A Appendix: Analysis of the Borrowing Constrained Agents

In general, for all $z \in Z$ there exists a current minimal accumulated asset level $k(z)$ above which agents are not borrowing constrained (see Figure A.1 below for illustration). For these agents, i.e. for those with $(k, z) \in [k, k(z)] \times Z$ and the next period savings $k^+$ equal to $k$, the Euler equation is satisfied in the form of inequality

$$u'(c) > \beta \sum_{z^+} u'(c^+) \left[1 + y^+_k - (\pi(x^+) + \pi'(x^+) x^+_k) x^+_k \right] Q(z, z^+),$$

where

$$c = (1 - \pi(x)) x + k - k,$$
$$c^+ = (1 - \pi(x^+)) x^+ + k - h(z^+, k; K, \pi, \pi'),$$
$$y^+_k = r(K),$$
$$x^+ = x(z^+, k; K),$$
$$x^+_k = x_k(z^+, k; K)$$

implying $h = k$.

Taking this into account we can define the extended Euler equation operator,

$$\mathcal{F}(h) = \begin{cases} 
\mathcal{F}(h) & \text{for } (z, k) \in Z \times [k(z), k(z)], \\
h & \text{for } (z, k) \in Z \times [k, k(z)], 
\end{cases}$$

and thus the operator equation in the form $\mathcal{F}(\tilde{h}) = 0$ determines the savings function with the segment of constrained savings, $\tilde{h}$.

For the sake of brevity we present here only the effect of the borrowing constrained agents at the lowest shock, $z$, with the next period capital $k$. The stationarity of the distribution functions implies

$$\lambda(z^+, k) = \int_{k}^{k(z)} \lambda(z, k) Q(z, z^+) dk + \sum_{z \neq z} \lambda(z, h^{-1}(z, k^+; K, \pi, \pi')) Q(z, z^+),$$

where $\lambda(z^+, k)$ is the mass of agents with the next period capital $k$ and the next-period shock $z^+$. Let us note such amended distribution function by $\tilde{\lambda}$. 

34
Figure A.1: Policy functions for the next period capital stock. An example with two productivity shocks $\bar{z} > \tilde{z}$. There is an exogenous lower bound $k_L$ and an endogenous upper bound $k^* < k_H$. The stationary distribution has a unique ergodic set $E = [k_L, k^*]$. Agents with shock $\tilde{z}$ and capital stock $k < k(\tilde{z})$ are borrowing constrained.

Clearly, it means that the amended distribution function has a discontinuity at $\tilde{\lambda}(\cdot, k)$ in the sense that

$$\tilde{\lambda}(z, k) > \lim_{k \downarrow k} \tilde{\lambda}(z, k),$$

for all $z \in Z$. However, since we assume here that functions are Lebesgue integrable, it follows that

$$\int_k^{\bar{k}(\pi)} \tilde{\lambda}(z, k)dk = \int_k^{\bar{k}(\pi)} \lambda(z, k)dk,$$

for any $z \in Z$ and the distribution functions $\tilde{\lambda}$ and $\lambda$ are equivalent. So we can simply consider only the distribution function $\lambda$ given by $\mathcal{L}$ in (11) and $h$ given by (10).
B Appendix: Proofs

B.1 Proof of Theorem 1

For the first order conditions for the Ramsey problem in (15)-(16) we define

\[ J(\varepsilon) = \int_{\bar{z}(\varepsilon)}^{\pi(\varepsilon)} \left\{ \mathcal{W}[z, x; \hat{K}, \hat{\pi}(x), \hat{\pi}'(x)] + \mu \mathcal{G}[z, x; \hat{K}, \hat{\pi}(x), \hat{\pi}'(x)] \right\} dx \]

\[ + \sum_{z \in Z \setminus \{\bar{z}(\varepsilon)\}} \int_{\hat{z}(\varepsilon)}^{\pi(\varepsilon)} \left\{ \mathcal{W}[z, x; \hat{K}, \hat{\pi}(x), \hat{\pi}'(x)] + \mu \mathcal{G}[z, x; \hat{K}, \hat{\pi}(x), \hat{\pi}'(x)] \right\} dx \]

\[ + \int_{\hat{z}(\varepsilon)}^{\bar{z}(\varepsilon)} \left\{ \mathcal{W}[z, x; \hat{K}, \hat{\pi}(x), \hat{\pi}'(x)] + \mu \mathcal{G}[z, x; \hat{K}, \hat{\pi}(x), \hat{\pi}'(x)] \right\} dx, \]

where

\[ \hat{\pi}(x) \equiv \pi^*(x) + \varepsilon \delta \pi(x), \]
\[ \hat{\pi}'(x) \equiv \pi'^*(x) + \varepsilon \delta \pi'(x), \]
\[ \hat{K} \equiv K^* + \varepsilon \delta K. \]

The dependence of the bounds on the value of shocks \( z \in Z \) makes our problem a little harder than the standard calculus of variations problem. However, as Theorem 1 states, we construct the variations (the perturbation of functions from the optimum) being zero at all (interior) bounds – see Figure B.1 below. Therefore, only the values of the government policy at the boundaries of the maximal interval, \( \pi(\bar{z}(\varepsilon)) \) and \( \pi(\bar{z}(\varepsilon)) \), are free while all other interior bounds are fixed. Then the condition

\[ J'(0) = \lim_{\varepsilon \to 0} \frac{dJ(\varepsilon)}{d\varepsilon} = 0 \]

leads to the following first-order conditions.

\[ J'(0) = \sum_{z \in Z} \int_{\hat{z}(\varepsilon)}^{\bar{z}(\varepsilon)} \left\{ \mathcal{W}_\pi[z, x; K^*, \pi^*(x), \pi'^*(x)] + \mu \mathcal{G}_\pi[z, x; K^*, \pi^*(x), \pi'^*(x)] \right\} \delta \pi(x) \]
\[ + \mathcal{W}_K[z, x; K^*, \pi^*(x), \pi'^*(x)] + \mu \mathcal{G}_K[z, x; K^*, \pi^*(x), \pi'^*(x)] \delta K \]
\[ - \mathcal{W}_z[z, x; K^*, \pi^*(x), \pi'^*(x)] \big|_{\hat{z}(\varepsilon)} + \mu \mathcal{G}_z[z, x; K^*, \pi^*(x), \pi'^*(x)] \big|_{\hat{z}(\varepsilon)} \big( -\delta \bar{z} \big) \]
\[ + \mathcal{W}[z, x; K^*, \pi^*(x), \pi'^*(x)] \big|_{\bar{z}(\varepsilon)} + \mu \mathcal{G}[z, x; K^*, \pi^*(x), \pi'^*(x)] \big|_{\bar{z}(\varepsilon)} \delta \bar{z}, \]

where the last two lines come from the fact that the lower and upper bounds are free.

