Online Appendix
High Marginal Tax Rates on the Top 1%? Lessons from a Life Cycle Model with Idiosyncratic Income Risk

By Fabian Kindermann and Dirk Krueger

A. Proofs of Propositions

A1. Proof of Proposition 1

We start from the definition of top 1% labor earnings tax revenue

$$T(\tau_h) = \tau_h(z_h - \bar{z}) - R, \tag{A1}$$

which we can, for the purpose of notation, also write as

$$T(\tau_h) = T(\bar{z}) + \tau_h(z_h - \bar{z}), \tag{A2}$$

with $T(\bar{z}) = -R$. Total differentiation yields

$$dT(\tau_h) = dT(\bar{z}) + d\tau_h(z_h - \bar{z}) + \tau_h dz_h. \tag{A3}$$

By some rearranging, we obtain

$$dT(\tau_h) = \frac{d\left[\frac{T(\bar{z})}{\bar{z}}\right]}{d(1-\tau_h)} \cdot \frac{(1-\tau_h)}{T(\bar{z})/\bar{z}} \cdot \frac{T(\bar{z})}{\bar{z}} \cdot \frac{\bar{z}}{1-\tau_h} \cdot d(1-\tau_h) \tag{A4}$$

$$- (z_h - \bar{z})d(1-\tau_h) + \frac{dz_h}{d(1-\tau_h)} \cdot \frac{1-\tau_h}{z_h} \cdot \frac{\tau_h z_h}{1-\tau_h} \cdot d(1-\tau_h). \tag{A5}$$

With the definitions as in Proposition 1, we immediately get

$$\frac{dT(\tau_h)}{d(1-\tau_h)} = \epsilon(\tau_a(\bar{z})) \cdot \frac{\tau_a(\bar{z})\bar{z}}{1-\tau_h} - (z_h - \bar{z}) + \epsilon(z_h) z_h \cdot \frac{\tau_h}{1-\tau_h}. \tag{A6}$$

The peak of the Laffer curve can then be found by setting $\frac{dT(\tau_h)}{d(1-\tau_h)} \overset{!}{=} 0$, which yields

$$\epsilon(\tau_a(\bar{z})) \cdot \frac{\tau_a(\bar{z})}{1-\tau_h} - \left(\frac{z_h}{\bar{z}} - 1\right) + \epsilon(z_h) \cdot \frac{z_h}{\bar{z}} \cdot \frac{\tau_h}{1-\tau_h} = 0. \tag{A7}$$

Using $\frac{z_h}{\bar{z}} = \frac{a}{a-1}$ and solving for τ_h gives us

$$(a-1)\cdot\epsilon(\tau_a(\bar{z}))\cdot\tau_a(\bar{z}) - (1-\tau_h) + \epsilon(z_h)\cdot a\cdot\tau_h = 0 \tag{A8}$$

from which we immediately get

$$\tau_{\text{Laffer}} := \tau_h = \frac{1-(a-1)\cdot\epsilon(\tau_a(\bar{z}))\cdot\tau_a(\bar{z})}{1+a\cdot\epsilon(z_h)}. \tag{A9}$$

$\square$

A2. Proof of Proposition 3

The optimization problem of a household with high labor productivity e_h reads

$$\max_{c_h,n_h} \frac{c_h^{1-\gamma}}{1-\gamma} - \lambda\frac{n_h^{1+\frac{1}{\chi}}}{1+\frac{1}{\chi}} \quad \text{s.t.} \quad c_h = e_h n_h - \tau_h(e_h n_h - \bar{z}) + R. \tag{A10}$$

The first-order conditions of this problem are

$$c_h^{-\gamma} = \mu \quad \text{and} \quad \lambda n_h^{\frac{1}{\chi}} = \mu(1-\tau_h)e_h, \tag{A11}$$

where μ is the Lagrange multiplier on the budget constraint. Combining these equations with the budget constraint yields the labor supply equation

$$n_h e_h - \tau_h(n_h e_h - \bar{z}) + R - \left[\frac{(1-\tau_h)e_h}{\lambda}\right]^{\frac{1}{\gamma}} n_h^{-\frac{1}{\gamma\chi}} = 0. \tag{A12}$$

Uncompensated labor supply elasticity. — Total differentiation with respect to e_h yields

$$\left\{(1-\tau_h)z_h + \frac{1}{\gamma\chi}\left[z_h - \tau_h(z_h-\bar{z}) + R\right]\right\}\frac{dn_h}{n_h} \tag{A13}$$

$$+\left\{(1-\tau_h)z_h - \frac{1}{\gamma}\left[z_h - \tau_h(z_h-\bar{z}) + R\right]\right\}\frac{de_h}{e_h} = 0. \tag{A14}$$

Rearranging leads to

$$\epsilon_h^u = \frac{dn_h}{de_h} \cdot \frac{e_h}{n_h} = \frac{-(1-\tau_h)z_h + \frac{1}{\gamma}\left[z_h - \tau_h(z_h - \bar{z}) + R\right]}{(1-\tau_h)z_h + \frac{1}{\gamma\chi}\left[z_h - \tau_h(z_h - \bar{z}) + R\right]} \tag{A15}$$

$$= \frac{(1-\gamma)(1-\tau_h)z_h + \tau_h\bar{z} + R}{\left(\gamma + \frac{1}{\chi}\right)(1-\tau_h)z_h + \frac{\tau_h\bar{z}+R}{\chi}}. \tag{A16}$$

Income elasticity of labor supply. — Total differentiation of the labor supply equation with respect to R yields

$$\left\{(1-\tau_h)z_h + \frac{1}{\gamma\chi}\left[z_h - \tau_h(z_h - \bar{z}) + R\right]\right\}\frac{dn_h}{n_h} + dR = 0 \tag{A17}$$

which immediately gives

$$\eta_h = \frac{dz_h}{dR}(1-\tau_h) = \frac{e_h \cdot dn_h}{dR}(1-\tau_h) \tag{A18}$$

$$= \frac{-(1-\tau_h)z_h}{(1-\tau_h)z_h + \frac{1}{\gamma\chi}\left[z_h - \tau_h(z_h - \bar{z}) + R\right]} \tag{A19}$$

$$= \frac{-\gamma(1-\tau_h)z_h}{\left(\gamma + \frac{1}{\chi}\right)(1-\tau_h)z_h + \frac{\tau_h\bar{z}+R}{\chi}}. \tag{A20}$$

Policy elasticity. — Before taking the total differential with respect to our tax reform, it is useful to formulate the labor supply equation in labor earnings terms as

$$z_h - T(\bar{z}) - \tau_h(z_h - \bar{z}) - \left[\frac{(1-\tau_h)}{\lambda}\right]^{\frac{1}{\gamma}} \cdot e_h^{\frac{1}{\gamma}\left[1+\frac{1}{\chi}\right]} z_h^{-\frac{1}{\gamma\chi}} = 0. \tag{A21}$$

Note that we again use the notation $T(\bar{z}) = -R$ from Appendix A.A1. Total differentiation with respect to the policy experiment then yields

$$dz_h - dT(\bar{z}) - d\tau_h(z_h - \bar{z}) - \tau_h dz_h \tag{A22}$$

$$- \frac{1}{\gamma} \cdot \frac{d(1-\tau_h)}{1-\tau_h} \cdot \left[\frac{(1-\tau_h)}{\lambda}\right]^{\frac{1}{\gamma}} \cdot e_h^{\frac{1}{\gamma}\left[1+\frac{1}{\chi}\right]} z_h^{-\frac{1}{\gamma\chi}} \tag{A23}$$

$$+ \frac{1}{\gamma\chi} \cdot \frac{dz_h}{z_h} \cdot \left[\frac{(1-\tau_h)}{\lambda}\right]^{\frac{1}{\gamma}} \cdot e_h^{\frac{1}{\gamma}\left[1+\frac{1}{\chi}\right]} z_h^{-\frac{1}{\gamma\chi}} = 0. \tag{A24}$$

