A Online Appendix

A.1 Proof of No Ordinal Restrictions in Assumption 2

Note that Assumption 2 is *not* required to hold at all levels of the gross interest rate, only the technological rate of return R, i.e. the return on storage without taxation.

Start with concave preferences $U(c_0, c_1; \theta)$ that satisfy Assumption 1. We construct a renormalization of preferences by applying the transformation $f^g(U(c_0, c_1; \theta); \theta)$. This renormalization ensures that Assumption 2 holds, and preserves concavity.

The functions $f^g(V; \theta)$ are constructed as follows. Fix θ^*. For any function g, for all I and θ, define $f^g(V; \theta)$ by

$$f^g_V(V(I, R; \theta); \theta)V_I(I, R; \theta) = g'(V(I, R; \theta^*))V_I(I, R; \theta^*),$$

or

$$f^g_V(V(I, R; \theta); \theta) = g'(V(I, R; \theta^*))\frac{V_I(I, R; \theta^*)}{V_I(I, R; \theta)}.$$

We can pick g so that $f^g_V(V; \theta)$ is decreasing in V for each θ so that concavity is preserved.

A.2 Proof of Proposition 1

Case (a) is immediate. We now prove case (b); case (c) is symmetric. Define $C_0(c_1, v; \theta)$ as $v = U^p(C_0(c_1, v; \theta), c_1; \theta)$. The planning problem is

$$\min \int \left(C_0(c_1(\theta), v(\theta); \theta) + \frac{1}{R}c_1(\theta) \right) f(\theta)d\theta$$

subject to c_1 non-decreasing and

$$\dot{v}(\theta) = U^p_\theta(C_0(c_1(\theta), v(\theta); \theta), c_1(\theta); \theta),$$

$$\int \lambda_\theta v(\theta)f(\theta)d\theta \geq V.$$

We study the relaxed planning problem

$$\min \int \left(C_0(c_1(\theta), v(\theta); \theta) + \frac{1}{R}c_1(\theta) \right) f(\theta)d\theta$$

subject to c_1 non-decreasing and

$$\dot{v}(\theta) = U^p_\theta(C_0(c_1(\theta), v(\theta); \theta), c_1(\theta); \theta) + r(\theta),$$

$$r(\theta) \geq 0,$$

$$\int \lambda_\theta v(\theta) f(\theta) d\theta \geq V.$$

The original problem imposes $r(\theta) = 0$ for all θ, but here we allow for $r(\theta) \geq 0$. Our first goal is to show that an interior solution to the relaxed problem features $r(\theta) = 0$ and, thus, coincide with the original planning problem.

Adapting Theorem 3.1 in Hellwig (2009) we form the Hamiltonian

$$H = \lambda_\theta v f(\theta) - \gamma \left(C_0(c_1, v; \theta) + \frac{1}{R} c_1 \right) f(\theta) + \mu(U_\theta^p(C_0(c_1, v; \theta), c_1, \theta) + r) + \chi q$$

where γ should be thought of the inverse of the multiplier ν on the promise keeping constraint. We obtain the following necessary conditions for an interior optimum:

$$\dot{\chi}(\theta) = \gamma \frac{1}{R} f(\theta) + \gamma f(\theta) C_{0,c_1}(\theta) - \mu(\theta) \left(U^p_{\theta,c_0}(\theta) C_{0,c_1}(\theta) + U^p_{\theta,c_1}(\theta) \right),$$

$$\dot{\mu}(\theta) = -\lambda_\theta f(\theta) + \gamma f(\theta) C_{0,v}(\theta) - \mu(\theta) U^p_{\theta,c_0}(\theta) C_{0,v}(\theta),$$

$$\mu(\underline{\theta}) = \mu(\bar{\theta}) = \chi(\underline{\theta}) = \chi(\bar{\theta}) = 0,$$

$$\mu(\theta) \leq 0 \quad \mu(\theta) r(\theta) = 0,$$

$$\chi(\theta) \leq 0,$$

$$\int \chi(\theta) dc_1(\theta) = 0.$$

The relaxed problem solves the original problem. To show that a solution to the relaxed problem features $r(\theta) = 0$ and, thus, coincide with the original planning problem, we argue by contradiction. Thus, assume that $r(\theta) > 0$ on a positive measure of points θ.