Using the following notation

\[ \mathcal{W}(z, x) \equiv \mathcal{W}[z, x; K^*, \pi^*(x), \pi'^*(x)] \]
Figure B.1: Variations Around Optimal Tax Schedule

and

\[ L(z, x) \equiv W(z, x) + \mu G(z, x), \]

integration by parts delivers

\[
\int_{\mathbb{Z}(z)}^{\pi(z)} L_{r'} (z, x) \delta \pi'(x) dx = \left[ L_{r'} (z, x) \delta \pi(x) \right]_{\mathbb{Z}(z)}^{\pi(z)} - \int_{\mathbb{Z}(z)}^{\pi(z)} \frac{d}{dx} L_{r'} (z, x) \delta \pi(x) dx.
\]

Thus we can rewrite the formula above in a more compact form as

\[
J'(0) = \sum_{z \in \mathbb{Z}} \int_{\mathbb{Z}(z)}^{\pi(z)} \left\{ \left[ L_p(z, x) - \frac{d}{dx} L_{r'}(z, x) \right] \delta \pi(x) + L_K(z, x) \delta \pi \right\} dx \quad \text{(B.1)}
\]

\[
+ \left[ L_{r'}(z, x) \delta \pi(x) \right]_{\mathbb{Z}(z)}^{\pi(z)} - L(z, x) \big|_{\mathbb{Z}(z)}^{\pi(z)} (-\delta \pi) + L(z, x) \big|_{\pi(z)}^{\pi(x)} \delta \pi.
\]

At the free upper bound, the variation at the end-value of the policy function, \( \delta \pi \), can be expressed as

\[ \delta \pi \equiv \pi (x + \delta x) - \pi^* (x) = \pi (x) + \pi'' (x) \delta x - \pi^* (x), \]

and

\[ \delta \pi (x) = \pi (x) - \pi'' (x) \delta x. \]
This implies that
\[ \delta \pi(x) = \delta \pi - \pi''(x) \delta x, \] (B.2)
i.e. the variance of the policy function at the upper bound can be expressed as a function of the variance at the end-value of policy function, \( \delta \pi \), and the variance at the end-value of the taxable activity value, \( \delta \pi \).

We can similarly specify the variation of the start-value of the policy function,
\[ \delta \pi(x) = -\delta \pi + \pi''(x) \delta x. \] (B.3)

The situation is more complicated at the equality constrained endpoint \( \pi(z) \), because the upper bound for capital, \( k \), is implicitly given by the saving function, \( k = h(z, k; K, \pi) \).

The total variance differential is
\[ k \left[ h_k(z, k; K, \pi) - 1 \right] + h_K(z, k; K, \pi) \delta K + h_\pi(z, k; K, \pi) \delta \pi = 0 \]
and thus
\[ \delta k = \left[ 1 - h_k(z, k; K, \pi) \right]^{-1} \left[ h_K(z, k; K, \pi) \delta K + h_\pi(z, k; K, \pi) \delta \pi \right]. \] (B.4)

As \( x = \{ x_k(z, k; K), x_K(z, k; K) \} \), we determine the variation
\[ \delta x = x_k(z, k; K) \delta K + x_K(z, k; K) \delta \pi, \] (B.5)
where we can further substitute for \( \delta k \) from (B.4)
\[ \omega_K \equiv \frac{\delta k}{\delta \pi} = \frac{h_K(z, k; K, \pi)}{1 - h_k(z, k; K, \pi)}, \]
\[ \omega_\pi \equiv \frac{\delta \pi}{\delta \pi} = \frac{h_\pi(z, k; K, \pi)}{1 - h_k(z, k; K, \pi)}. \]

Going back to (B.1) and using (B.2) and (B.3) we obtain
\[ J'(0) = \sum_{z \in Z} \int_{\pi(z)} \left[ \left( L_s(z, z) - \frac{d}{dx} L_{s'}(z, x) \right) \delta \pi(x) + L_{K}(z, x) \delta K \right] dx \] (B.6)
\[ + \ L_{s'}(z, z) \left[ \frac{d}{dx} \delta \pi(x) \right]_{\pi(z)} [\pi'(x) L_{s'}(z, x) - L(z, x)]_{\pi(z)} \delta \pi + \ L_{s'}(z, x) \left[ \delta \pi(x) \right]_{\pi(z)} [\pi'(x) L_{s'}(z, x) - L(z, x)]_{\pi(z)} \delta \pi. \]

Since the upper bound is equality constrained, \( \delta \pi \) is not independent and we need to
use (B.5) to obtain
\[
J'(0) = \sum_{z \in \mathbb{Z}} \int_\pi^x \left\{ \left[ L_\pi(z, x) - \frac{d}{dx} L_{\pi'}(z, x) \right] \delta \pi(x) + L_K(z, x) \delta K \right\} dx \quad (B.7)
\]
\[
+ L_{\pi'}(z, x) |_{\pi(x)} \delta \pi + [L(z, x) - \pi'(x) L_{\pi'}(z, x)] |_{\pi(x)} \delta K
\]
\[
+ L_{\pi'}(\pi(x)) |_{\pi(x)} \delta \pi + [L(\pi(x)) - \pi'(x) L_{\pi'}(\pi(x))] |_{\pi(x)} \left\{ (\pi_K + \pi_K \pi_K) \delta K + \pi_K \pi_K \delta \pi \right\}.
\]
For the variation of the aggregate capital \( \delta K \) we use equation (14)
\[
\delta K = \sum_{z \in \mathbb{Z}} \int_\pi^x \left\{ K_\pi(z, x) \delta \pi(x) + K_{\pi'}(z, x) \delta \pi'(x) + K_K(z, x) \delta K \right\} dx
\]
\[
- K(z, x) |_{\pi(x)} \delta \pi + K(\pi(x)) |_{\pi(x)} \delta \pi,
\]
again using integration by parts and the conditions for the free boundary points,
\[
\delta K = \sum_{z \in \mathbb{Z}} \int_\pi^x \left\{ \left[ K_\pi(z, x) - \frac{d}{dx} K_{\pi'}(z, x) \right] \delta \pi(x) + K_K(z, x) \delta K \right\} dx
\]
\[
+ K_{\pi'}(z, x) |_{\pi(x)} \delta \pi' + [K(z, x) - \pi'(x) K_{\pi'}(z, x)] |_{\pi(x)} \delta K
\]
\[
+ K_{\pi'}(\pi(x)) |_{\pi(x)} \delta \pi + [K(\pi(x)) - \pi'(x) K_{\pi'}(\pi(x))] |_{\pi(x)} \left\{ (\pi_K + \pi_k \pi_K) \delta K + \pi_k \pi_K \delta \pi \right\}.
\]

The variation of the aggregate capital is
\[
\delta K = \Psi_K \left\{ \sum_{z \in \mathbb{Z}} \int_\pi^x \left\{ \left[ K_\pi(z, x) - \frac{d}{dx} K_{\pi'}(z, x) \right] \delta \pi(x) \right\} dx
\]
\[
+ K_{\pi'}(z, x) |_{\pi(x)} \delta \pi' + \left\{ K_{\pi'}(\pi(x)) |_{\pi(x)} + [K(z, x) - \pi'(x) K_{\pi'}(z, x)] |_{\pi(x)} \pi_k \pi_K \right\} \delta \pi \right\},
\]
where
\[
\Psi_K^{-1} \equiv 1 - \sum_{z \in \mathbb{Z}} \int_\pi^x K_K(z, x) dx - [K(z, x) - \pi'(x) K_{\pi'}(z, x)] |_{\pi(x)} \pi_k \pi_K
\]
\[
- [K(\pi(x)) - \pi'(x) K_{\pi'}(\pi(x))] |_{\pi(x)} \left( \pi_K + \pi_k \pi_K \right).
\]