Some rearranging gives us

$$\text{(A25)} \quad \left\{\gamma(1-\tau_h)z_h + \frac{1}{\chi}\left[(1-\tau_h)z_h + (\tau_h - \tau_a(\bar{z}))\bar{z}\right]\right\}\frac{dz_h}{z_h}$$

$$\text{(A26)} \quad + \left\{\gamma(1-\tau_h)(z_h - \bar{z}) - \left[(1-\tau_h)z_h + (\tau_h - \tau_a(\bar{z}))\bar{z}\right]\right\} \cdot \frac{d(1-\tau_h)}{1-\tau_h}$$

$$\text{(A27)} \quad - \gamma\frac{d\tau_a(\bar{z})}{d(1-\tau_h)} \cdot \frac{1-\tau_h}{\tau_a(\bar{z})} \cdot \tau_a(\bar{z}) \cdot \bar{z} \cdot \frac{d(1-\tau_h)}{1-\tau_h} = 0.$$

Hence, we obtain with $-\tau_a(\bar{z})\bar{z} = R$

(A28)

$$\epsilon(z_h) = \frac{dz_h}{d(1-\tau_h)} \cdot \frac{1-\tau_h}{z_h}$$

$$\text{(A29)} \quad = \frac{(1-\gamma)(1-\tau_h)z_h + (\tau_h - \tau_a(\bar{z}))\bar{z}}{\left(\gamma + \frac{1}{\chi}\right)(1-\tau_h)z_h + \frac{(\tau_h - \tau_a(\bar{z}))\bar{z}}{\chi}} + \frac{\gamma(1-\tau_h)\bar{z}_h + \gamma\tau_a(\bar{z})\bar{z}\epsilon(\tau_a(\bar{z}))}{\left(\gamma + \frac{1}{\chi}\right)(1-\tau_h)z_h + \frac{(\tau_h - \tau_a(\bar{z}))\bar{z}}{\chi}}$$

$$\text{(A30)} \quad = \frac{(1-\gamma)(1-\tau_h)z_h + \tau_h\bar{z} + R}{\left(\gamma + \frac{1}{\chi}\right)(1-\tau_h)z_h + \frac{\tau_h\bar{z}+R}{\chi}} + \frac{\gamma(1-\tau_h)\bar{z}_h + \gamma\tau_a(\bar{z})\bar{z}\epsilon(\tau_a(\bar{z}))}{\left(\gamma + \frac{1}{\chi}\right)(1-\tau_h)z_h + \frac{\tau_h\bar{z}+R}{\chi}}$$

$$\text{(A31)} \quad = \epsilon_h^u - \frac{-\gamma(1-\tau_h)z_h}{\left(\gamma + \frac{1}{\chi}\right)(1-\tau_h)z_h + \frac{\tau_h\bar{z}+R}{\chi}} \cdot \frac{\bar{z}}{z_h} \cdot \left[1 + \frac{\tau_a(\bar{z})}{1-\tau_h} \cdot \epsilon(\tau_a(\bar{z}))\right]$$

$$\text{(A32)} \quad = \epsilon_h^u - \eta_h \cdot \frac{\bar{z}}{z} \cdot \left[1 + \frac{\tau_a(\bar{z})}{1-\tau_h} \cdot \epsilon(\tau_a(\bar{z}))\right].$$

□

Comparison with the Saez (2001) result. — The formula for the Laffer tax rate can hence be written as

$$\text{(A33)} \quad \tau_{\text{Laffer}} = \frac{1 - (a-1) \cdot \tau_a(\bar{z}) \cdot \epsilon(\tau_a(\bar{z}))}{1 + a \cdot \epsilon_h^u - \eta_h \cdot (a-1) \cdot \left[1 + \frac{\tau_a(\bar{z})}{1-\tau_h} \cdot \epsilon(\tau_a(\bar{z}))\right]}.$$

With $\epsilon(\tau_a(\bar{z})) = 0$ as in Saez (2001), the formula reduces to

(A34)

$$\tau_{\text{Laffer}} = \frac{1}{1 + a\epsilon_h^u - \eta_h \cdot (a-1)} = \frac{1}{1 + \epsilon_h^u + \left[\epsilon_h^u - \eta_h\right] \cdot (a-1)} = \frac{1}{1 + \epsilon_h^u + \epsilon_h^c \cdot (a-1)},$$

with $\epsilon_h^u = \epsilon_h^c + \eta$ and ϵ_h^c being the compensated labor supply elasticity. This is the same as in equation (9) in Saez (2001, p. 212) with $\bar{g} = 0$.

A3. Further Discussion of the Policy Elasticity

In this appendix we discuss the implications of Proposition 3 in greater detail. First, this proposition clarifies that the specific tax system and reform matters for the size of the income effect and thus the policy elasticity. In contrast to the purely proportional tax system studied in Corollary 4, if $\bar{z} > 0$, then the income effect from the tax change is smaller since taxes are only lowered above the threshold $\bar{z}$. In fact, household exactly at the threshold $z_h = \bar{z}$ experience no income effect at all (recall we still assume $\tau_a(\bar{z}) = 0$), and thus the policy elasticity $\epsilon(z_h) = \epsilon_h^u - \eta_h \cdot \frac{\bar{z}}{z_h} = \epsilon_h^u - \eta_h = \epsilon_h^c$ is governed exclusively by the Hicksian compensated labor supply elasticity. A household with a greater z_h experiences a larger negative income effect on leisure (a larger positive income effect on labor). Consequently, the extra $\eta_h \cdot \frac{1-a}{a} = \eta_h \cdot \frac{\bar{z}}{z_h}$ declines with z_h, and the labor earnings reaction becomes smaller. If z_h is large relative to $\bar{z}$, the labor earnings reaction is approximately the same as in a proportional tax system, as the tax payment $T(\bar{z})$ on income below the threshold is small relative to the total tax bill $T(z)$.

Second, if in addition other parts of the tax schedule adjust to the change in the top rate, then $\epsilon(\tau_a(\bar{z})) \neq 0$. Even if the top earner was exactly at the top threshold $\bar{z}$, she would experience an additional income effect on labor supply due to the tax change for her income below the threshold. In the case of $\epsilon(\tau_a(\bar{z})) < 0$, these additional income effects make labor supply more elastic to the tax reform (i.e. increase the policy elasticity $\epsilon(z_h)$), and thus, ceteris paribus, reduce the maximal tax revenue and tax rate at which the peak of the top Laffer curve is attained.

A4. Theoretical Welfare Results in Section 2.4

We proceed in four steps to prove the results in Section 2.4. First, in the next subsection we characterize the peak of the Laffer curve for the specification of the simple model used in Section 2.4, since it is needed for the proof of Proposition 5. Then, in Section A.A4 we derive the condition in Proposition 5 that insures that high-productivity individuals are at least as well-off as low-productivity individuals even at that Laffer tax rate and that low-productivity individuals have labor income $z_l < \bar{z}$. Then we prove Proposition 5 (in Section A.A4) and Proposition ref(in Section A.A4).