We claim that if we can find a set E of positive measure such that for all $\theta \in E$ we have

$$r(\theta) > 0,$$

$$\mu(\theta) = 0,$$

$$\dot{\mu}(\theta) = 0,$$

and $c_1(\cdot)$ strictly increasing at θ then we are done. First we show that we can find such a set under the assumption that $r(\theta) > 0$ on a set of positive measure. Let E be a set of positive measure on which $r(\theta) > 0$. Then by the necessary conditions we know that $\mu(\theta) = 0$ on E. Since μ, v, χ are all absolutely continuous they are differentiable almost everywhere. Thus we can assume without loss of generality that μ, χ and v are all differentiable on E. Since $\mu(\theta) \leq 0$ we know that $\dot{\mu}(\theta) = 0$ on E. If we have that $c_1(\theta)$ is strictly increasing on a positive measure subset of E then we are done. Thus suppose not, so that c_1 is not strictly increasing at almost every of E. Without loss of generality we can suppose that c_1 is not strictly increasing at every point of E. In other words given $\theta \in E$ we can choose ϵ small enough so that $c_1(\theta - \epsilon) = c_1(\theta + \epsilon)$. Now by our preliminary fact we know that E contains an accumulation point, call it θ_0. By the preceding argument there exists an $\epsilon > 0$ so that c_1 is constant on the interval $(\theta_0 - \epsilon, \theta_0 + \epsilon)$. Let $\overline{c}_1 = c_1(\theta_0)$. Now since $U_{c_0} > 0$ and v is differentiable at θ_0 the implicit function theorem says that $c_0(\theta)$ defined by

$$v(\theta) = U^p(c_0(\theta), \overline{c}_1; \theta)$$

is differentiable at θ_0. Then

$$\dot{v}(\theta_0) = U^p_{c_1}(c_0(\theta_0), \overline{c}_1; \theta)\dot{c}_0(\theta_0) + U^p_\theta(c_0(\theta_0), \overline{c}_1; \theta).$$

We also know that

$$\dot{v}(\theta_0) = U^p_\theta(c_0(\theta_0), \overline{c}_1; \theta) + r(\theta_0) > U^p_\theta(c_0(\theta_0), \overline{c}_1; \theta).$$

Since $r(\theta_0) > 0$. Since $U^P_{c_0} > 0$ this implies that $\dot{c}_0(\theta_0) > 0$. Since θ_0 is an accumulation point of E there exists a sequence $\theta_n \in E$ with $\theta_n \to \theta_0$. We can further suppose that we have either $\theta_n \nearrow \theta_0$ or $\theta_n \searrow \theta_0$. We take the case $\theta_n \nearrow \theta_0$. Using the fact that for all $\theta \in E$ we have $\mu(\theta) = \dot{\mu}(\theta) = 0$ we have that

$$\gamma = \lambda(\theta) U^p_{c_0}(c_0(\theta), c_1(\theta); \theta).$$

Using the fact that $\dot{c}_0(\theta_0) > 0$ we know that for all large enough n we have $c_0(\theta_n) < c_0(\theta_0)$ and $c_1(\theta_n) = c_1(\theta)$. But then using the concavity of the utility function, the fact that $U^p_{c_0,\theta} < 0$ and λ_θ is decreasing we see that

$$\begin{aligned}\gamma &= \lambda_{\theta_0} U^p_{c_0}(c_0(\theta_0), c_1(\theta_0); \theta_0), \\ &\leq \lambda_{\theta_0} U^p_{c_0}(c_0(\theta_0), c_1(\theta_0); \theta_n), \\ &< \lambda_{\theta_0} U^p_{c_0}(c_0(\theta_n), c_1(\theta_n); \theta_n), \\ &\leq \lambda_{\theta_n} U^p_{c_0}(c_0(\theta_n), c_1(\theta_n); \theta_n),\end{aligned}$$

which is a contradiction. Thus on the original set E we must have had a positive measure set of points A such that c_1 was strictly increasing at these points. Consider the new set $E' = E \cap A$ which also has positive measure. Then for all $\theta \in E'$ we have that

$$\begin{aligned} r(\theta) &> 0, \\ \mu(\theta) &= 0, \\ \dot{\mu}(\theta) &= 0, \end{aligned}$$

the function $c_1(\theta)$ is strictly increasing and μ, v, χ are all differentiable. Since c_1 is increasing at θ we know that we must have $\chi(\theta) = 0$. But since χ is differentiable at θ and $\chi(\theta) = 0$ and $\chi \leq 0$ it must be that $\dot{\chi}(\theta) = 0$.

From here onwards we restrict to points $\theta \in E'$. Using the fact that $\mu(\theta) = \dot{\mu}(\theta) = \dot{\chi}(\theta) = 0$ tells us that

$$-\lambda_\theta f(\theta) + \gamma C_{0,v}(\theta) f(\theta) = 0.$$

and

$$\gamma \frac{1}{R} f(\theta) + \gamma f(\theta) C_{0,c_1}(\theta) = 0.$$

Rearranging and using the definition of $C_0(c_1, v; \theta)$, these two equations are equivalent to

$$\gamma = \lambda_\theta U^p_{c_0}(c_0(\theta), c_1(\theta); \theta), \tag{1}$$

$$R = \frac{U^p_{c_0}(c_0(\theta), c_1(\theta); \theta)}{U^p_{c_1}(c_0(\theta), c_1(\theta); \theta)}. \tag{2}$$