Substituting the formula for \( \delta K \) into (B.7) we get
\[
\delta J = \sum_{z \in \mathbb{Z}} \int_\pi^x \left[ L_\pi(z, x) + \Psi K_{\pi'}(z, x) - \frac{d}{dx} (L_{\pi'}(z, x) + \Psi K_{\pi'}(z, x)) \right] \delta \pi(x) dx
\]
\[
+ [L_{\pi'}(z, x) + \Psi K_{\pi'}(z, x)] |_{\pi(x)} \delta \pi' + \left\{ [L_{\pi'}(\pi(x)) + \Psi K_{\pi'}(\pi(x))] |_{\pi(x)} \pi_k \pi_K \right\} \delta \pi,
\]
\[
+ L(z, x) + \Psi K(\pi(x)) - \pi'(x) (L_{\pi'}(\pi(x)) + \Psi K_{\pi'}(\pi(x))) |_{\pi(x)} \pi_k \pi_K \right\} \delta \pi,
\]
\[
39
\]
with
\[
\Psi = \frac{\delta L}{\delta K} = \Psi_k \left\{ \sum_{z \in Z} \int_{\pi(z)} L_K(z, x) dx + \left[ L (z, x) - \pi' (x) L_{\pi'} (z, x) \right]_{\pi(z)} \frac{\partial K}{\partial x} \right. \\
+ \left. \left[ L (z, x) - \pi' (x) L_{\pi'} (z, x) \right]_{\pi(z)} \left( x_K + x_k \right) \right\}.
\]

Now, in order to get FOCs we assign \( \delta J = 0 \). Since the first term is zero for any \( \delta \pi (x) \) for all \( x \), the following terms must be zero
\[
\sum_{z \in Z} \left[ L (z, x) + \Psi K (z, x) \right]_{\pi(z)} \left[ \frac{1}{x_k \omega_k} - \pi' (x) \right] \left[ L_{\pi'} (z, x) + \Psi K_{\pi'} (z, x) \right]_{\pi(z)} = 0.
\]

Note that
\[
\frac{\delta x}{\delta \pi}^{-1} = \frac{1}{x_k \omega_k}.
\]

If we denote the modified Lagrange function by \( \tilde{L} (z, x) \equiv L (z, x) + \Psi K (z, x) \), then clearly the derived FOCs are those in (19)-(23). Q.E.D.

**B.1.1 Definition of Terms in Theorem 1**

Recall that according to (13)
\[
W [z, x; K, \pi (x), \pi' (x)] = W \left[ u \left( c \left( z, k (z, x; K); K, \pi, \pi' \right) \right) \right] \lambda \left[ z, k (z, x; K); K, \pi, \pi' \right] k_x (z, x; K),
\]
then
\[
W_\pi (z, x) \delta \pi (x) = \lim_{\varepsilon \to 0} \frac{dW [z, x; K, \tilde{\pi} (x), \pi' (x)]}{d\varepsilon},
\]
with
\[
c (z, k (z, x; K); K, \tilde{\pi}, \pi') = y (z, k (z, x; K); K) - \tilde{\pi} (x) + k (z, x; K) - h (z, k (z, x; K); K, \tilde{\pi}, \pi'),
\]
\[
y (z, k (z, x; K); K) = r (K) k (z, x; K) + w (K) z,
\]
and \( \tilde{\pi} (x) \equiv \pi (x) + \varepsilon \delta \pi (x) \).

To simplify notation we will omit the obvious arguments, writing \( c = c \left( z, k (z, x; K) \right), c_{\pi} = c_{\pi} \left( z, k (z, x; K) \right), \lambda = \lambda \left( z, k (z, x; K) \right) \) and so on. Thus,
\[
W_\pi = \{ W' [u (c)] u' (c) c_{\pi} \lambda + W [u (c)] \lambda_{\pi} \} k_x,
\]
where \( c_{\pi} = -x - h_{\pi} \), and the Fréchet derivatives of the savings policy and the distribution function, \( h_{\pi} \) and \( \lambda_{\pi} \), are given by Lemmas 1 and 2, respectively.
Similarly,
\[ W_{\pi'} = \{ W' [u (c)] u' (c) c_{\pi'} \lambda + W [u (c)] \lambda_{\pi'} \} k_x, \]  
where \( c_{\pi'} = -h_{\pi'} \), for the Fréchet derivatives \( h_{\pi'} \) and \( \lambda_{\pi'} \). Finally, for \( h_K \) and \( \lambda_K \),
\[ W_K = \{ W' [u (c)] u' (c) c_K \lambda + W [u (c)] [\lambda_K + \lambda_k k_K] \} k_x + W [u (c)] \lambda k_x K, \]
where
\[ r_K = F_{11} (K, \bar{L}), \]
\[ w_K = F_{21} (K, \bar{L}), \]
\[ c_K = r_K k + w_K z + [1 + r - h_k] k_K - h_K. \]

For the equation
\[ \frac{d}{dx} W_{\pi'} (z, x) = W_{\pi', \pi'} (z, x) \pi'' (x) + W_{\pi', \pi} (z, x) \pi' (x) + W_{\pi' x} (z, x), \]
we obtain, using \( W = W [u (c)], W' = W' [u (c)] u' (c), \) and \( W'' = W'' [u (c)] [u' (c)]^2 + W' [u (c)] u'' (c), \)
\[ W_{\pi', \pi'} (z, x) = \{ W'' [c_{\pi'}] \lambda + W' [c_{\pi', \pi'} \lambda + 2 c_{\pi'} \lambda_{\pi'}] + W \lambda_{\pi', \pi'} \} k_x, \]
\[ W_{\pi', \pi} (z, x) = \{ W'' c_{\pi'} c_{\pi} \lambda + W' [c_{\pi', \pi} \lambda + c_{\pi'} \lambda_{\pi'} + c_{\pi} \lambda_{\pi'}] + W \lambda_{\pi', \pi} \} k_x, \]
\[ W_{\pi' x} (z, x) = \{ W'' c_{\pi'} c_{\pi} \lambda + W' [c_{\pi', \pi} \lambda + c_{\pi'} \lambda_{\pi'} + c_{\pi} \lambda_{\pi'}] + W \lambda_{\pi' x} \} k_x + \{ W' c_{\pi'} \lambda + W \lambda_{\pi'} \} k_{xx}, \]
with \( c_{\pi', \pi'} = -h_{\pi', \pi'}, c_{\pi'} = -h_{\pi', \pi} x, c_{\pi' x} = -h_{\pi' x} k_x (z, x), \) and \( \lambda_{\pi' x} = \lambda_{\pi' x} k_x (z, x). \)

According to (14),
\[ G [z, x; K, \pi (x), \pi' (x)] = [\pi (x) x - g y (z, k (z), x; K); K)] \lambda [z, k (z, x; K); K, \pi, \pi'] k_x (z, x; K), \]
and therefore,
\[ G_{\pi} (z, x) = \{ x \lambda + [\pi (x) x - g y] \lambda_{\pi} \} k_x, \]
\[ G_{\pi'} (z, x) = \{ [\pi (x) x - g y] \lambda_{\pi'} \} k_x, \]
\[ G_K (z, x) = \{ -g [r_k k + r_k + w_k z] \lambda + [\pi (x) x - g y] [\lambda_K + \lambda_k k_K] \} k_x + [\pi (x) x - g y] \lambda k_x K, \]
and
\[ \frac{d}{dx} G_{\pi'} (z, x) = G_{\pi', \pi'} (z, x) \pi'' (x) + G_{\pi', \pi} (z, x) \pi' (x) + G_{\pi' x} (z, x), \]
\[ G_{\pi', \pi'} (z, x) = \{ [\pi (x) x - g y] \lambda_{\pi', \pi'} \} k_x, \]
\[ G_{\pi', \pi} (z, x) = \{ x \lambda_{\pi} + [\pi (x) x - g y] \lambda_{\pi'} \} k_x, \]
\[ G_{\pi' x} (z, x) = \{ [\pi (x) x - g y k_x] \lambda_{\pi'} + [\pi (x) x - g y] \lambda_{\pi' x} \} k_x + [\pi (x) x - g y] \lambda_{\pi' x} k_{xx}. \]
Lastly, according to (14)
\[ K[z, x; K, \pi(x), \pi'(x)] = k(z, x; K) \lambda [z, k(z, x; K); K, \pi, \pi'] k_x(z, x; K), \]
and thus
\[
\begin{align*}
K_\pi(z, x) &= k\lambda_x k_x, \\
K_{\pi'}(z, x) &= k\lambda_{x'} k_x, \\
K_K(z, x) &= \{k_K + k [\lambda_k k_K + \lambda_K]\} k_x + k \lambda k_{xK},
\end{align*}
\]
and
\[
\frac{d}{dx} K_\pi'(z, x) = K_\pi''(z, x) \pi''(x) + K_\pi'(z, x) \pi'(x) + K_\pi(z, x),
\]
\[
K_{\pi''}(z, x) = k\lambda_{x''} k_x,
\]
\[
K_{\pi'}(z, x) = k\lambda_{x'} k_x,
\]
\[
K_{\pi'}(z, x) = \{k^2 + kk_{xx}\} \lambda_{x'} + k\lambda_{x'} k_{xK}.
\]