Revenue Maximization: The Laffer Curve Revisited. — The revenue maximization problem, given the optimal labor supply choice of top income earners, can be stated as

$$\max_{\tau_h} \tau_h \left[[(1-\tau_h)]^{\chi} [e_h]^{1+\chi} - \bar{z} \right] \tag{A35}$$

with first order condition

$$\text{(A36)} \qquad \left[(1-\tau_h)\right]^{\chi}\left[e_h\right]^{1+\chi} - \bar{z} = \chi\tau_h\left[1-\tau_h\right]^{\chi-1}\left[e_h\right]^{1+\chi}.$$

Thus[45]

$$\text{(A37)} \qquad \frac{1-\tau_h}{\tau_h} = \chi + \frac{\bar{z}}{\tau_h\left[1-\tau_h\right]^{\chi-1}\left[e_h\right]^{1+\chi}} > \chi$$

and thus the revenue maximizing tax rate satisfies

$$\text{(A38)} \qquad \tau_h^{\text{Laffer}} < \frac{1}{1+\chi} = \tau_h^{\text{Laffer}}(\bar{z}. = 0)$$

To state the revenue-maximizing rate more concisely, recall that the Pareto coefficient is defined as

$$\text{(A39)} \qquad \frac{a}{a-1} = \frac{e_h n_h}{\bar{z}} = \frac{\left[1-\tau_h\right]^{\chi}\left[e_h\right]^{1+\chi}}{\bar{z}} \quad \text{or} \quad \frac{a-1}{a} = \frac{\bar{z}}{\left[1-\tau_h\right]^{\chi}\left[e_h\right]^{1+\chi}}.$$

Note that if $\bar{z}=0$, then $a=1$ and as $\bar{z}\to z_h$ then $a\to\infty$. Then the revenue-maximizing tax rate satisfies

$$\text{(A40)} \qquad \frac{1-\tau_h}{\tau_h} = \chi + \frac{a-1}{a}\frac{1-\tau_h}{\tau_h} \quad \text{or} \quad \tau_h = \frac{1}{1+a(\tau_h)\chi}.$$

precisely as predicted by the Saez formula. But it is important to note that

$$\text{(A41)} \qquad \frac{a-1}{a} = \frac{\bar{z}}{\left[1-\tau_h\right]^{\chi}\left[e_h\right]^{1+\chi}} = \frac{a-1}{a}(\tau_h) \quad \text{or}$$

$$\text{(A42)} \qquad a(\tau_h;\bar{z}) = \frac{\left[1-\tau_h\right]^{\chi}\left[e_h\right]^{1+\chi}}{\left[1-\tau_h\right]^{\chi}\left[e_h\right]^{1+\chi}-\bar{z}} = \frac{1}{1-\frac{\bar{z}}{\left[1-\tau_h\right]^{\chi}\left[e_h\right]^{1+\chi}}}.$$

We observe that $a(\tau_h;\bar{z})$ is a strictly increasing function of the tax rate τ_h and a strictly increasing function of the threshold $\bar{z}$. Thus the right hand side of $(A40)$ is continuous and strictly decreasing in τ_h, strictly positive at $\tau_h=0$ and tends to 0 as τ_h tends to 1. Thus there is a unique positive revenue-maximizing tax rate τ_h^{Laffer} characterized by $(A40)$, and since the right hand side is strictly decreasing in $\bar{z}$, so is τ_h^{Laffer}. At $\bar{z}=0$, we find $\tau_h^{\text{Laffer}}(\bar{z}=0)=\frac{1}{1+\chi}$.

[45] The previous equation also insures that at the revenue-maximizing tax rate (and thus at any tax rate lower than that) labor income of the top income earners $n_h e_h$ is strictly higher than the threshold $\bar{z}$.

Insuring that High Productivity Households Are Better Off. — In principle, a high-productivity worker could work at lower productivity e_l and not pay taxes. We now state a condition insuring that this is not in their interest. Furthermore, we maintained the assumption that low-productivity workers, at their optimal labor supply, have income less than the tax threshold $\bar{z}$. We now provide a sufficient condition on parameters insuring that both assumptions implicit in our analysis are satisfied

For a given tax rate τ_h a high-productivity worker is better off than working as a low-productivity worker if

$$\frac{[e_l]^{1+\chi}}{1+\chi} + R \leq \frac{[(1-\tau_h)e_h]^{1+\chi}}{1+\chi} + \tau_h \bar{z} + R \tag{A43}$$

and thus

$$\frac{[e_l]^{1+\chi}}{1+\chi} \leq \frac{[(1-\tau_h)e_h]^{1+\chi}}{1+\chi} + \tau_h \bar{z}. \tag{A44}$$

Since the welfare-maximizing tax rate cannot exceed the revenue- and thus transfer-maximizing tax rate, a sufficient (but by no means necessary) condition for the welfare analysis that this condition is satisfied at the peak of the Laffer curve rate $\tau_h^{\text{Laffer}}(\bar{z}=0) = \frac{1}{1+\chi}$ since for all $\bar{z} \geq 0$ and all $\tau_h \leq \tau_h^{\text{Laffer}}(\bar{z}=0)$

$$\frac{[(1-\tau_h)e_h]^{1+\chi}}{1+\chi} + \tau_h \bar{z} \geq \frac{\left[\frac{\chi}{1+\chi} e_h\right]^{1+\chi}}{1+\chi}. \tag{A45}$$

Thus a sufficient condition such that (for all $\bar{z}$) high-income individuals have higher welfare from post-tax consumption and labor than low-income individuals is

$$\frac{[e_l]^{1+\chi}}{1+\chi} \leq \frac{\left[\frac{\chi}{1+\chi} e_h\right]^{1+\chi}}{1+\chi} \tag{A46}$$

or

$$e_l \leq \frac{\chi}{1+\chi} e_h. \tag{A47}$$

Note that this is a (potentially very loose) sufficient condition, and a much tighter (but $\bar{z}$-specific) condition could be obtained. Furthermore, for low-productivity workers to have earnings below the threshold at their optimal labor supply requires that

$$e_l n_l = e_l^{1+\chi} =\leq \bar{z} \tag{A48}$$

and thus

$$e_l \leq \bar{z}^{\frac{1}{1+\chi}}. \tag{A49}$$

Thus a sufficient condition insuring that both low-productivity workers do not have taxable income, and that their realized utility does not dominate that of high-productivity workers is, combining equations (A47) and (A49)

$$e_l \leq \min \left\{ \bar{z}^{\frac{1}{1+\chi}}, \frac{\chi}{1+\chi} e_h \right\} \tag{A50}$$

This is the sufficient condition imposed in Proposition

Utilitarian Optimum: Proof of Proposition 5. — Now turn to Utilitarian welfare, defined in the main text as

$$\mathcal{W}(\tau_h) = \Phi_l \frac{\left(\frac{[e_l]^{1+\chi}}{1+\chi} + R\right)^{1-\gamma}}{1-\gamma} + (1-\Phi_l) \frac{\left(\frac{[(1-\tau_h)e_h]^{1+\chi}}{1+\chi} + \tau_h \bar{z} + R\right)^{1-\gamma}}{1-\gamma} \tag{A51}$$

$$R(\tau_h) = (1-\Phi_l)\tau_h \left[[1-\tau_h]^\chi [e_h]^{1+\chi} - \bar{z} \right]. \tag{A52}$$

Trivially $\mathcal{W}(\tau_h)$ is independent of Ψ (item 1 of the proposition).

Taking first order conditions with respect to the tax rate τ_h and rearranging yields

$$\Theta(\tau_h) := \left(\frac{[(1-\tau_h)e_h]^{1+\chi} + (1+\chi)\tau_h \bar{z} + (1+\chi)R(\tau_h)}{[e_l]^{1+\chi} + (1+\chi)R(\tau_h)} \right)^\gamma \tag{A53}$$

$$= \frac{1-\Phi_l}{\Phi_l} \left(\frac{[e_h]^{1+\chi} [1-\tau_h]^\chi - \bar{z}}{\frac{dR(\tau_h)}{d\tau_h}} - 1 \right) := \Gamma(\tau_h) \tag{A54}$$

with

$$\frac{dR(\tau_h)}{d\tau_h} = (1-\Phi_l) \left([1-\tau_h]^\chi [e_h]^{1+\chi} - \bar{z} - \tau_h \chi [1-\tau_h]^{\chi-1} [e_h]^{1+\chi} \right) \tag{A55}$$

$$= (1-\Phi_l) [e_h]^{1+\chi} [1-\tau_h]^\chi \left(\left[1 - \frac{\bar{z}}{[1-\tau_h]^\chi [e_h]^{1+\chi}} \right] - \frac{\chi \tau_h}{1-\tau_h} \right) \tag{A56}$$

and thus

$$\Theta(\tau_h) := \left(\frac{[(1-\tau_h)e_h]^{1+\chi} + (1+\chi)\tau_h \bar{z} + (1+\chi)R(\tau_h)}{[e_l]^{1+\chi} + (1+\chi)R(\tau_h)} \right)^{\gamma} \tag{A57}$$

$$= \frac{1-\Phi_l}{\Phi_l} \left(\frac{1}{(1-\Phi_l)\left(1 - \frac{a(\tau_h)\chi\tau_h}{1-\tau_h}\right)} - 1 \right) := \Gamma(\tau_h). \tag{A58}$$

The existence, uniqueness and comparative statics in Proposition 5 then follow the properties of the functions $(\Gamma(\tau_h), \Theta(\tau_h))$.