The second equation tells us that we can write

$$v(\theta) = V^p(I(\theta), R, \theta)$$

for some $I(\theta)$ since $c_0(\theta)$ and $c_1(\theta)$ are chosen as they would be in the parent's optimal problem. Now we claim that $I(\theta)$ is decreasing on E'. This follows from the fact that

$$\frac{\gamma}{\lambda(\theta)} = U^p_{c_0}(c_0(\theta), c_1(\theta); \theta) = V^p_I(I(\theta), R; \theta)$$

and the fact that V is concave and λ is decreasing. Since E' has positive measure it contains a limit point θ. Suppose that there exists a sequence $\theta_n \in E$ with $\theta_n \searrow \theta$. The case $\theta_n \nearrow \theta$ is symmetric. Then since v is differentiable at θ we know that

$$\dot{v}(\theta) = \lim_{n\to\infty} \frac{v(\theta_n) - v(\theta)}{\theta_n - \theta}$$

since V is differentiable in I and θ we see that

$$\begin{aligned}v(\theta_n) - v(\theta) &= V^p(I(\theta_n), R; \theta_n) - V^p(I(\theta), R; \theta) \\ &= V_I^p(I(\theta), R; \theta)(I(\theta_n) - I(\theta)) + V_\theta^p(I(\theta), R; \theta)(\theta_n - \theta) + \varepsilon(\theta_n - \theta)\end{aligned}$$

where the $\varepsilon(\theta_n - \theta)$ is a second order error term so that $\frac{\varepsilon(\theta_n-\theta)}{\theta_n-\theta} \to 0$. Thus

$$\begin{aligned}\dot{v}(\theta) &= \lim_{n\to\infty} \frac{V_I^p(I(\theta), R; \theta)(I(\theta_n) - I(\theta)) + V_\theta^p(I(\theta), R; \theta)(\theta_n - \theta)}{\theta_n - \theta} \\ &= V_\theta^p(I(\theta), R; \theta) + V_I^p(I(\theta), R; \theta) \cdot \lim_{n\to\infty} \frac{(I(\theta_n) - I(\theta))}{\theta_n - \theta} \\ &\leq V_\theta^p(I(\theta), R; \theta)\end{aligned}$$

since $I(\theta_n) - I(\theta) \leq 0$. Thus we see that $\dot{v}(\theta) \leq V_\theta^p(I(\theta), R; \theta)$. But since

$$\dot{v}(\theta) = V_\theta^p(I(\theta), R; \theta) + r(\theta) = U_{c_0}^p(c_0(\theta), c_1(\theta); \theta) + r(\theta)$$

and $r(\theta) > 0$ we have a contradiction. Thus it must have been that $r = 0$ almost surely so that the solution to the relaxed problem coincides with the solution to the original problem.

The relaxed problem features positive taxes. In the proof, we will make repeated use of the fact that over an interval where there is bunching, the tax rate τ is increasing. This is a direct consequence of the single crossing condition in Assumption 1.

Since μ, v, χ are all absolutely continuous they are differentiable almost everywhere. Thus there is a full measure set Ω of θ such that μ, χ and v are all differentiable. Consider $\theta \in \Omega$.

Suppose that c_1 is strictly increasing at θ. Then $\chi(\theta) = \dot{\chi}(\theta) = 0$, and we get using equation 4 together with the fact that $\mu(\theta) \leq 0$ and single crossing that $\tau(\theta) \geq 0$.

Now suppose that c_1 is not strictly increasing at θ. Consider the greatest interval around θ over which c_1 is constant. Then the tax rate τ is increasing over this interval so that it is comprised between its limit values at the bounds θ_l and θ_h of the interval.

If $\theta_l > \underline{\theta}$, then the function c_1 must be strictly increasing at θ_l so that $\tau(\theta_l) \geq 0$. If c_1 is continuous at θ_l, then so are c_0 and τ. We can conclude that $\tau(\theta) \geq 0$. Suppose now that c_1 is not continuous at θ_l. If $\lim_{\tilde{\theta}\to\theta_l^+} \tau(\tilde{\theta}) \geq 0$, then we have $\tau(\theta) \geq 0$. Otherwise $\lim_{\tilde{\theta}\to\theta_l^+} \tau(\tilde{\theta}) < 0$, from which we derive a contradiction by constructing a new allocation that satisfies the constraints of the relaxed planning problem but achieves lower cost. We first construct a new allocation $(\hat{c}_0, \hat{c}_1)$ which coincides with the old one except at points $\tilde{\theta} \in (\theta_l, \theta_h)$ such that $\tau(\tilde{\theta}) < 0$, in which case we pick $(\hat{c}_0(\tilde{\theta}), \hat{c}_1(\tilde{\theta}))$ so that $U^p(\hat{c}_0(\tilde{\theta}), \hat{c}_1(\tilde{\theta}); \tilde{\theta}) = U^p(c_0(\tilde{\theta}), c_1(\tilde{\theta}); \tilde{\theta})$