For the marginal effect \( \Psi \) of the aggregate capital stock on social welfare, the marginal effects at the lower and upper bounds are defined for \( L \) in equation (22), as
\[
L_K(z, \underline{\pi}(z)) \equiv [(L(z, x) - \pi'(x) L_{\pi'}(z, x)) x_K]_{x=\underline{\pi}(z)},
\]
\[
L_L(z, \overline{\pi}(z)) \equiv [(L(z, x) - \pi'(x) L_{\pi'}(z, x)) (x_K + x_k \omega_K)]_{x=\pi(z)},
\]
and similarly for \( K \) in equation (23),
\[
K_K(z, \underline{\pi}(z)) \equiv [(K(z, x) - \pi'(x) K_{\pi'}(z, x)) x_K]_{x=\underline{\pi}(z)},
\]
\[
K_L(z, \overline{\pi}(z)) \equiv [(K(z, x) - \pi'(x) K_{\pi'}(z, x)) (x_K + x_k \omega_K)]_{x=\pi(z)}.
\]

### B.2 Proof of Theorem 2

First, we need to prove that the second-order sufficient Legendre condition for a maximum in the optimal tax schedule problem has the form \( \sum_{z \in Z} L_{x'''}(z, x) < 0 \) for all \( x \in [\underline{\pi}(z), \overline{\pi}(z)] \). See Luenberger (1969), Ok (2007) or Gelfand and Fomin (2000) as standard references.

Applying the second-order Taylor expansion of \( J \),
\[ J'(\varepsilon) = J(0) + J'(0) \varepsilon + J''(0) \varepsilon^2 + o(\varepsilon^2), \]
we obtain
\[
J''(0) = \frac{1}{2} \sum_{z \in Z} \int_{\underline{\pi}(z)}^{\overline{\pi}(z)} \left\{ L_{\pi'''}(z, x) [\delta \pi'(x)]^2 + 2L_{\pi'''}(z, x) \delta \pi(x) \delta \pi'(x) + L_{\pi'''}(z, x) [\delta \pi'(x)]^2 \right. \]
\[ + 2L_{K\pi}(z, x) \delta K \delta \pi(x) + 2L_{K\pi}(z, x) \delta K \delta \pi'(x) + L_{KK}(z, x) (\delta K)^2 \} \ dx, \quad (B.11) \]
where
\[ L(z, x) \equiv W(z, x) + \mu G(z, x). \]

The term \( L_K \) can be written as
\[
\sum_{z \in \mathcal{Z}} \int_{\mathcal{Z}(z)} L_K(z, x) \delta K \delta \pi(x) \, dx = \delta K \sum_{z \in \mathcal{Z}} \int_{\mathcal{Z}(z)} L_K(z, x) \delta \pi(x) \, dx,
\]
(B.12)

the term \( L_{K'} \) as
\[
\sum_{z \in \mathcal{Z}} \int_{\mathcal{Z}(z)} L_{K'}(z, x) \delta K \delta \pi'(x) \, dx = \delta K \sum_{z \in \mathcal{Z}} \int_{\mathcal{Z}(z)} L_{K'}(z, x) \delta \pi'(x) \, dx
\]
\[
= -\delta K \left\{ \sum_{z \in \mathcal{Z}} \int_{\mathcal{Z}(z)} \frac{d}{dx} L_{K'}(z, x) \delta \pi(x) \, dx \right\},
\]
(B.13)

and similarly, the term \( L_{K''} \) is,
\[
\sum_{z \in \mathcal{Z}} \int_{\mathcal{Z}(z)} L_{K''}(z, x) \delta (K')^2 \, dx = (\delta K)^2 \sum_{z \in \mathcal{Z}} \int_{\mathcal{Z}(z)} L_{K''}(z, x) \, dx.
\]
(B.14)

Adding (B.12) and (B.13), we obtain
\[
\sum_{z \in \mathcal{Z}} \int_{\mathcal{Z}(z)} L_K(z, x) \delta K \delta \pi(x) \, dx + \sum_{z \in \mathcal{Z}} \int_{\mathcal{Z}(z)} L_{K'}(z, x) \delta K \delta \pi'(x) \, dx
\]
\[
= \delta K \left\{ \sum_{z \in \mathcal{Z}} \int_{\mathcal{Z}(z)} \left[ L_K(z, x) - \frac{d}{dx} L_{K'}(z, x) \right] \delta \pi(x) \, dx \right\}.
\]

Finally,
\[
\sum_{z \in \mathcal{Z}} \int_{\mathcal{Z}(z)} 2L_{\pi'}(z, x) \delta \pi(x) \delta \pi'(x) = \sum_{z \in \mathcal{Z}} \int_{\mathcal{Z}(z)} L_{\pi'}(z, x) \frac{d}{dx} [\delta \pi(x)]^2
\]
\[
= -\sum_{z \in \mathcal{Z}} \int_{\mathcal{Z}(z)} \frac{d}{dx} L_{\pi'}(z, x) [\delta \pi(x)]^2.
\]

Substituting the variation of the aggregate capital from equation (23) for \( \delta K \),
\[
\delta K = \Psi_K \left\{ \sum_{z \in \mathcal{Z}} \int_{\mathcal{Z}(z)} \left[ \mathcal{K}_\pi(z, x) - \frac{d}{dx} \mathcal{K}_{\pi'}(z, x) \right] \delta \pi(x) \, dx \right\},
\]
(B.15)
the Taylor expansion (B.11) can be expressed as

\[
J''(0) = \sum_{z \in Z} \int_{\Xi(z)} \left\{ P(z, x) [\delta \pi'(x)]^2 + Q(z, x) [\delta \pi(x)]^2 \right\} \, dx 
+ \left( \sum_{z \in Z} \int_{\Xi(z)} A(z, x) \delta \pi(x) \, dx \right) \left( \sum_{z \in Z} \int_{\Xi(z)} B(z, x) \delta \pi(x) \, dx \right) 
+ \Phi \left( \sum_{z \in Z} \int_{\Xi(z)} A(z, x) \delta \pi(x) \, dx \right)^2,
\]

where

\[
P(z, x) \equiv \frac{1}{2} L_{\pi' \pi'}(z, x),
Q(z, x) \equiv \frac{1}{2} \left[ L_{\pi \pi}(z, x) - \frac{d}{dx} L_{\pi' \pi'}(z, x) \right],
A(z, x) \equiv \Psi_K \left[ K_{\pi}(z, x) - \frac{d}{dx} K_{\pi'}(z, x) \right],
B(z, x) \equiv L_{K\pi}(z, x) - \frac{d}{dx} L_{K\pi'}(z, x),
\Phi \equiv \sum_{z \in Z} \int_{\Xi(z)} L_{KK}(z, x) \, dx.
\]