Since $a(\tau_h)$ is strictly increasing in τ_h, the function $\Gamma(\tau_h)$ is continuous, strictly increasing on $[0, \tau_h^{\text{Laffer}})$ and with

$$\Gamma(\tau_h = 0) = \frac{1-\Phi_l}{\Phi_l} \left(\frac{1}{1-\Phi_l} - 1 \right) = 1 \tag{A59}$$

$$\lim_{\tau_h \to \tau_h^{\text{Laffer}}} \Gamma(\tau_h) = \frac{1-\Phi_l}{\Phi_l} \left(\frac{1}{(1-\Phi_l)\left(1 - \frac{\frac{a\chi}{1+a\chi}}{\frac{a\chi}{1+a\chi}}\right)} - 1 \right) = \infty. \tag{A60}$$

Finally, $\Gamma(\tau_h)$ is independent of γ and e_h/e_l, but depends on e_h through $a(\tau_h)$. Therefore, in the comparative statics results with respect to inequality e_h/e_l we had to state in the proposition that when changing e_h the threshold $\bar{z}$ is also changed such that top income relative to threshold income $n_h e_h/\bar{z} = z_h/\bar{z}$ and thus a remains unchanged.

Turning to the function $\Theta(\tau_h)$ we first note that it is continuous and strictly decreasing[46] on $[0, \tau_h^{\text{Laffer}}]$, with

$$\Theta(\tau_h = \tau_h^{\text{Laffer}}) < \infty \tag{A62}$$

$$\Theta(\tau_h = 0) = \left(\frac{[e_h]^{1+\chi}}{[e_l]^{1+\chi}} \right)^{\gamma} \geq 1, \tag{A63}$$

[46] $\Theta(\tau_h)$ is strictly decreasing in τ_h since, taking the derivatives of the numerator and the denominator, we obtain

$$\begin{aligned} &\frac{d\,[(1-\tau_h)e_h]^{1+\chi} + (1+\chi)\tau_h\bar{z} + (1+\chi)R(\tau_h)}{d\tau_h} \\ &\quad = -(1+\chi)(1-\tau_h)^{\chi}\,[e_h]^{1+\chi} + (1+\chi)\bar{z} + (1+\chi)\frac{dR(\tau_h)}{d\tau_h} \\ &\quad \leq (1+\chi)\frac{dR(\tau_h)}{d\tau_h} = \frac{d\,[e_l]^{1+\chi} + (1+\chi)R(\tau_h)}{d\tau_h} \end{aligned} \tag{A61}$$

since for all $\tau_h \leq \tau_h^{\text{Laffer}}$ we have $(1-\tau_h)^{\chi}\,[e_h]^{1+\chi} = z_h \geq \bar{z}$.

with equality only if $\gamma = 0$. Thus, $\Theta(\tau_h)$ and $\Gamma(\tau_h)$ intersect only once in $[0, \tau_h^{\text{Laffer}}]$, at $\tau_h = 0$ if $\gamma = 0$ and at $\tau_h \in (0, \tau_h^{\text{Laffer}})$ if $\gamma > 0$. This establishes item 2 in the proposition.

Finally, the comparative statics properties in item 3 are established as follows. Since the ratio defining $\Theta(\tau_h)$ is strictly larger than 1, an increase in γ shifts $\Theta(\tau_h)$ up without changing $\Gamma(\tau_h)$, and thus increases the Utilitarian tax rate τ_h^U. Finally, assume that e_l/e_h increases, reducing inequality, and maintain the assumption that $a = n_h e_h/\bar{z} = z_h/\bar{z}$ remains unchanged. Then $\Gamma(\tau_h)$ remains unchanged and

$$\Theta(\tau_h) = \left(\frac{\frac{[1-\tau_h)e_h]^{1+\chi}+\tau_h\bar{z}}{[1-\tau_h]^{\chi}[e_h]^{1+\chi}} + (1+\chi)\Phi_h\tau_h\left[1 - \frac{a-1}{a}\right]}{\frac{[e_l]^{1+\chi}}{[1-\tau_h]^{\chi}[e_h]^{1+\chi}} + (1+\chi)\Phi_h\tau_h\left[1 - \frac{a-1}{a}\right]} \right)^{\gamma} \tag{A64}$$

$$= \left(\frac{1 - \frac{\tau_h}{a} + (1+\chi)\Phi_h\frac{\tau_h}{a}}{\frac{[e_l/e_h]^{1+\chi}}{[1-\tau_h]^{\chi}} + (1+\chi)\Phi_h\frac{\tau_h}{a}} \right)^{\gamma} \tag{A65}$$

and thus the $\Theta(\tau_h)$ curve shifts down, reducing the Utilitarian tax rate τ_h^U.

Social Optimum: Proof of Proposition 6. — For part 1, the CEV transfers of the ex-ante heterogeneous types are given by $T_l = R$ and $T_h = \left[(1-\tau_h)^{1+\chi} - 1\right]\frac{(e_h)^{1+\chi}}{1+\chi} + \tau_h\bar{z} + R$, and thus, exploiting the government budget constraint (17), direct calculations give

$$\mathcal{V}(\tau_h, \Psi = 0) = \Phi_l T_l(\tau_h) + (1-\Phi_l)T_h(\tau_h) =$$
$$= [1-\Phi_l]\frac{(e_h)^{1+\chi}}{1+\chi}\left[(1+\chi\tau_h)(1-\tau_h)^{\chi} - 1\right], \tag{A66}$$

and we readily observe that $\mathcal{V}(\tau_h = 0; \Psi = 0) = 0$ and $\mathcal{V}(\tau_h; \Psi = 0)$ has a unique maximum at $\tau_h = 0$ with $\mathcal{V}'(\tau_h, \Psi = 0) = 0$, unless of course $\chi = 0$ (inelastic labor supply, no distortions) or $\Phi_l = 1$ (no rich households and thus no distortionary taxation and no transfers) in which case $\mathcal{V}(\tau_h; \Psi = 0) = 0$ for all τ_h.

Part 2 follows directly from the fact that $\mathcal{V}(\tau_h; \Psi = 1) = \mathcal{W}(\tau_h)$ and from Proposition 5.

