1-H_K(rk^*))-\int_t e_K^c(rk^*,t)rk^*\frac{T_K'(rk^*)}{1-T_K'(rk^*)+rk^*e_K^c(rk^*,t)T_K''(rk^*)}h_K(rk^*)\int_i \delta_i g_i e^{-\delta_i t}didt$$
$$-\int_t\int_{rk^*}^{\infty}\eta(s,t)\frac{T_K'(s)}{1-T_K'(s)+se_K^c(s,t)T_K''(s)}h_K(s)ds\int_i \delta_i g_i e^{-\delta_i t}didt - \int_{i:rk_i\geq rk^*} g_i di]$$

Thus at the optimum, the optimal marginal tax schedule is characterized by the differential equation:

$$(1 - H_K(rk^*)) - \int_t e^c_K(rk^*,t)rk^* \frac{T'_K(rk^*)}{1 - T'_K(rk^*) + rk^* e^c_K(rk^*,t) T''_K(rk^*)} h_K(rk^*) \int_i \delta_i g_i e^{-\delta_i t} didt$$
$$- \int_t \int_{rk^*}^{\infty} \eta(s,t) \frac{T'_K(s)}{1 - T'_K(s) + s e^c_K(s,t) T''_K(s)} h_K(s) ds \int_i \delta_i g_i e^{-\delta_i t} didt - \int_{i: rk_i \geq rk^*} g_i di = 0 \quad \text{(A3)}$$

A.3 Optimal Taxation with Horizontal Equity Concerns.

In this section, we formally consider optimal capital and labor taxation under horizontal equity concerns.

As derived in Section 2.3.4, the optimal revenue-maximizing rates are: $\tau^R_L = \frac{1}{1+e_L}$ and $\tau^R_K = \frac{1}{1+e_K}$. Without loss of generality, we suppose that capital is more elastic so that $\tau^R_K < \tau^R_L$. The optimal linear comprehensive tax on income is, as derived in (16):

$$\tau_Y = \frac{1 - \bar{g}_Y}{1 - \bar{g}_Y + e_Y} \quad \text{with} \quad \bar{g}_Y = \frac{\int_i g_i \cdot y_i}{\int_i y_i}$$

Suppose that the distribution of capital and labor income is dense enough, so that at every total income level $y = rk + z$, there are agents with $y = rk$ (capital income only) and $y = z$ (labor income only).

Generalized social welfare weights that capture horizontal equity concerns are such that:

(i) If $\tau_L = \tau_K$, then g_i are standard, for instance $g_i = u_{ci}$ for all agents. Any reform that changes taxes should put zero weight on those who after the reform are such that $\tau_L z_i + \tau_K r k_i < \max_j \{\tau_L z_j + \tau_K r k_j | z_j + r k_j = z_i + r k_i\}$, i.e., on those who pay less taxes at a given total income $y = rk_i + z_i$, or, equivalently, have the highest disposable income and consumption at any income. This means that if labor taxes are increased, $g_i = 0$ for those with any positive capital income at each total income level. Conversely, increasing capital taxes will yield $g_i = 0$ for those individuals with some labor income at each total income level.

(ii) If $\tau_L > \tau_K$, then all the social welfare weights are concentrated on those with $\tau_L z_i + \tau_K r k_i > \max_j \{\tau_L z_j + \tau_K r k_j | z_j + r k_j = z_i + r k_i\}$, i.e., on those agents with only labor income. Conversely, if $\tau_L < \tau_K$, all the social welfare weights are on agents with only capital income.

Suppose that, starting from a situation with $\tau_L = \tau_K$ we introduce a small tax break on capital income, $d\tau_K < 0$. Capital income earners now get an unfair advantage and all the weight is concentrated on those with no capital income (equivalently, everyone with $k_i > 0$ receives a weight $g_i = 0$). As a result, a small tax break on capital can only be optimal if it raises tax revenue and, hence, allows to lower the tax rate on labor income as well. This can only occur if $\tau_Y > \tau_K^R$, i.e., the optimal comprehensive tax rate is above the revenue-maximizing rate on capital income.

Proposition 1. ***Optimal labor and capital taxation with horizontal equity concerns.***

(i) If $\tau_Y \leq \tau_K^R$, taxing labor and capital income at the same comprehensive rate $\tau_L = \tau_K = \tau_Y$ is the unique optimum.

(ii) If $\tau_Y > \tau_K^R$, a differential tax system with the capital tax rate set to the revenue maximizing rate $\tau_K = \tau_K^R < \tau_L$ (with both τ_K and τ_L smaller than τ_Y) is the unique optimum.

Proof. Let us consider the two cases in turn.

(i) If $\tau_Y \leq \tau_K^R$.