If \( \pi \) is a maximum then for all perturbations \( \delta \pi \) the quantity \( J''(0) \) must be negative. If we consider a family of perturbations \( \delta \pi \), parameterized by a small \( \varepsilon > 0 \), then

\[
\sum_{z \in Z} \int_{\Xi(z)} P(z, x) [\delta \pi'(x)]^2 \, dx
\]

is not bounded as \( \varepsilon \rightarrow 0 \) due to \( \delta \pi' \) which is of order \( 1/\varepsilon \). On the other hand, all other terms in (B.16) containing \( \delta \pi \) are bounded and, therefore, dominated by

\[
\sum_{z \in Z} \int_{\Xi(z)} P(z, x) [\delta \pi'(x)]^2 \, dx.
\]

This implies that the necessary condition for a maximum is the so-called Legendre condition

\[
\sum_{z \in Z} L_{\pi' \pi'}(z, x) < 0 \quad \text{for all } x \in [\underline{z}(\bar{x}), \pi(\bar{x})].
\]

Next we need to prove that this condition is satisfied for our Calculus of Variations Ramsey Problem given in Definition 5. We show that the second-order sufficient condition is satisfied for any shape of the policy schedule \( \pi \). The definition of \( L = \mathcal{W} + \mu \mathcal{G} \) gives us

\[
L_{\pi' \pi'} = \mathcal{W}_{\pi' \pi'} + \mu \mathcal{G}_{\pi' \pi'}.
\]

The expression for \( \mathcal{W} \) in (17) leads to

\[
\mathcal{W}_{\pi'} = \{ u'(c) c_{\pi'} \lambda + u(c) \lambda_{\pi'} \} k_x,
\]

and

\[
\mathcal{W}_{\pi' \pi'} = \left\{ \left[ u''(c) (c_{\pi'})^2 + u'(c) c_{\pi' \pi'} \right] \lambda + 2u'(c) c_{\pi' \pi'} \lambda_{\pi'} + u(c) \lambda_{\pi' \pi'} \right\} k_x.
\]

Assuming the same specification as the optimal total income tax problem\(^{30} \) with

\(^{30}\)Setting the positive linear aggregator \( W \equiv 1 \) is innocuous for the proof under the generality of the
\[ c(z, x; \pi) = y - \pi(x)x + k(x) - h(z, k(x); \pi) \] and \( k_x = 1/x_k > 0 \) which implies that \( c_{\pi'} = -h_{\pi'} \) and \( c_{\pi'\pi'} = -h_{\pi'\pi'} \). Since \( G = [\pi(x)x - gy] \lambda k_x \), we obtain \( G_{\pi'\pi'} = [\pi(x) - gy] \lambda_{\pi'\pi'} k_x \), and

\[
L_{\pi'\pi'} = \{ [u''(c)(c_{\pi'})^2 + u'(c)c_{\pi'\pi'}] \lambda + 2u'(c)c_{\pi'}\lambda_{\pi'} + [u(c) + \mu(\pi(x)x - gy)] \lambda_{\pi'\pi'} \} k_x.
\]

When we consider the higher order terms \( c_{\pi'\pi'} \) and \( \lambda_{\pi'\pi'} \) negligible, the above equation reduces to the following condition

\[
L_{\pi'\pi'} = \{ u''(c)(c_{\pi'})^2 \lambda + 2u'(c)c_{\pi'}\lambda_{\pi'} \} k_x,
\]

where the first term is clearly negative due to \( u''(c) < 0 \). The sign of the second term follows from the definition of the Fréchet derivative \( h_{\pi'} = \delta h/\delta \pi' \), where the policy function for the next-period capital depends implicitly on the tax schedule function and its derivative in the Euler equation operator \( \mathcal{F} \). We consider what is the effect on the next-period capital function \( h \) when \( \pi' \) shifts marginally. If this marginal shift of \( \pi' \) lowers savings, \( h_{\pi'} < 0 \), the change in the related segment of the stationary distribution has the same sign, \( \lambda_{\pi'} < 0 \). Vice versa, an increase in savings, \( h_{\pi'} > 0 \), implies \( \lambda_{\pi'} > 0 \). Because from (B.8) the effect on consumption \( c_{\pi'} = -h_{\pi'} \), the second term in (B.17), \( c_{\pi'}\lambda_{\pi'} \), is always negative. Therefore, \( L_{\pi'\pi'}(z, k; \pi) < 0 \) for each \( z \in Z \) and \( k \in [\bar{k}(z), \bar{\pi}(\bar{z})] \), thus \( \sum_{z \in Z} L_{\pi'\pi'}(z, k) < 0 \) for each \( k \in [\bar{k}(z), \bar{\pi}(\bar{z})] \).

Second, we need to prove that there are no conjugate points on the interval \([\bar{x}(\bar{z}), \bar{x}(\bar{z})]\). Following Liberzon (2011) and considering only the second-order terms, equation (B.16) can be written as

\[
\int_a^b \left\{ R(x) [\delta \pi(x)]^2 + S(x) [\delta \pi'(x)]^2 \right\} dx,
\]

where

\[
R(x) = \sum_{z \in Z} Q(z, x) = \frac{1}{2} \sum_{z \in Z} \left[ L_{\pi\pi'}(z, x) - \frac{d}{dx} L_{\pi\pi'}(z, x) \right],
\]

\[
S(x) = \sum_{z \in Z} P(z, x) = \frac{1}{2} \sum_{z \in Z} L_{\pi'\pi'}(z, x),
\]

with \( a \equiv \bar{x}(\bar{z}) \) and \( b \equiv \bar{x}(\bar{z}) \).

For every differentiable function \( w = w(x) \) we know that

\[
0 = \left[ w\delta \pi^2 \right]_a^b = \int_a^b \frac{d}{dx} (w\delta \pi^2) dx = \int_a^b \left\{ w'\delta \pi^2 + 2w\delta \pi \delta \pi' \right\} dx.
\]

\[
\text{conditions.}
\]

\[31\] For this definition we use the Gateaux derivative which is the same as the Fréchet derivative if it is continuous.
As \( \delta \pi (a) = \delta \pi (b) = 0 \),

\[
\int_a^b \left\{ R(x) [\delta \pi (x)]^2 + S(x) [\delta \pi' (x)]^2 \right\} \, dx = \int_a^b \left\{ [R(x) + w' (x)] [\delta \pi (x)]^2 + 2w (x) \delta \pi (x) \delta \pi' (x) + S(x) [\delta \pi' (x)]^2 \right\} \, dx.
\]

Finding a \( w \) to make the integrand on the right-hand side into a perfect square means that

\[
S (R + w') = w^2,
\] (B.18)

which is a quadratic differential equation for the unknown function \( w \)

\[
\int_a^b S(x) \left[ \frac{w(x)}{S(x)} \delta \pi (x) + \delta \pi' (x) \right]^2 \, dx
\]

with \( S(x) < 0 \). Excluding the conjugate points is equivalent to ensure the existence of a solution to equation (B.18) on the whole interval \([a, b] \). The first step is to transform the quadratic differential equation (B.18) into a linear second-order differential equation using a substitution

\[
w(x) = -\frac{S(x) v' (x)}{v(x)},
\] (B.19)

where \( v \) is a new unknown twice differentiable function. In this way, equation (B.18) transforms into

\[
S \left( R - \frac{d}{dx} \frac{(Sv') v - S (v')^2}{v^2} \right) = \frac{S^2 (v')^2}{v^2}.
\]