Now we prove part 3. We recall that our welfare measure is given by

$$\mathcal{V}(\tau_h; \Psi) = (1-\Psi)\left[\Phi_l T_l(\tau_h) + (1-\Phi_l)T_h(\tau_h)\right] + \Psi T_u(\tau_h) \tag{A67}$$
$$= (1-\Psi)\mathcal{V}(\tau_h; \Psi = 0) + \Psi T_u(\tau_h). \tag{A68}$$

We have characterized $\mathcal{V}(\tau_h; \Psi = 0)$ in equation A66, and see that it is continuously differentiable and strictly decreasing in τ_h. Now we need to characterize

$T_u(\tau_h)$, which solves the equation

$$LHS(T_u) = \Phi_l \frac{\left[\frac{[e_l]^{1+\chi}}{1+\chi} + T_u\right]^{1-\gamma}}{1-\gamma} + (1-\Phi_l)\frac{\left[\frac{[e_h]^{1+\chi}}{1+\chi} + T_u\right]^{1-\gamma}}{1-\gamma}$$
$$\stackrel{!}{=} V_u(\tau_h) = \Phi_l V_l(\tau_h) + (1-\Phi_l)V_h(\tau_h) = \mathcal{W}(\tau_h). \tag{A69}$$

Thus the right hand side of this equation is simply Utilitarian social welfare characterized in Proposition 5, and thus is continuous and strictly increasing in $\tau_h \in [0, \tau_h^U)$, reaching its maximum at τ_h^U and is strictly decreasing thereafter. The left hand side is strictly increasing and continuous in T_u, and independent of τ_h. Furthermore

$$LHS(T_u = 0) = \mathcal{W}(\tau_h = 0) \tag{A70}$$

and thus $T_u(\tau_h = 0) = 0$. By the implicit function theorem $T_u(\tau_h)$ defined implicitly in equation (A69) is a differentiable function with

$$T_u'(\tau_h) = \frac{\mathcal{W}'(\tau_h)}{\Phi_l\left[\frac{[e_l]^{1+\chi}}{1+\chi} + T_u(\tau_u)\right]^{-\gamma} + (1-\Phi_l)\left[\frac{[e_h]^{1+\chi}}{1+\chi} + T_u(\tau_h)\right]^{-\gamma}} \tag{A71}$$

and thus $T_u(\tau_u)$ is differentiable in τ_h, strictly increasing $\tau_h \in [0, \tau_h^U)$ and strictly decreasing for $\tau_h > \tau_h^U$.

Therefore

$$\mathcal{V}(\tau_h; \Psi) = (1-\Psi)\mathcal{V}(\tau_h; \Psi = 0) + \Psi T_u(\tau_h) \tag{A72}$$

is the convex combination of two continuous functions in τ_h, one with weight $(1-\Psi)$ that is strictly decreasing on $[0, \tau_h^U]$, the other with weight Ψ that strictly increasing on $[0, \tau_h^U)$, and flat at τ_h^U. Finally

$$\mathcal{V}'(\tau_h = 0; \Psi) = (1-\Psi)\mathcal{V}'(\tau_h = 0; \Psi = 0) + \Psi T_u'(\tau_h = 0) \tag{A73}$$
$$= 0 + \Psi T_u'(\tau_h = 0) > 0 \tag{A74}$$

as long as $\Psi > 0$. Thus welfare is strictly increasing in τ_h at $\tau_h = 0$ as long as $\Psi > 0$. We conclude that

$$\tau_h^*(\Psi) = \arg\max_{\tau_h} \mathcal{V}(\tau_h; \Psi) \tag{A75}$$

is continuous and strictly increasing on $\Psi \in (0, 1)$ and satisfies $\tau_h^*(\Psi) \in (0, \tau_h^U)$

and

$$\lim_{\Psi\to 0} \tau_h^*(\Psi) = \tau_h^*(\Psi = 0) = 0 \tag{A76}$$

$$\lim_{\Psi\to 1} \tau_h^*(\Psi) = \tau_h^*(\Psi = 1) = \tau_h^U. \tag{A77}$$

B. Details of the Computational Approach

In order to solve the model outlined in this paper, we need three distinct algorithms: one that determines policy and value functions, one that solves for equilibrium quantities and prices, and one that delivers compensation payments.

B1. Computation of Policy and Value Functions

We solve for policy and value functions using the method of endogenous grid points Formally, these functions exist on the state space

(B1)
$$(j, s, \alpha, \eta, a) \in \{1, \ldots, J\} \times \{n, c\} \times \{-\sigma_\alpha, +\sigma_\alpha\} \times \{\eta_{s,1}, \ldots, \eta_{s,7}\} \times [0, \infty].$$

In order to be able to represent them on a computer, we however have to discretize the continuous elements of the state space, namely the asset dimension. For this purpose we chose a set of discrete points $\{\hat{a}^1 \ldots, \hat{a}^{100}\}$ such that the state space above can be approximated by

(B2)
$$(j, s, \alpha, \eta, a) \in \{1, \ldots, J\} \times \{n, c\} \times \{-\sigma_\alpha, +\sigma_\alpha\} \times \{\eta_{s,1}, \ldots, \eta_{s,7}\} \times \{\hat{a}^1 \ldots, \hat{a}^{100}\}.$$

Note that the choice of $\hat{a}^i$ is not straightforward. Specifically we let

$$\hat{a}^i = \bar{a} \cdot \frac{(1 + g_a)^{i-1} - 1}{(1 + g_a)^{99} - 1}, \tag{B3}$$

which leaves us with two parameters that define our discrete grid space. $\bar{a}$ is the upper limit of the asset grid which we chose such that no individual in our simulated model would like to save more than this amount.[47] A g_a of 0 would result in equidistantly spaced grid points Setting $g_a > 0$ the distance between two successive grid points $\hat{a}^i$ and $\hat{a}^{i+1}$ grows at the rate g_a in i. In our preferred parameterization we let $g_a = 0.08$. We consequently located many grid points at the lower end of the grid space where borrowing constraints may occur and therefore policy functions may have kinks or be sharply curved. At the upper end of the grid space where policy and value functions are almost linear, we

[47]In our model this leads to $\bar{a} = 1800$.

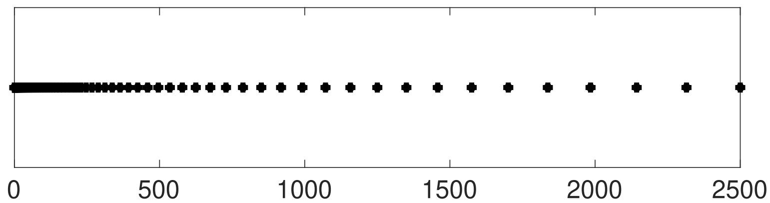

FIGURE B1. DISCRETIZED ASSET STATE SPACE

consequently use a much smaller amount of points. Figure B1 visualizes our discrete asset grid.

The discretization of the asset state space makes the solution for policy and value functions feasible via backward induction. We start out by solving the optimization problem at the last possible age an individual may have J. Since the agent is retired and dies with certainty, she will consume all her remaining resources and work zero hours,

$$c(J,s,\alpha,\eta,\hat{a}^i) = \frac{(1+r_n)\hat{a}^i + p(s,\alpha,\eta)}{1+\tau_c} \quad , \quad n(J,s,\alpha,\eta,\hat{a}^i) = 0 \quad , \tag{B4}$$

$$a'(J,s,\alpha,\eta,\hat{a}^i) = 0 \quad \text{for all } i = 1,\ldots,100. \tag{B5}$$

In order to simplify the computation of the value function we will actually keep track of two different value functions, the one for consumption and the one for labor. This is possible due to the additive separability assumption we made. Consequently we have

$$v_c(J,s,\alpha,\eta,\hat{a}^i) = \frac{[c(J,s,\alpha,\eta,\hat{a}^i)]^{1-\gamma}}{1-\gamma} \quad \text{and} \quad v_n(J,s,\alpha,\eta,\hat{a}^i) = 0. \tag{B6}$$