and $R = \frac{U^p_{c_0}(\hat{c}_0(\tilde{\theta}),\hat{c}_1(\tilde{\theta});\tilde{\theta})}{U^p_{c_1}(\hat{c}_0(\tilde{\theta}),\hat{c}_1(\tilde{\theta});\tilde{\theta})}$ so that $\hat{\tau}(\tilde{\theta}) = 0$. Then define an ironed version of this allocation by setting $\hat{\hat{c}}_1(\tilde{\theta}) = \hat{c}_1(\phi(\tilde{\theta}))$ and $\hat{\hat{c}}_0(\tilde{\theta}) = \hat{c}_0(\phi(\tilde{\theta}))$ where $\phi(\tilde{\theta}) = \arg\max_{\tilde{\theta}' < \tilde{\theta}} \hat{c}_1(\tilde{\theta}')$. Then this allocation satisfies the constraints of the relaxed planning problem but has a lower cost.

Suppose now that $\theta_l = \underline{\theta}$. We have $\chi(\underline{\theta}) = 0$. Then for every $\epsilon > 0$, we can find $\tilde{\theta}_\epsilon$ in $[\underline{\theta}, \underline{\theta} + \epsilon) \cap \Omega$, such that $\dot{\chi}(\tilde{\theta}) \leq 0$. Because $\lim_{\epsilon \to 0} \chi(\tilde{\theta}_\epsilon) = 0$, we conclude using equation (3) that $\lim_{\epsilon \to 0} \tau(\tilde{\theta}_\epsilon) \geq 0$. Since τ is increasing on $(\underline{\theta}, \theta_h)$, this allows us to conclude that $\tau(\theta) \geq \lim_{\epsilon \to 0} \tau(\tilde{\theta}_\epsilon) \geq 0$.

A.3 Derivation of Optimal Tax Formula and Proposition 2

Define $C_0(c_1, v; \theta)$ as $v = U^p(C_0(c_1, v; \theta), c_1; \theta)$. We have $C_{0,c_1} = -\frac{U^p_{c_1}}{U^p_{c_0}}$, and $C_{0,v} = \frac{1}{U^p_{c_0}}$. We adapt Theorem 3.1 in Hellwig (2009). We form the Hamiltonian

$$H = (\lambda_\theta v + \alpha_\theta U^c(c_1)) f(\theta) - \gamma \left(C_0(c_1, v; \theta) + \frac{1}{R} c_1 \right) f(\theta) + \mu U^p_\theta(C_0(c_1, v; \theta), c_1, \theta) + \chi q$$

We have the following necessary conditions:

$$\dot{\chi}(\theta) = -\alpha_\theta U^c_{c_1}(\theta) f(\theta) + \gamma \frac{1}{R} f(\theta) + \gamma f(\theta) C_{0,c_1}(\theta) - \mu(\theta) \left(U^p_{\theta,c_0}(\theta) C_{0,c_1}(\theta) + U^p_{\theta,c_1}(\theta) \right),$$
$$\dot{\mu}(\theta) = -\lambda_\theta f(\theta) + \gamma f(\theta) C_{0,v}(\theta) - \mu(\theta) U^p_{\theta,c_0}(\theta) C_{0,v}(\theta),$$

$$\mu(\underline{\theta}) = \mu(\bar{\theta}) = \chi(\underline{\theta}) = \chi(\bar{\theta}) = 0,$$

$$\chi(\theta) \leq 0,$$

and the complementary slackness condition

$$\int \chi(\theta) dc_1(\theta) = 0.$$

Using the definition for

$$\tau(\theta) = R \frac{U^p_{c_1}(\theta)}{U^p_{c_0}(\theta)} - 1$$

we can rewrite the first equation as

$$\gamma \tau(\theta) = -R \frac{\dot{\chi}(\theta)}{f(\theta)} - \alpha_\theta R U^c_{c_1}(\theta) - \frac{\mu(\theta)}{f(\theta)} R U^c_{c_1}(\theta) \left(\frac{U^p_{\theta,c_1}(\theta)}{U^p_{c_1}(\theta)} - \frac{U^p_{\theta,c_0}(\theta)}{U^p_{c_0}(\theta)} \right). \tag{3}$$