To see why $\tau_L = \tau_K = \tau^*$ is an equilibrium, suppose that we tried to lower the tax rate on capital income. Then, all the weight will concentrate on people with only labor income, which will then in turn make it optimal to increase the tax on capital again.

This equilibrium is unique. There is no other equilibrium with equal taxes on capital and labor that can raise more revenue with a lower tax rate, by definition of τ_Y as the optimal rate on comprehensive income. There is also no equilibrium with non-equal tax rates on capital and labor. Suppose that we tried to set (without loss of generality) $\tau_K < \tau_L$. Then to raise enough revenue we would require that $\tau_K < \tau_Y < \tau_L$. Since capital owners are now advantaged, all the social welfare weight concentrates on people with only labor income. Since then a fortiori $\tau_K < \tau_K^R$, increasing τ_K would mean that more revenue would be raised, which would allow us to lower τ_L, which is good since all weight is on people with only labor income.

(ii) If $\tau_Y > \tau_K^R$.

In this case, the equilibrium has $\tau_K = \tau_K^R < \tau_Y$ and $\tau_Y > \tau_L > \tau_K^R$. Clearly this is an equilibrium since we cannot decrease τ_L without losing revenue and we cannot raise more revenue through τ_K (since it is already set at the revenue-maximizing rate for the capital tax base). In addition, we cannot decrease τ_K further without increasing τ_L, which is not desirable since it would benefit people capital income earners, who already receive a weight of zero.

This equilibrium is also unique. If we set $\tau_L = \tau_K$ equal, we should set them equal to τ_Y which is the optimal tax rate on comprehensive income. But then, since τ_K is now above its revenue maximizing rate, we could lower both τ_K and τ_L without losing revenues, so this would not be an equilibrium. On the other hand, as long as we set $\tau_K < \tau_L$, capital income earners get zero weight and the only possibility is to go all the way to $\tau_K = \tau_K^R$ since only people with only labor income have a non-zero weight.

□

As a result, horizontal equity concerns will be a force pushing towards the comprehensive income tax system derived in Section 2.3.4. In the text, we provided an efficiency argument in favor of a tax on comprehensive income (based on income shifting opportunities) while the argument here is based on equity considerations. With horizontal equity preferences, deviations from a comprehensive income tax system can only be justified if they raise more revenue and generate a Pareto-improvement, which drastically reduces the scope for them. In **?** we argue that this is akin to a generalized Rawlsian principle whereby discrimination against some groups (e.g., capital owners versus labor providers) is only permissible if it makes the group discriminated against better off, i.e., if it generates a Pareto improvement.

A.3.1 Horizontal Equity with Nonlinear Taxation

The same reasoning as for linear taxation with horizontal equity also applies to nonlinear taxes. Starting from a comprehensive tax system $T_Y(z + rk)$ as derived in Section 2.3.4, lowering the tax rate on capital income, conditional on a given total income level, will generate a horizontal inequity and concentrate all social weight on those with no capital income conditional on that total income level. Such a preferential tax break for capital income earners will only be acceptable if it generates more revenue and allows to lower the tax rate on labor income as well. We show this below.

Formally, suppose that we start from the optimal tax on comprehensive income, $T_Y(rk+z)$, as derived in Section 2.3.4 which does not discriminate between capital and labor income conditional on total income. We say that a tax system unambigously favors capital (respectively, labor) at income level $y = rk + z$, if for any (rk, z) such that $y = rk + z$, and any $\varepsilon \in]0, z]$, $T_Y(rk, z) > T(rk + \varepsilon, z - \varepsilon)$ (having more capital income, conditional on a given total income leads to lower taxes). (Note that it may be the case that a tax system favors capital only at some y levels or only at some rk, z ranges..)

Denote a change in the tax by $\delta T(rk, z)$.

A deviation $\delta T(rk, z)$ is said to introduce horizontal inequity, if, starting from a comprehensive tax system $T_Y(rk+z)$, the resulting tax system $T_Y(z+rk)+\delta T(rk.z)$ cannot be expressed as $\tilde{T}_Y(rk + z)$ for some function $\tilde{T}_Y$.

With nonlinear taxes, we can again define the generalized social welfare weights as follows.

i) If there is a comprehensive tax $T_Y(z + rk)$, then everybody has standard weights, such as, for instance, $g_i = u_{ci}$. For any deviation $\delta T(rk, z)$ that introduces horizontal inequity, the weights concentrate on the agents who pay the highest tax at a given total income level, i.e., on those with $T_Y(z_i + rk_i) + \delta T(rk_i, z_i) = \max_j \{T_Y(z_j + rk_j) + \delta T(rk_j, z_j)|z_j + rk_j = rk_i + z_i\}$ (which is equivalent to putting all the weight on the agent(s) with lowest disposable income at any total income level).