Knowing the policy and value function in the last period of life, we can now iterate backward over ages to determine the remaining household decisions. Since the algorithm is very similar for retired and working individuals, we will restrict ourselves to the case of workers. Assume that we had already calculated policy and value functions at age $j+1$. The problem we need to solve for an individual at state (j,s,α,η,a) then reads

$$\begin{aligned} \max_{c,n,a'} \quad & \frac{c^{1-\gamma}}{1-\gamma} - \alpha\frac{n^{1+\chi}}{1+\chi} \\ & + \beta\psi_{j+1}\sum_{\eta'}\pi_s(\eta'|\eta)\left[v_c(j+1,s,\alpha,\eta',a') - v_l(j+1,s,\alpha,\eta',a')\right] \end{aligned} \tag{B7}$$

subject to the constraints

$$(1+\tau_c)c + a' + T(we(j,s,\alpha,\eta)n) + T_{ss}(we(j,s,\alpha,\eta)n) = (1+r_n)a + b_j(s,\eta) + we(j,s,\alpha,\eta)n \tag{B8}$$

as well as $0 \le n \le 1$ and $a' \ge 0$. The first order conditions (ignoring the constraint on n and the borrowing constraint) then are

$$c = [\lambda(1+\tau_c)]^{-1/\gamma} \tag{B9}$$

$$\alpha n^{\chi} = \lambda we(j,s,\alpha,\eta)\left[1 - T'(we(j,s,\alpha,\eta)n) - T'_{ss}(we(j,s,\alpha,\eta)n)\right] \tag{B10}$$

$$\lambda = \beta\psi_{j+1}(1+r'_n)(1+\tau'_c)\sum_{\eta'} c(j+1,s,\alpha,\eta',a'^{-\gamma}. \tag{B11}$$

We now apply the method of endogenous grid points as follows: We assume that savings for tomorrow would amount to $a' = \hat{a}^i$ for all $i = 1,\ldots,100$. Under this assumption, we can compute for each combination of (s,α,η) the respective λ from the last first-order condition. λ then defines a certain level of consumption $c^e(j,s,\alpha,\eta,\hat{a}^i)$ and labor supply $n^e(j,s,\alpha,\eta,\hat{a}^i)$.[48] Plugging these into the budget constraint, we can determine the endogenous grid point as

$$a^e(j,s,\alpha,\eta,\hat{a}^i) = \frac{1}{1+r_n}\Big[(1+\tau_c)c^e(j,s,\alpha,\eta,\hat{a}^i) + a' + T(we(j,s,\alpha,\eta)n) + T_{ss}(we(j,s,\alpha,\eta)n) - b_j(s,\eta) - we(j,s,\alpha,\eta)n^e(j,s,\alpha,\eta,\hat{a}^i)\Big]. \tag{B12}$$

Finally, we can compute the value functions as

$$v^e_c(j,s,\alpha,\eta,\hat{a}^i) = \frac{[c^e(j,s,\alpha,\eta,\hat{a}^i)]^{1-\gamma}}{1-\gamma} + \beta\psi_{j+1}\sum_{\eta'}\pi_s(\eta'|\eta)v_c(j+1,s,\alpha,\eta',\hat{a}_i) \tag{B13}$$

$$v^e_n(j,s,\alpha,\eta,\hat{a}^i) = \frac{\alpha[n^e(j,s,\alpha,\eta,\hat{a}^i)]^{1+\chi}}{1+\chi} + \beta\psi_{j+1}\sum_{\eta'}\pi_s(\eta'|\eta)v_n(j+1,s,\alpha,\eta',\hat{a}_i). \tag{B14}$$

[48] Note that we can not solve for labor supply analytically due to the non linearity of the labor earnings tax schedule. Instead we use a quasi-Newton root finding routine to determine the solution to the respective first order condition. We thereby have to respect the constraint $0 \le n \le 1$ as well as the fact that there is a cap on contributions to the social security system. However, due to the additive separability of the utility function in consumption and labor supply, the constraints on n will not affect the individual's choice of consumption c.

Using the interpolation data

(B15)
$$\left\{a^e(j,s,\alpha,\eta,\hat{a}^i), c^e(j,s,\alpha,\eta,\hat{a}^i)\right\}_{i=1}^{100}, \quad \left\{a^e(j,s,\alpha,\eta,\hat{a}^i), n^e(j,s,\alpha,\eta,\hat{a}^i)\right\}_{i=1}^{100},$$
(B16)
$$\left\{a^e(j,s,\alpha,\eta,\hat{a}^i), v_c^e(j,s,\alpha,\eta,\hat{a}^i)\right\}_{i=1}^{100}, \quad \left\{a^e(j,s,\alpha,\eta,\hat{a}^i), v_n^e(j,s,\alpha,\eta,\hat{a}^i)\right\}_{i=1}^{100},$$
(B17)

we can finally determine the (discrete) policy and value functions

$$c(j,s,\alpha,\eta,\hat{a}^i), \quad n(j,s,\alpha,\eta,\hat{a}^i), \quad v_c(j,s,\alpha,\eta,\hat{a}^i) \quad \text{and} \quad v_n(j,s,\alpha,\eta,\hat{a}^i) \tag{B18}$$

for each today's asset value $\hat{a}^i, i = 1, \dots, 100$ by piece-wise linear interpolation.[49]

Before applying this interpolation scheme, we however check for the occurrence of liquidity constraints. Liquidity constraints occur if $a^e(j,s,\alpha,\eta,0) > 0$. In this case, we extend the above interpolation data by another point of value 0 on the left. The policy and value functions at this point are determined under the assumption that $a = a' = 0$, i.e. the policy function values solve the equation system

(B19)
$$\frac{c^{-\gamma}}{1+\tau_c} = \lambda$$
$$\alpha n^\chi = \lambda w e(j,s,\alpha,\eta)\left[1 - T'(we(j,s,\alpha,\eta)n) - T'_{ss}(we(j,s,\alpha,\eta)n)\right] \tag{B20}$$
(B21)
$$(1+\tau_c)c = b_j(s,\eta) + we(j,s,\alpha,\eta)n - T(we(j,s,\alpha,\eta)n) - T_{ss}(we(j,s,\alpha,\eta)n).$$

B2. Determining Aggregate Quantities and Prices

Our algorithm to determine aggregate quantities and prices follows closely the Gauss-Seidel method already proposed in Auerbach and Kotlikoff (1987). Specifically, in order to determine an equilibrium path of the economy, we start with an initial guess of quantities $\{K_t, L_t\}_{t\geq 0}$ as well as tax rates $\{\tau_l, \tau_{ss,t}\}_{t\geq 0}$ and transfers $\{Tr_t\}_{t\geq 0}$. Our algorithm then iterates over the following steps:

1) Determine the factor prices $\{r_t, w_t\}_{t\geq 0}$ that correspond to the quantities $\{K_t, L_t\}_{t\geq 0}$.
2) Solve the household optimization problem using these factor prices and the guesses for tax rates. Determine the measure of households.

[49]We do not interpolate v_c^e and v_n^e directly, but rather $[(1-\gamma)v_c^e]^{1/(1-\gamma)}$ and $[(1+\chi)v_n^e]^{1/(1+\chi)}$ and then transform them back to their original shape. This leads to much more accurate results in the high curvature region of the asset grid.

3) Solve for the tax rate τ_l that balances the intertemporal budget constraint of the government by means of a quasi-Newton root finding method. Then calculate the path of government debt $\{B_t\}_{t\geq 0}$.

4) Determine the budget balancing payroll tax rates $\tau_{ss,t}$ using the social security system's sequential budget constraints.

5) Calculate lump-sum transfers Tr such that the sum of transfers equals the sum of bequests left by the non-surviving households.

6) Determine the new quantities $\{K_t^{new}, L_t^{new}\}_{t\geq 0}$ by aggregating individual decisions. Calculate updated quantities through

$$K_t = (1-\omega)K_t + \omega K_t^{new} \quad \text{and} \quad L_t = (1-\omega)L_t + \omega L_t^{new}. \tag{B22}$$

ω thereby serves as a damping factor. Our preferred value for ω is 0.3.

7) Check whether the economy is in equilibrium, i.e.

$$\max_{t\geq 0} \left| \frac{Y_t - C_t - I_t - G_t}{Y_t} \right| < \varepsilon. \tag{B23}$$

This means that the relative difference between aggregate demand and supply of goods should be smaller than a given tolerance level. If this is not the case, start with the updated guesses of quantities, tax rates and transfers at step 1. If this is the case, we have found an equilibrium path of the economy. To determine the initial equilibrium we use a tolerance level of $\varepsilon = 10^{-9}$ while for the transition path we set $\varepsilon = 10^{-6}$.