The result in Proposition 2 follows immediately from the fact that $\mu(\theta)$ is zero at the extremes. As long as $c_1(\theta)$ is strictly increasing, then we have $\chi(\theta) = 0$ and $\dot{\chi}(\theta) = 0$ so that using $\gamma = \frac{1}{\nu}$, we have

$$\tau(\theta) = -\nu\alpha_\theta R U^c_{c_1}(\theta) - \nu\frac{\mu(\theta)}{f(\theta)} R U^c_{c_1}(\theta)\left(\frac{U^p_{\theta,c_1}(\theta)}{U^p_{c_1}(\theta)} - \frac{U^p_{\theta,c_0}(\theta)}{U^p_{c_0}(\theta)}\right). \tag{4}$$

A.4 Proof of Proposition 3

We fix the weights α_θ, and solve the following system

$$\gamma\frac{1}{U^p_{c_0}(\theta)} = -\alpha_\theta\frac{U^c_{c_1}(\theta)}{U^p_{c_1}(\theta)} + \gamma\frac{1}{R}\frac{1}{U^p_{c_1}(\theta)}, \tag{5}$$

$$U^p(\theta) = v(\theta), \tag{6}$$

$$U^p_\theta(\theta) = \dot{v}(\theta). \tag{7}$$

Given $v(\theta)$, equations 5 and 6 pin down $c_0(\theta)$ and $c_1(\theta)$. Equation 7 can then be seen as a differential equation in $v(\theta)$.

If the solution of this system is such that $c_1(\theta)$ is increasing in θ, then the corresponding allocation is incentive compatible, and solves the planning problem for parental weights λ_θ given by

$$\gamma\frac{1}{U^p_{c_0}(\theta)} = \lambda_\theta.$$

For $\alpha_\theta = \alpha$ constant, we know that the allocation is incentive compatible. Indeed the allocation can be constructed by confronting agents with a nonlinear tax on bequests given by

$$T'\left(\frac{c_1 - I_1}{R}\right) = -\nu\alpha R U^c_{c_1}(c_1).$$

The bequest tax T is convex, and hence the resulting budget set is concave. The corresponding allocation is incentive compatible by construction.

A.5 Proof of Proposition 4

The planning problem is (normalizing $\lambda_\theta = 1$)

$$\min \int \left(C_0(c_1(\theta), v(\theta); \theta) + \frac{1}{R}c_1(\theta)\right) f(\theta) d\theta$$

subject to c_1 non-decreasing and

$$c_1(\theta) \geq \underline{c},$$

$$\dot{v}(\theta) = U^p_\theta(C_0(c_1(\theta), v(\theta); \theta), c_1(\theta); \theta),$$

$$\int v(\theta)dF(\theta) \geq V.$$

As in the proof of Proposition 1, we study the relaxed planning problem

$$\min \int \left(C_0(c_1(\theta), v(\theta); \theta) + \frac{1}{R} c_1(\theta) \right) f(\theta) d\theta$$

subject to c_1 non-decreasing and

$$c_1(\theta) \geq \underline{c},$$

$$\dot{v}(\theta) = U^p_\theta(C_0(c_1(\theta), v(\theta); \theta), c_1(\theta); \theta) + r(\theta),$$

$$r(\theta) \geq 0,$$

$$\int v(\theta)dF(\theta) \geq V.$$

The original problem imposes $r(\theta) = 0$ for all θ, but here we allow for $r(\theta) \geq 0$. The necessary conditions can be derived by adapting Theorem 3.1 in Hellwig (2009). Indeed, his setup explicitly allows for a constraint such as $c_1(\theta) \geq \underline{c}$.

We claim that a solution to the relaxed problem must feature $r(\theta) = 0$ and, thus, coincide with the original planning problem. The proof of this claim is essentially identical to that laid out in the proof of Proposition 1. The presence of the new constraint $c_1(\theta) \geq \underline{c}$ does not change the key arguments involved. Indeed, the necessary conditions are identical except that $\chi(\underline{\theta})$ is not required to be zero.

Following the proof of Proposition 1 it follows immediately that $\tau(\theta) \geq 0$ for all $\theta \geq \theta^*$. Finally we can use Proposition 9 to conclude that we cannot have $\tau(\theta) = 0$ almost surely for $\theta \geq \theta^*$.

A.6 Proof of Proposition 5

The proposition is a direct application of Proposition 6 specialized to the logarithmic utility case.