Multiplying both sides by \( v \) (nonzero) and dividing by \( S \) (negative) yields a Jacobi equation

\[
Rv = \frac{d}{dx} (Sv').
\] (B.20)

Now we need to prove that any solution \( v \) to the Jacobi equation does not vanish anywhere on \([a, b] \) since then the desired solution \( w \) to the original Ricatti equation (B.18) is given by equation (B.19). Under the condition that \( R \) is close to zero in comparison to \( S \) (results will be qualitatively the same if \( R \) is bounded and small relative to \( S \)), equation (B.20) simplifies to \( d/dx (Sv') = 0 \), implying that \( Sv' = A \) where \( A \) is a constant, and

\[
v(x) = \int_a^x \frac{A}{S(t)} \, dt.
\]

As function \( v \) is arbitrary, we can assume initial conditions \( v(a) = 0 \) and \( v'(a) = 1 \). Using the condition \( v'(a) = 1 \), we have \( S(a) = A \) and

\[
v(x) = \int_a^x \frac{S(a)}{S(t)} \, dt.
\]

Now we need to analyze \( S(t) \) for \( t \in [a, b] \) and show that it is bounded from below and
above, i.e.

\[ 0 > -S \geq -S(t) \geq -\overline{S} > -\infty. \]  \hspace{2cm} (B.21)

From (B.17)

\[
S(x) = \frac{1}{2} \sum_{z \in \mathbb{Z}} \left\{ u''(c(z, k(x))) (c_{x'}(z, k(x)))^2 \lambda(z, k(x)) + 2u'(c(z, k(x))) c_{x'}(z, k(x)) \lambda_{x'}(z, k(x)) \right\}. \quad (B.22)
\]

Consider first the terms with consumption. Using the Regularity Condition, the lowest value of consumption is when a household is borrowing constrained at \( k = 0 \) with total income \( \zeta = (1 - \varepsilon) w_\zeta \) for any \( 0 < \varepsilon < 1 \). Clearly, consumption cannot be larger than the maximum income equal to \( \overline{\zeta} = w \tau + r \overline{k} \). Thus, \( 0 < \zeta \leq c \leq \overline{\zeta} < \infty \) and \( -\infty < u''(\zeta) \leq u''(c) \leq u''(\overline{\zeta}) < 0 \). \( \lambda \) is equal to zero at the endogenous upper bound, \( \overline{k} \). For the same reason \( \lambda_{x'} \) is different from zero. For \( z = \overline{\zeta} \) and \( k \in [0, \overline{k}(\overline{\zeta})] \), \( k^+ = 0 \) and \( h_{x'}(\overline{\zeta}, k) = -c_{x'}(\overline{\zeta}, k) = 0 \) but for \( z = \overline{\zeta} \), \( c_{x'}(\overline{\zeta}, k) \) is nonzero for any \( k \). Thus, \( S \) given by (B.22) is bounded as in (B.21) and

\[
v(x) = \int_a^x \frac{S(a)}{S(t)} dt \geq \int_a^x \frac{S(a)}{\overline{S}} dt = \frac{S(a)}{\overline{S}} (x-a) > 0 \quad \text{for all } x > a.
\]

Therefore, there are no conjugate points to \( a \) on the interval \([a, b]\) and the first-order conditions for the optimal income tax schedule given by the Euler-Lagrange condition are also the sufficiency conditions and the obtained solution is a maximum. Q.E.D.

### B.3 Proof of Lemma 1

From equation (13) the Euler equation operator with the variation \( \tilde{\pi} \equiv \pi + \varepsilon \delta \pi \) is

\[
\mathcal{F}(h, \pi)(z, k; K, \tilde{\pi}(x), \pi'(x)) \equiv u'(c(z, k; K, \tilde{\pi}, \pi')) - \beta \sum_{z^+} u'(c^+(z^+, h(z, k; K, \tilde{\pi}, \pi')) K(z^+, h(z, k; K, \tilde{\pi}, \pi')) Q(z, z^+),
\]

and

\[
\mathcal{F}_\pi(z, k) \delta \pi = \lim_{\varepsilon \to 0} \frac{d\mathcal{F}[z, k; K, \tilde{\pi}(x), \pi'(x)]}{d\varepsilon},
\]

Using abbreviated notation \( x = x(z, k), c = c(z, k), h = h(z, k), h^+ = h(z^+, h), x^+ = x(z^+, h), y^+ = y(z^+, h), \pi^+ = \pi(z^+, h), c^+ = c^+(z, z^+, k), \) and \( R^+ = R^+(z, z^+, k) \),

\[
\mathcal{F}_\pi(z, k) = u''(c) c_{x'} - \beta \sum_{z^+} \left\{ u''(c^+) c_{x'}^+ R^+ + u'(c^+) c_{x'}^+ R_{x'}^+ \right\} Q(z, z^+) = 0,
\]

47
where
\[ c = y - \tilde{\pi}(x)x + k - h, \]
\[ y = rK + wz, \]
\[ c^+ = y^+ - \tilde{\pi}(x^+)_x + h - h^+, \]
\[ R^+ = 1 + y^k - [\tilde{\pi}(x^+) + \pi'(x^+)x^+]x^+_k. \]

Terms for \( F_\pi \) are above and in equation (24) are
\[ c_\pi = -x - h_\pi, \]
\[ c_\pi^+ = [1 + r - h_k^+ - [\pi'x + \pi^+]x^+_k]h_\pi - h^+_\pi - x^+, \]
\[ R^+_\pi = rh_{k\pi} - [1 + 2\pi'(x^+) + \pi''(x^+)x^+]x^+_k h_\pi - [\pi'^x + \pi^+]x^+_k h_\pi. \]

For
\[ F_\pi'(z, k) = u''(c) c_\pi' - \beta \sum_{z^+} \{ u''(c^+) c_\pi^+ R^+ + u'(c^+) R^+_\pi \} Q(z, z^+) = 0, \]
we use terms
\[ c_\pi' = -h_\pi', \]
\[ c_\pi'^+ = \{ 1 + r - h_k^+ - [\pi'x + \pi^+]x^+_k \}h_\pi' - h^+_\pi', \]
\[ R^+_\pi' = rh_{k\pi'} - \{ x^+ + [2\pi'(x^+) + \pi''(x^+)x^+]x^+_k h_\pi' \}x^+_k - [\pi'^x + \pi^+]x^+_k h_\pi'. \]

For
\[ F_K(z, k) = u''(c) c_K - \beta \sum_{z^+} \{ u''(c^+) c_K^+ R^+ + u'(c^+) R^+_K \} Q(z, z^+) = 0, \]
the terms are, as well as for equation (24) in the case of \( F_K \),
\[ c_K = rKk + wKz - [\pi'(x)x + \pi(x)]x_K - h_K, \]
\[ c_K^+ = rKk + wKz^+ - [\pi'^x + \pi^+]x^+_K h_K - h^+_K, \]
\[ R^+_K = rKk + rh_{KK} - \{ 2\pi'(x^+) + \pi''(x^+)x^+ \}x^+_K [x^+_K + x^+_K h_K] - [\pi'^x + \pi^+]x^+_K h_K. \]