B3. Calculation of Compensating Transfers

The calculation of compensating transfers is straightforward. In order to do so, we use a quasi-Newton root finding method that numerically determines the solutions to the equations

$$v_1(j, s, \alpha, \eta, a + \Psi_0(j, s, \alpha, \eta, a)) = v_0(j, s, \alpha, \eta, a) \tag{B24}$$

and

$$Ev_t(j = 1, s, \alpha, \bar{\eta}, \Psi_t) = Ev_0(j = 1, s, \alpha, \bar{\eta}, 0), \tag{B25}$$

respectively. Note that in each iteration of the root finding method, we have to solve for the optimal household decisions.

C. Definition of a Stationary Recursive Competitive Equilibrium

DEFINITION 7: *Given government expenditures G, government debt B, a tax system characterized by (τ_c, τ_k, T) and a social security system characterized by $(\tau_{ss}, \bar{z}_{ss})$, a stationary recursive competitive equilibrium with population growth is a collection of value and policy functions (v, c, n, a') for the household, optimal input choices (K, L) of firms, transfers b, prices (r, w) and an invariant probability measure Φ with the following properties:*

1) *[Household maximization]: Given prices (r, w), transfers b_j given by (29) and government policies $(\tau_c, \tau_k, T, \tau_{ss}, \bar{z}_{ss})$, the value function v satisfies the Bellman equation (24), and (c, n, a') are the associated policy functions.*

2) *[Firm maximization]: Given prices (r, w), the optimal choices of the representative firm satisfy*

$$r = \Omega\epsilon \left[\frac{L}{K}\right]^{1-\epsilon} - \delta_k \tag{C1}$$

$$w = \Omega(1-\epsilon) \left[\frac{K}{L}\right]^{\epsilon}. \tag{C2}$$

3) *[Government Budget Constraints]: Government policies satisfy the government budget constraints (26) and (27).*

4) *[Market clearing]:*

   a) *The labor market clears:*

$$L = \int e(j, s, \alpha, \eta) n(j, s, \alpha, \eta, a) d\Phi \tag{C3}$$

   b) *The capital market clears*

$$(1 + g_n)(K + B) = \int a'(j, s, \alpha, \eta, a) d\Phi \tag{C4}$$

   c) *The goods market clears*

$$Y = \int c(j, s, \alpha, \eta, a) d\Phi + (g_n + \delta)K + G \tag{C5}$$

5) *[Consistency of Probability Measure Φ]: The invariant probability measure is consistent with the population structure of the economy, with the exogenous processes π_s, and the household policy function $a'(.)$. A formal definition is provided in Appendix D.*

D. Definition of Invariant Probability Measure

First we construct the share of the population in each age group. Let $\tilde{\mu}_1 = 1$, and for each $j \in \{2, \ldots, J\}$ define recursively

$$\tilde{\mu}_j = \frac{\psi_j \tilde{\mu}_{j-1}}{1+g_n}. \tag{D1}$$

Then the share of the population in each age group is given by

$$\mu_j = \frac{\tilde{\mu}_j}{\sum_\iota \tilde{\mu}_\iota}. \tag{D2}$$

Next, we construct the measure of households of age 1 across characteristics (s, α, η, a). By assumption (see the calibration section, Section III of the paper) newborn households enter the economy with zero assets, $a = 0$ and at the mean idiosyncratic productivity shock $\bar{\eta}$. The share of college-educated households is exogenously given by ϕ_c and $\phi_n = 1 - \phi_c$, and the fixed effect is drawn from a discrete pdf $\phi_s(\alpha)$. Thus

$$\Phi(\{j=1\}, \{\alpha\}, \{s\}, \{\bar{\eta}\}, \{0\}) = \mu_1 \phi_s \phi_s(\alpha) \tag{D3}$$

for $s = \{n, c\}$ and zero else.

Finally we construct the probability measure for all ages $j > 1$. For all Borel sets of assets $\mathcal{A}$ we have

$$\begin{aligned} &\Phi(\{j+1\}, \{\alpha\}, \{s\}, \{\eta'\}, \mathcal{A}) \\ &\qquad = \frac{\psi_{j+1}\pi_s(\eta'|\eta)}{1+g_n} \int \mathbb{1}_{\{a'(j,s,\alpha,\eta,a) \in \mathcal{A}\}} \Phi(\{j\}, \{\alpha\}, \{s\}, \{\eta\}, da) \end{aligned} \tag{D4}$$

where

$$\int \mathbb{1}_{\{a'(j,s,\alpha,\eta,a) \in \mathcal{A}\}} \Phi(\{j\}, \{\alpha\}, \{s\}, \{\eta\}, da) \tag{D5}$$

is the measure of assets a today such that, for fixed (j, s, α, η), the optimal choice today of assets for tomorrow, $a'(j, s, \alpha, \eta, a)$ lies in $\mathcal{A}$.

E. Details of the Calibration

E1. Markov Chain for Labor Productivity

The Markov chain governing idiosyncratic labor productivity for both education groups is given by

$s = n$

i, j	1	2	3	4	5	6	7
1	0.969909	0.029317	0.000332	0.000002	0.000000	0.000440	0.000000
2	0.007329	0.970075	0.021989	0.000166	0.000000	0.000440	0.000000
3	0.000055	0.014659	0.970130	0.014659	0.000055	0.000440	0.000000
4	0.000000	0.000166	0.021989	0.970075	0.007329	0.000440	0.000000
5	0.000000	0.000002	0.000332	0.029317	0.969909	0.000440	0.000000
6	0.000000	0.000000	0.002266	0.000000	0.000000	0.970000	0.027734
7	0.000000	0.000000	0.000000	0.000000	0.000000	0.288746	0.711254
$\exp(\eta_{n,i})$	0.1354	0.3680	1.0000	2.7176	7.3853	19.7204	654.0124

and

$s = c$

i, j	1	2	3	4	5	6	7
1	0.960937	0.029046	0.000329	0.000002	0.000000	0.009686	0.000000
2	0.007261	0.961102	0.021786	0.000165	0.000000	0.009686	0.000000
3	0.000055	0.014524	0.961157	0.014524	0.000055	0.009686	0.000000
4	0.000000	0.000165	0.021786	0.961102	0.007261	0.009686	0.000000
5	0.000000	0.000002	0.000329	0.029046	0.960937	0.009686	0.000000
6	0.000000	0.000000	0.047247	0.000000	0.000000	0.949922	0.002831
7	0.000000	0.000000	0.000000	0.000000	0.000000	0.288746	0.711254
$\exp(\eta_{c,i})$	0.2362	0.4860	1.0000	2.0575	4.2334	8.3134	654.0124

E2. Numerical Computation of Policy Elasticities

In order to be able to apply the formula for the Laffer tax rate proposed in Proposition 1 in our full quantitative simulation model, we have to calculate the policy elasticities $\epsilon(z_h)$ and $\epsilon(\tau_a(\bar{z}))$. To this end, we proceed in the following steps:

1) We start from the initial equilibrium described in Section IV and compute a transition path that results from keeping the top tax rate at its initial equilibrium level of $\tau_h = 0.396$, but setting the top tax threshold $\bar{z}_h$ such that exactly the top 1% earners are hit by the top rate. For each period $t \geq 1$ of the transition, we can then calculate average top 1% earnings as

$$\text{(E1)} \quad z^0_{h,t} = \int \mathbb{1}_{z_t(j,s,\alpha,\eta,a) \geq \bar{z}_t} \; z_t(j,s,\alpha,\eta,a) \; d\Phi_t$$
$$\text{with} \quad z_t(j,s,\alpha,\eta,a) = w_t e(j,s,\alpha,\eta) n_t(j,s,\alpha,\eta,a).$$

The average tax rate at the top earnings threshold then is

$$\tau_a^0(\bar{z}_t) = \frac{T_t(\bar{z}_t)}{\bar{z}_t}. \tag{E2}$$

2) We now increase the top marginal net-of-tax rate by an amount δ and calculate a new equilibrium path. From this, we obtain a new value for average top 1% earnings $z^1_{h,t}$ and a new value for the average tax rate $\tau_a^1(\bar{z}_t)$. In our numerical calculations, we use $\delta = 0.01$.