A.7 Proof of Proposition 6

We define the after-tax interest rate $\tilde{R} = \frac{R}{1+\tau}$. The planning problem is[1]

$$\min \int \left(c_0(I, \tilde{R}, \theta) + \frac{1}{R} c_1(I, \tilde{R}, \theta) \right) f(\theta) d\theta$$

[1] If $\alpha_\theta = 0$ this case amounts to a many-person Ramsey tax problem, as in Diamond (1975).

subject to

$$\int \left(\lambda_\theta V^p(I,\tilde{R},\theta) + \alpha_\theta U^c(c_1(I,\tilde{R},\theta))\right) f(\theta)d\theta \geq V,$$

where $c_0(I,\tilde{R},\theta)$ and $c_1(I,\tilde{R},\theta)$ are the uncompensated demand functions and $V^p(I,\tilde{R},\theta)$ is the indirect utility function. We denote by $W(I,\tilde{R},\theta) = \lambda_\theta V^p(I,\tilde{R},\theta) + \alpha_\theta U^c(c_1(I,\tilde{R},\theta)))$. Finally we denote by $c_0^c(u,\tilde{R};\theta)$ and $c_1^c(u,\tilde{R};\theta)$ the compensated demand functions, and by $\varepsilon_{c_1,\tilde{R}}(I,\tilde{R},\theta)$ the compensated elasticity of c_1 to the after tax interest rate $\tilde{R}$.

We now proceed to prove Proposition 6. We use

$$c_{0,I}(\theta) + \frac{1}{\tilde{R}} c_{1,I}(\theta) = 1,$$

$$c_{0,\tilde{R}}(\theta) + \frac{1}{\tilde{R}} c_{1,\tilde{R}}(\theta) - \frac{1}{\tilde{R}^2} c_1(\theta) = 0,$$

$$c_{1,\tilde{R}}(\theta) = c^c_{1,\tilde{R}}(\theta) + \frac{1}{\tilde{R}^2} c_1(\theta) c_{1,I}(\theta),$$

$$c_{0,\tilde{R}}(\theta) = c^c_{0,\tilde{R}}(\theta) + \frac{1}{\tilde{R}^2} c_1(\theta) c_{0,I}(\theta),$$

$$V^p_{\tilde{R}}(\theta) = \frac{1}{\tilde{R}^2} V^p_I(\theta) c_1(\theta).$$

We find

$$\int \left[\left(\frac{1}{R} - \frac{1}{\tilde{R}}\right) c_{1,\tilde{R}}(\theta) + \frac{1}{\tilde{R}^2} c_1(\theta) \left(1 - \nu V^p_I(\theta)\right) - \alpha_\theta \nu U^c_{c_1}(\theta) c_{1,\tilde{R}}(\theta) \right] f(\theta)d\theta = 0,$$

$$\left(\frac{1}{\tilde{R}} - \frac{1}{R}\right) = \frac{\int \left[\frac{1}{\tilde{R}^2} c_1(\theta) \left(1 - \nu V^p_I(\theta)\right) - \alpha_\theta \nu U^c_{c_1}(\theta) c_{1,\tilde{R}}(\theta)\right] f(\theta)d\theta}{\int c_{1,\tilde{R}}(\theta) f(\theta)d\theta}.$$

Similarly, we have

$$\int \left[1 + \left(\frac{1}{R} - \frac{1}{\tilde{R}}\right) c_{1,I}(\theta) - \nu V^p_I(\theta) - \alpha_\theta \nu U^c_{c_1}(\theta) c_{1,I}(\theta)\right] f(\theta)d\theta = 0,$$

$$\left(\frac{1}{\tilde{R}} - \frac{1}{R}\right) = \frac{\int \left[1 - \nu V^p_I(\theta) - \alpha_\theta \nu U^c_{c_1}(\theta) c_{1,I}(\theta)\right] f(\theta)d\theta}{\int c_{1,I}(\theta) f(\theta)d\theta}.$$

After some manipulations, we get

$$\frac{1}{R} - \frac{1}{\tilde{R}} = \frac{\frac{\nu}{\tilde{R}^2} \mathrm{Cov}\left(c_1(\theta), W_I(\theta)\right) + \int \alpha_\theta \nu U^c_{c_1}(\theta) c^c_{1,\tilde{R}}(\theta) f(\theta)d\theta}{\int c^c_{1,\tilde{R}}(\theta) f(\theta)d\theta + \frac{1}{\tilde{R}^2} \mathrm{Cov}\left(c_1(\theta), c_{1,I}(\theta)\right)},$$

which can be transformed into

$$\frac{\tau}{1+\tau} = -\nu\tilde{R}\frac{\frac{1}{\tilde{R}^2}\text{Cov}\left(c_1(\theta), W_I(\theta)\right) + \int \alpha_\theta U^c_{c_1}(\theta) c^c_{1,\tilde{R}}(\theta) f(\theta) d\theta}{\int c^c_{1,\tilde{R}}(\theta) f(\theta) d\theta + \frac{1}{\tilde{R}^2}\text{Cov}\left(c_1(\theta), c_{1,I}(\theta)\right)},$$

and finally

$$\frac{\tau}{1+\tau} = -\nu \frac{\frac{\text{Cov}(c_1(\theta), W_I(\theta))}{\int \varepsilon_{c_1,\tilde{R}}(\theta) c_1(\theta) f(\theta) d\theta} + \tilde{R}\frac{\int \alpha_\theta U^c_{c_1}(\theta)\varepsilon_{c_1,\tilde{R}}(\theta) c_1(\theta) f(\theta) d\theta}{\int \varepsilon_{c_1,\tilde{R}}(\theta) c_1(\theta) f(\theta) d\theta}}{1 + \frac{1}{\tilde{R}}\frac{\text{Cov}\left(c_1(\theta), c_{1,I}(\theta)\right)}{\int \varepsilon_{c_1,\tilde{R}}(\theta) c_1(\theta) f(\theta) d\theta}}.$$