For
\[ F_\pi'(z, k) = u''(c) c_\pi'^2 + u''(c) c_\pi', \]
\[ -\beta \sum_{z^+} \{ u''(c^+) [c_\pi'^+]^2 + u''(c^+) c_\pi'^+ \} R^+ + 2u''(c^+) c_\pi^+ R^+_\pi + u'(c^+) R^+_\pi' \} Q(z, z^+) = 0, \]
we use, as well in equation (24) in the case of $\mathcal{F}_{\pi',\pi}$, the terms

$$
c_{\pi',\pi} = -h_{\pi',\pi},
$$

$$
c_{\pi',\pi}^+ = \left[ 1 + \frac{1}{h_{k\pi'} + h_{kk}h_{\pi'}} \right] h_{\pi'} + \left(1 + r - h_k^+\right) h_{\pi';\pi'} - \left( h_{k\pi}^+ h_{\pi'} + h_{\pi';\pi'}^+ \right)
$$

$$
- \left[ \pi' \left( x^+ \right) x^+ + \pi \left( x^+ \right) \right] \left( x_{kk}^+ h_{\pi'}^2 + x_{k\pi'}^+ h_{\pi'} \right),
$$

$$
R_{\pi',\pi}^+ = rh_{k\pi';\pi} - \left[ 3 \left( x_k^+ \right)^2 + x_{kk}^+ \right] h_{\pi'} - \left[ 3 \pi'' \left( x^+ \right) x^+ \right] \left( x_k^+ \right)^3 \left( h_{\pi'}^2 \right)
$$

$$
- \left[ \pi' \left( x^+ \right) x^+ + \pi \left( x^+ \right) \right] \left[ x_{kkk}^+ \left( h_{\pi'}^2 \right)^2 + x_{kk}^+ h_{\pi';\pi'} \right].
$$

Finally,

$$
\mathcal{F}_{\pi',\pi} (z, k) = u'' (c) c_{\pi,\pi} c_{\pi'} + u'' (c) c_{\pi',\pi} - \beta \sum_{z^+} \left\{ u'' \left( c^+ \right) c_{\pi,\pi}^+ c_{\pi'}^+ + u'' \left( c^+ \right) c_{\pi',\pi}^+ \right\} R^+
$$

$$
+ u'' \left( c^+ \right) \left[ c_{\pi,\pi}^+ R_{\pi}^+ + c_{\pi'}^+ R_{\pi'}^+ \right] + u' \left( c^+ \right) R_{\pi,\pi}^+ Q(z, z^+) = 0.
$$

Terms for $\mathcal{F}_{\pi',\pi}$ above and in equation (24) are

$$
c_{\pi',\pi} = -h_{\pi',\pi},
$$

$$
c_{\pi',\pi}^+ = \left[ 1 + \pi'' \left( x^+ \right) x^+ + 2 \pi' \left( x^+ \right) \right] \left( x_{kk}^+ h_{\pi'}^2 + x_{k\pi'}^+ h_{\pi'} \right),
$$

$$
R_{\pi',\pi}^+ = rh_{k\pi';\pi} - \left[ \left( x_k^+ \right)^2 + x_{kk}^+ \right] h_{\pi'} - \left[ 3 \pi'' \left( x^+ \right) x^+ \right] \left( x_k^+ \right)^3 h_{\pi';\pi'}
$$

$$
- \left[ \pi' \left( x^+ \right) x^+ + \pi \left( x^+ \right) \right] \left[ x_{kkk}^+ \left( h_{\pi'}^2 \right)^2 + x_{kk}^+ h_{\pi';\pi'} \right].
$$

Q.E.D.

B.4 Proof of Lemma 2

Using equation (13), the stationary distribution operator is

$$
\mathcal{L}(h, \lambda, \pi) \equiv \lambda (z^+, k^+; K, \pi, \pi') - \sum_z \lambda \left[ z, h^{-1} (z, k^+; K, \pi, \pi'); K, \pi, \pi' \right] Q(z, z^+),
$$

and so

$$
\mathcal{L}_\pi \left( z^+, k^+ \right) \delta \pi = \lim_{\varepsilon \to 0} \frac{d\mathcal{L} \left[ z^+, k^+; K, \pi (x), \pi'(x) \right]}{d\varepsilon},
$$

with $\pi \equiv \pi + \varepsilon \delta \pi$. Therefore, abbreviating for $(K, \pi, \pi')$,

$$
\mathcal{L}_\pi \left( z^+, k^+ \right) = \lambda_\pi \left( z^+, k^+ \right) - \sum_z \left[ \lambda_\pi \left( z, h^{-1} (z, k^+) \right) + \lambda_k \left( z, h^{-1} (z, k^+) \right) h_{k}^{-1} (z, k^+) \right] = 0
$$

49
\[ h^{-1}_\pi(z, k^+) = h_{\pi}(z, h^{-1}(z, k^+)) / h_k(z, h^{-1}(z, k^+)) \]

Similarly, we can derive the total F-derivative of the Euler equation operator with respect to the derivative of government policy function,

\[ L_{\pi'}(z^+, k^+) = \lambda_{\pi'}(z^+, k^+) - \sum_z \left[ \lambda_{\pi'}(z, h^{-1}(z, k^+)) + \lambda_k(z, h^{-1}(z, k^+)) h^{-1}_\pi(z, k^+) \right] = 0 \]

with \( h^{-1}_{\pi'}(z, k^+) = h_{\pi'}(z, h^{-1}(z, k^+)) / h_k(z, h^{-1}(z, k^+)) \), and

\[ L_K(z^+, k^+) = \lambda_K(z^+, k^+) - \sum_z \left[ \lambda_K(z, h^{-1}(z, k^+)) + \lambda_k(z, h^{-1}(z, k^+)) h^{-1}_K(z, k^+) \right] = 0, \]

with \( h^{-1}_K(z, k^+) = h_K(z, h^{-1}(z, k^+)) / h_k(z, h^{-1}(z, k^+)) \).

Further, we can derive the following total F-derivative of the Euler equation operator

\[ L_{\pi'\pi'}(z^+, k^+) = \lambda_{\pi'\pi'}(z^+, k^+) - \sum_z \left\{ \lambda_{\pi'\pi'}(z, h^{-1}(z, k^+)) + \lambda_{\pi'k}(z, h^{-1}(z, k^+)) h^{-1}_{\pi'}(z, k^+) \right\} \]

\[ + \left[ \lambda_{\pi'k}(z, h^{-1}(z, k^+)) + \lambda_{kk}(z, h^{-1}(z, k^+)) \right] \left[ h^{-1}_\pi(z, k^+) \right]^2 \]

\[ + \lambda_k(z, h^{-1}(z, k^+)) h^{-1}_{\pi'}(z, k^+) \right\} = 0 \]

where \( \left. h^{-1}_{\pi'\pi'}(z, k^+) = h_{\pi'\pi'}(z, h^{-1}(z, k^+)) / h_{kk}(z, h^{-1}(z, k^+)) \right\} \).

Finally,

\[ L_{\pi'\pi}(z^+, k^+) = \lambda_{\pi'\pi}(z^+, k^+) - \sum_z \left\{ \lambda_{\pi'\pi}(z, h^{-1}(z, k^+)) + \lambda_{\pi'k}(z, h^{-1}(z, k^+)) h^{-1}_{\pi}(z, k^+) \right\} \]

\[ + \left[ \lambda_{\pi'k}(z, h^{-1}(z, k^+)) + \lambda_{kk}(z, h^{-1}(z, k^+)) \right] \left[ h^{-1}_\pi(z, k^+) \right]^2 \]

\[ + \lambda_k(z, h^{-1}(z, k^+)) h^{-1}_{\pi\pi}(z, k^+) \right\} = 0 \]

where \( h^{-1}_{\pi'\pi}(z, k^+) = h_{\pi'\pi}(z, h^{-1}(z, k^+)) h_{kk}(z, h^{-1}(z, k^+)) \).