Hence, the weights also need to depend on $\delta T(z, rk)$, the direction of the tax reform.

ii) If the tax is such that $T(rk, z)$ cannot be expressed as $\tilde{T}_Y(rk + z)$ for some function $\tilde{T}_Y$, then the weights concentrate on those with
$T(z_i, rk_i) = \max_j \{T(z_j, rk_j)|z_j + rk_j = rk_i + z_i\}$, i.e., on the agents which pay the highest tax (equivalently, have the lowest disposable income) conditional on total income.

Equilibria:

Suppose that, at the comprehensive tax rate, no small reform $\delta T(rk.z)$ that introduces horizontal equity and favors capital (according to our definitions above) can increase total tax revenues, i.e., for all $\delta T(rk, z)$ that favor capital and introduce horizontal inequity, the alternative tax system $\tilde{T}(rk, z) = T(rk + z) + \delta T(rk, z)$ is such that:

$$\int_i T_Y(rk_i(T) + z_i(T))di > \int_i \tilde{T}_Y(rk_i(\tilde{T}) + z_i(\tilde{T}))di$$

where naturally, the choices $z_i(T)$ and $k_i(T)$ depend on the tax system T. Then the unique equilibrium has the comprehensive tax system in place, as derived in 2.3.4. No horizontal inequity can be an equilibrium unless it introduces a Pareto improvement.

Suppose on the other hand that if the revenue maximizing tax rate on capital, $T_K^R(rk)$ were implemented, and a labor income tax $T_L(z)$ was used to complement it, more revenue could be raised than with the tax on comprehensive income $T_Y(rk, z)$ and the tax burden on all agents would be lower than under the comprehensive income tax. Then, the optimum is to set differential taxes on capital and labor income, with the capital tax at its optimal revenue-maximizing schedule. Horizontal inequity is an equilibrium because it generates a Pareto improvement.

A.4 Progressive Consumption Taxes

The progressive consumption tax is defined on an exclusive basis as $t_C(.)$ such that

$$\dot{k} = \bar{r}k + z - [c + t_c(c)]$$

Equivalently, we can again define the inclusive consumption tax $T_C(y)$ on pre-tax resources y devoted to consumption such that $c + t_c(c) = y$ is equivalent to $c = y - T_C(y)$, i.e., $y \to y - T_C(y)$ is the inverse function of $c \to c + t_c(c)$ and hence $1 + t_C' = 1/(1 - T_C')$.

The case of a progressive consumption tax is most easily explained with inelastic labor income (possibly heterogenous across individuals). Real wealth k^r in the presence of the progressive consumption tax is:

$$k^r(k) = k - \frac{T_C(\bar{r}k + z) - T_C(z)}{\bar{r}}$$

Recall that real wealth is defined as nominal wealth adjusted for the price of consumption. There are to see why the above is the right expression. First, wealth k provides an income stream $\bar{r}k$ which translates into extra permanent consumption equal to the income minus the tax paid on the extra consumption $\bar{r}k-[T_C(\bar{r}k+z)-T_C(z)]$ which can be capitalized into wealth k^r by dividing by $\bar{r}$. If labor income is heterogeneous across agents, then $k^r(k,z)$ should also be indexed by z. Another way to see this is to ask what the capital k^r would be that would yield the same disposable income as the nominal capital under the consumption tax. Disposable income in terms of real capital k^r is $\bar{r}k^r - T_C(z)$. Disposable income expressed in terms of nominal capital is: $\bar{r}k - T_C(\bar{r}k+z)$. These two must be equal, which yields the expression for k^r above. k^r has three natural properties: with no consumption tax, real and nominal wealth are equal, $dk^r/dk = 1 - T'_C$, i.e., and extra dollar of nominal wealth is worth $1 - T'_C$ in real terms, and $k^r(0) = 0$.

In that case, we have in steady-state

$$c = \bar{r}k + z - T_C(\bar{r}k + z) = \bar{r}k^r + z - T_C(z)$$

and the first order condition for utility maximization is $a'_i(k^r) = \delta - \bar{r}$. Hence, real capital is chosen to satisfy the same condition as nominal capital when there is no consumption tax. Put differently, any consumption tax will be undone by agents in terms of their savings and will have no effect on the real value of their wealth held (and, hence, by definition of the real wealth, on their purchasing power). Hence, the consumption tax is equivalent to a tax on labor income only.

The equivalence is not exact with elastic labor supply, as in that case, the marginal consumption tax depends on the labor choice and the first-order condition for labor income is $h'_i(z) = 1 - T'_C(\bar{r}k + z) + a'_i(k^r)[T'_C(\bar{r}k + z) - T'_C(z)]/\bar{r}$.