A.8 Proof of Proposition 7

We apply Proposition 9 to the logarithimic utility case to prove Proposition 7. We have

$$c_1(I, \tilde{R}, \theta) = \min(I\tilde{R}\theta, \exp(\underline{u})),$$

$$\lambda_\theta V^p(I, \tilde{R}, \underline{u}, \theta) = \begin{cases} \lambda_\theta \left[(1-\theta)\log\left(I - \frac{\exp(\underline{u})}{\tilde{R}}\right) + \theta \underline{u}\right] & \text{if} \quad I\tilde{R}\theta \leq \exp(\underline{u}), \\ \lambda_\theta \left[\log I + (1-\theta)\log(1-\theta) + \theta \log(\tilde{R}\theta)\right] & \text{if} \quad I\tilde{R}\theta \geq \exp(\underline{u}), \end{cases}$$

so that

$$\lambda_\theta V^p_I(I, \tilde{R}, \underline{u}, \theta) = \begin{cases} \frac{\lambda_\theta(1-\theta)}{I - \frac{\exp(\underline{u})}{\tilde{R}}} & \text{if} \quad I\tilde{R}\theta \leq \exp(\underline{u}), \\ \frac{\lambda_\theta}{I} & \text{if} \quad I\tilde{R}\theta \geq \exp(\underline{u}). \end{cases}$$

Note that for $I\tilde{R}\theta \leq \exp(\underline{u})$, we have

$$\frac{\lambda_\theta(1-\theta)}{I - \frac{\exp(\underline{u})}{\tilde{R}}} = \frac{\lambda_\theta(1-\theta)\tilde{R}}{I\tilde{R}(1-\theta) + I\tilde{R}\theta - \exp(\underline{u})} = \frac{\lambda_\theta}{I + \frac{I\tilde{R}\theta - \exp(\underline{u})}{\tilde{R}(1-\theta)}} > \frac{\lambda_\theta}{I},$$

and $\frac{I\tilde{R}\theta - \exp(\underline{u})}{\tilde{R}(1-\theta)}$ is negative and increasing in θ. Hence $\text{Cov}(c_1(\theta), V^p_I(\theta))$ is negative even when λ_θ is constant. Also, $c_{1,I}(\theta) = \tilde{R}\theta$ for $\theta \geq \frac{\exp(\underline{u})}{I\tilde{R}}$ and 0 otherwise. Thus, both $c_1(\theta)$ and $c_{1,I}(\theta)$ are increasing in θ, implying that $\text{Cov}\left(c_1(\theta), c_{1,I}(\theta)\right) \geq 0$.

A.9 Proof of Proposition 8

We first show that $V^p_I(I, R, \underline{u}; \theta)$ is higher for types such that the constraint $U^c(c_1(I, R, \underline{u}; \theta)) = \underline{u}$ binds. Let C^c_1 be the inverse function of U^c. We have

$$V^p(I, R, \underline{u}; \theta) = \max U(I - \frac{1}{R}c_1, c_1)$$

subject to

$$c_1 \geq C_1^c(\underline{u}).$$

The FOC implies that

$$\frac{U_{c_1}}{U_{c_0}} \leq \frac{1}{R},$$

with an equality if $c_1(I, R, \underline{u}; \theta) > C_1^c(\underline{u})$, i.e. if $\theta > \theta^*$. Consider $\theta < \theta^*$ so that $c_1(I, R, \underline{u}; \theta) = C_1^c(\underline{u})$. We have

$$V_I^p(I, R, \underline{u}; \theta) = U_{c_0}^p(I - \frac{1}{R}C_1^c(\underline{u}), C_1^c(\underline{u})).$$

Hence

$$U_{c_1}^c \frac{d}{d\underline{u}} V_I^p = -\frac{1}{R} U_{c_0,c_0}^p + U_{c_0,c_1}^p,$$

where we have omitted the arguments for brevity.