3) The relevant elasticities for the Laffer tax rate formula then are

(E3)
$$\epsilon_t(z_{h,t}) = \frac{z^1_{h,t} - z^0_{h,t}}{\delta} \cdot \frac{0.604}{z^0_{h,t}} \quad \text{and} \quad \epsilon_t(\tau_a(\bar{z}_t)) = \frac{\tau_a^1(\bar{z}_t) - \tau_a^0(\bar{z}_t)}{\delta} \cdot \frac{0.604}{\tau_a^0(\bar{z}_t)}.$$

The elasticities we derive from this procedure are listed in the columns *initial* within Table 9. We then repeat the above exercise, but instead of starting from the initial equilibrium tax rate $\tau_h = 0.396$, we start from the actual Laffer tax rate. The resulting elasticities are then shown in the columns *final*.

E3. The Social Security System

We use the pension formula for the US social security system to calculate pension payments. Specifically, for a given average labor earnings $\tilde{z}$ we set

(E4)
$$p(s,\alpha,\eta) = f(\tilde{z}) = \begin{cases} r_1\tilde{z} & \text{if } \tilde{z} < b_1 y^{med} \\ r_1 b_1 y^{med} + r_2(\tilde{z} - b_1 y^{med}) & \text{if } \tilde{z} < b_2 y^{med} \\ r_1 b_1 y^{med} + r_2(b_2 - b_1) y^{med} + r_3(\tilde{z} - b_2 y^{med}) & \text{otherwise} \end{cases}$$

Here r_1, r_2, r_3 are the respective replacement rates and b_1 and b_2 the bend points. We express these points in terms of median household *income* y^{med} which is the median of income from labor and assets (including bequests and pension payments). We use $y^{med} = 50,000$ as a reference value for this (see US Census Bureau for 2009). Consequently, the bend points are $b_1 = 0.184$ and $b_2 = 1.144$ and the respective replacement rates are $r_1 = 0.90$, $r_2 = 0.32$ and $r_3 = 0.15$. The maximum amount of pension benefit a household can receive is therefore $30,396$, or 0.608 times the median income. All data is taken from the information site of the social security system for 2012. Finally, we calibrate the contribution cap of the pension system $\bar{z}_{ss}$ in order to obtain a contribution rate of 12.4 percent.

F. Additional Figures

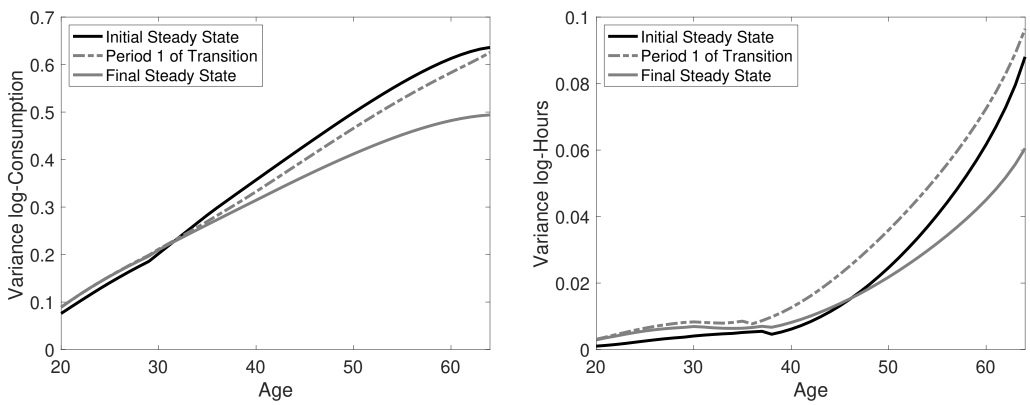

Figure F1. Variance of Consumption and Hours over the Life Cycles, Entire Population

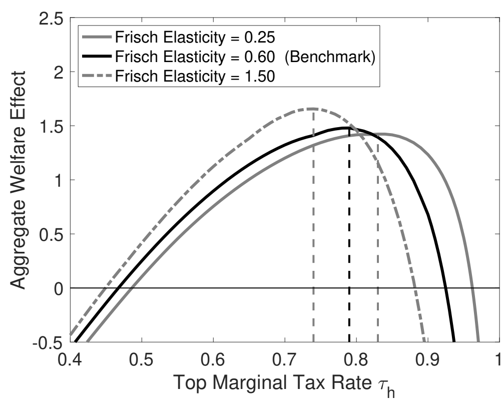

Figure F2. Aggregate Welfare as Function of τ_h, Different Frisch Elasticities

G. Sensitivity analysis

When doing sensitivity analysis, we have to partly recalibrate the model in order to make results comparable. For each different specification of the model we therefore recalibrate the technology level Ω such that the wage rate for effective labor is again equal to $w = 1$ as well as the depreciation rate δ_k such that the interest rate remains at 4%. The former ensures stability of our computational algorithm, the latter is necessary to guarantee equal weights of generations in the social welfare function. Finally we recalibrate the taste parameter for the disutility of labor λ so that average hours worked remain at 33% of the time endowment. We furthermore do some specific adjustments for different sensitivity scenarios which we outline in the following.

G1. Size of the Income Effect

When we impose log preferences the relationship between hours worked and individual labor productivity changes dramatically. As a consequence we have to completely recalibrate the total income process. The following table shows which probabilities and productivity levels we have to choose in this case to obtain the same fit for the earnings and wealth distribution in our model:

$s = n$

1	0.969945	0.029318	0.000332	0.000002	0.000000	0.000403	0.000000
2	0.007330	0.970111	0.021990	0.000166	0.000000	0.000403	0.000000
3	0.000055	0.014660	0.970166	0.014660	0.000055	0.000403	0.000000
4	0.000000	0.000166	0.021990	0.970111	0.007330	0.000403	0.000000
5	0.000000	0.000002	0.000332	0.029318	0.969945	0.000403	0.000000
6	0.000000	0.000000	0.012043	0.000000	0.000000	0.969903	0.018054
7	0.000000	0.000000	0.000000	0.000000	0.000000	0.269999	0.730001
$\exp(\eta_{n,i})$	0.1722	0.4149	1.0000	2.4101	5.8085	18.0227	374.1023

and

$s = c$

i, j	1	2	3	4	5	6	7
1	0.960202	0.029024	0.000329	0.000002	0.000000	0.010444	0.000000
2	0.007256	0.960366	0.021769	0.000164	0.000000	0.010444	0.000000
3	0.000055	0.014513	0.960421	0.014513	0.000055	0.010444	0.000000
4	0.000000	0.000164	0.021769	0.960366	0.007256	0.010444	0.000000
5	0.000000	0.000002	0.000329	0.029024	0.960202	0.010444	0.000000
6	0.000000	0.000000	0.068922	0.000000	0.000000	0.928130	0.002948
7	0.000000	0.000000	0.000000	0.000000	0.000000	0.269999	0.730001
$\exp(\eta_{c,i})$	0.2809	0.5300	1.0000	1.8867	3.5597	6.3118	374.1023

G2. Persistence of High Productivity States

To make the highest productivity state completely permanent we again have to adjust the transition probabilities in our model. This time we assume that only at age 30 there is a certain probability that individuals can climb up to the highest productivity region. This probability is the same for each individual of an education level. In order to determine this probability we calculate the fraction of individuals in the highest productivity region between the ages 30 and j_r for each education level in the benchmark model. We then choose the probability to get a permanent very high income shock in the sensitivity model such that the fraction of households in the highest income region is exactly the same as in the benchmark model.