Since $\frac{U_{c_1}^p}{U_{c_0}^p} \leq \frac{1}{R}$ and $U_{c_0,c_0}^p \leq 0$, we get

$$U_{c_1}^c \frac{d}{d\underline{u}} V_I^p \geq -\frac{U_{c_1}^p}{U_{c_0}^p} U_{c_0,c_0}^p + U_{c_0,c_1}^p \geq 0,$$

where the second inequality follows from the assumption that c_1 is a normal good. Hence $\frac{d}{d\underline{u}} V_I^p \geq 0$. Using the fact then when $\underline{u}$ is low enough, the constraint $U^c(c_1(I, R, \underline{u}; \theta)) = \underline{u}$ ceases to bind and then $V_I^p(I, R, \underline{u}; \theta)$ is independent of θ, we conclude that $V_I^p(I, R, \underline{u}; \theta)$ is higher for types such that the constraint $U^c(c_1(I, R, \underline{u}; \theta)) = \underline{u}$ binds.

We now use this observation to prove that a small positive tax is beneficial. We find it easier to work in the prime where we maximize welfare subject to a resource constraint rather than in the dual. Imagine that we change R to $R + dR$. To satisfy the resource constraint, we need to change I to $I + dI$, where $dI = -\frac{dR}{R^2} \int c_1(I, R, \underline{u}; \theta) f(\theta) d\theta$. The change in welfare is given by

$$dW = \int \lambda_\theta (V_R^p(\theta) dR + V_I^p(\theta) dI) f(\theta) d\theta,$$

which using $V_R^p = \frac{c_1}{R^2} V_I^p$, we can rewrite as

$$dW = \int \lambda_\theta (c_1(\theta) V_I^p(\theta) \frac{dR}{R^2} - V_I(\theta) \frac{dR}{R^2} \int c_1(\theta') f(\theta') d\theta') f(\theta) d\theta,$$

or

$$dW = \frac{dR}{R^2} \int \lambda_\theta (c_1(\theta) V_I^p(\theta) - V_I^p(\theta) \int c_1(\theta') f(\theta') d\theta') f(\theta) d\theta,$$

which when $\lambda_\theta = \bar{\lambda}$ is constant, can be rewritten as

$$dW = \frac{dR}{R^2}\bar{\lambda}\text{Cov}(c_1(\theta), V_I^p(\theta)).$$

Our previous result shows that $\text{Cov}(c_1(\theta), V_I^p(\theta)) < 0$, assuming normality of good c_1. We conclude that starting with no tax, a small positive tax increases welfare.

A.10 Proof of Proposition 9

The planning problem is

$$\min \int \left(c_0(I, \tilde{R}, \underline{u}, \theta) + \frac{1}{R} c_1(I, \tilde{R}, \underline{u}, \theta) \right) f(\theta) d\theta$$

subject to

$$\int \lambda_\theta V^p(I, \tilde{R}, \underline{u}, \theta) f(\theta) d\theta \geq V,$$

where we have defined $\underline{u} = U^c(R\underline{B})$ and $V^p(I, \tilde{R}, \underline{u}, \theta) = \max_{c_0, c_1} U^p(c_0, c_1; \theta)$ subject to $c_0 + \frac{c_1}{\tilde{R}} \leq I$ and $U^c(c_1(\theta)) \geq \underline{u}$, with demands $c_0(I, \tilde{R}, \underline{u}, \theta)$ and $c_1(I, \tilde{R}, \underline{u}, \theta)$. The dual problem $\min_{c_0, c_1} \left(c_0 + \frac{c_1}{\tilde{R}} \right)$ subject to $U = U^p(c_0, c_1; \theta)$ and $U^c(c_1(\theta)) \geq \underline{u}$ gives associated compensated demands $c_0^c(U, \tilde{R}, \underline{u}, \theta)$ and $c_1^c(U, \tilde{R}, \underline{u}, \theta)$, and by $\varepsilon_{c_1, \tilde{R}}(I, \tilde{R}, \theta)$ the compensated elasticity of c_1 to the after tax interest rate $\tilde{R}$.

The proof of Proposition 9 follows exactly the same steps as that of Proposition 6. Indeed, this follows from the fact that the Slutsky relations between uncompensated and compensated demands are still valid $c_{1,\tilde{R}}^c = -\frac{c_1}{\tilde{R}^2} c_{1,I} + c_{1,\tilde{R}}$.[2]

References

Diamond, P. A., “A many-person Ramsey tax rule,” *Journal of Public Economics*, November 1975, *4* (4), 335–342.

Hellwig, Martin, “A Maximum Principle for Control Problems with Monotonicity Constraints,” Working Paper, Max Planck Insittute for Research on Collective Goods August 2009.

[2]Indeed using a variation of dI and $d\tilde{R}$ such that $V_I \frac{dI}{d\tilde{R}} + V_{\tilde{R}} = 0$, we get $c_{1,\tilde{R}}^c = -\frac{V_{\tilde{R}}}{V_I} c_{1,I} + c_{1,\tilde{R}}$, with $\frac{V_{\tilde{R}}}{V_I} = \frac{c_1}{\tilde{R}^2}$.