Q.E.D.

### B.5 Proof of Proposition 1

Using the definition for \( \tilde{L} \) in equation (18), the boundary first-order condition in (21) can be expressed as

\[ \tilde{L}_{\pi'}(z, \tilde{x}(z)) = W_{\pi'}(z, \tilde{x}(z)) + \mu G_{\pi'}(z, \tilde{x}(z)) + \Psi K_{\pi'}(z, \tilde{x}(z)) = 0. \]

The term \( W_{\pi'} \) in (B.8) evaluated at \( (z, \tilde{x}(z)) \) gives

\[ W_{\pi'}(z, \tilde{x}(z)) = \left\{ W'[u(c(z, \tilde{k}))] u'(c(z, \tilde{k})) c_{\pi'}(z, \tilde{k}) \lambda(z, \tilde{k}) \right\} + W[u(c(z, \tilde{k}))] \lambda_{\pi'}(z, \tilde{k}) k_{\pi}(z, \tilde{x}(z)). \]

(B.23)

From \( c_{\pi'} = -h_{\pi'} \) it follows that

\[ c_{\pi'}(z, \tilde{k}) = -h_{\pi'}(z, \tilde{k}) = 0, \]
since the ‘variation’ of the savings policy function related to the lowest shock, \( h(\bar{z}, \cdot) \), with respect to the slope of the government policy function, \( h_{\sigma'} \), at \((\bar{z}, \bar{k})\) is clearly zero. Note that the saving function of the borrowing-constrained agents is flat and equal to zero. It implies that the first term in (B.23) is also equal to zero. Using \( \mathcal{G}_{\sigma'} \) and \( \mathcal{K}_{\sigma'} \) from (B.9) and (B.10), respectively, we get

\[
W[u(c(\bar{z}, \bar{k}))] + \mu [\pi(x(\bar{z}))x(\bar{z}) - gy(\bar{z}, \bar{k})] + \Psi \bar{k} = 0,
\]

where we used the fact that both \( \lambda_{\sigma'}(\bar{z}, \bar{k}) \) and \( k_x(\bar{z}, x(\bar{z})) \) are non-zero. The result of the Proposition follows. Q.E.D.

**B.6 Proof of Proposition 2**

We see that terms \( \bar{L} \) and \( \bar{L}_{\sigma'} \) appear in the boundary first-order condition in (20). Using the definition of \( \bar{L} \) in (18), we get

\[
\bar{L}(\bar{z}, \bar{x}(\bar{z})) = \{ W[u(c(\bar{z}, \bar{k}))] + \mu [\pi(x(\bar{z}))x(\bar{z}) - gy(\bar{z}, \bar{k})] + \Psi \bar{k} \} \lambda(\bar{z}, \bar{k}) = 0,
\]

where we have used (17), (14), and (14) for \( \mathcal{W}, \mathcal{G}, \) and \( \mathcal{K} \), respectively, and also the fact that at the upper endogenous limit on capital, \( \bar{k}, \lambda(\bar{z}, \bar{k}) = 0 \). Thus the first-order boundary condition (20) becomes

\[
\left( \pi'(x(\bar{z})) - \frac{k_x(\bar{z}, x(\bar{z}))}{\omega_{x}(\bar{z}, x(\bar{z}))} \right) \bar{L}_{\sigma'}(\bar{z}, \bar{x}(\bar{z})) = 0.
\]

If we assume that \( \pi'(x(\bar{z})) - \frac{k_x(\bar{z}, x(\bar{z}))}{\omega_{x}(\bar{z}, x(\bar{z}))} > 0 \), then the condition above is satisfied if the second term \( \bar{L}_{\sigma'}(\bar{z}, \bar{x}(\bar{z})) \) is equal to zero. By the inspection of \( \mathcal{W}_{\sigma'} \) in (B.8) we see that since \( \lambda(\bar{z}, \bar{k}) = 0 \) then again the first term in (B.8) is zero and

\[
W[u(c(\bar{z}, \bar{k}))] + \mu [\pi(x(\bar{z}))x(\bar{z}) - gy(\bar{z}, \bar{k})] + \Psi \bar{k} = 0,
\]

where we again used the fact that both \( \lambda_{\sigma'}(\bar{z}, \bar{k}) \) and \( k_x(\bar{z}, x(\bar{z})) \) are non-zero. This implies the result of the Proposition. Q.E.D.

**B.7 Proof of Theorem 3**

We need to show that the first order approach to each agent’s maximization problem is valid. First, agents maximize over a quasi-convex set: \( \Psi = \{ x \in B : 0 \leq x \leq \psi(k, z) \text{ for all } (k, z) \in B \times Z \} \). If the function \( \psi \) is increasing and quasi-concave, then the set \( \Psi \) is quasi-convex. Further, we need to satisfy Assumptions 18.1 in Stokey, Lucas, and Prescott (1989), particularly that (i) \( \beta \in (0, 1) \); (ii) utility function \( u : \mathbb{R}_+ \rightarrow \mathbb{R} \) is twice continuously differentiable, strictly increasing and strictly concave function; (iii) for some \( \bar{k}(z) > 0, \psi(k, z) - k \) is strictly positive on \([0, \bar{k}(z)]\) and strictly negative for \( k > \bar{k}(z) \), where the value \( \bar{k} \), the maximum sustainable capital stock out of after-tax income for any agent, is defined as \( \bar{k} = \max\{\bar{k}(z_1), \ldots, \bar{k}(z_J)\} \); and, (iv) given the tax-schedule function
the right-hand side of the Euler equation is strictly positive

$$\beta \sum_{z^+} u'(\psi(k^+(k, z), z^+) - k^+(k^+(k, z), z^+)) \psi_1(k^+(k, z), z^+) Q(z, z^+) > 0,$$

where

$$\psi_1(k^+(k, z), z^+) Q(z, z^+) = (1 - \tau(y(k^+(k, z), z^+))) - \tau'(y(k^+(k, z), z^+)) y(k^+(k, z), z^+)) r^+] Q(z, z^+),$$

It can be easily checked that the assumptions (i)-(iii) are satisfied from our previous assumptions and the model. The assumption (iv) follows directly from the fact that \( \psi \) is increasing in \( k \), i.e. \( \psi_1 > 0 \).

The other assumptions needed for proving the existence of a stationary recursive competitive equilibrium (see Assumption 18.2 in Stokey, Lucas, and Prescott (1989)) are satisfied: (i) the equilibrium marginal return on capital for any \( k \in B \) is finite (in our case the interest rate \( r \)); and (ii) that \( \lim_{c \to 0} u'(c) = \infty \).

Then to prove the Schauder’s Theorem, let \( C(B, Z) \) be the set of continuous bounded functions \( h : B \times Z \to B \) and define a subset \( F = \{ h \in C(B, Z) \} \) where the function \( h \) satisfies \( 0 \leq h(k, z) \leq \psi(k, z) \), all \( (k, z) \in B \times Z \), and \( h \) and \( \psi - h \) are nondecreasing. Note that \( B \times Z \) is a bounded subset of \( \mathbb{R}^2 \) and that the family of functions \( F \) is nonempty, closed, bounded, and convex. Define an operator \( T \) on \( F \)

$$u'(\psi(k, z) - (Th)(k, z)) = \beta \sum_{z^+} u'(\psi((Th)(k, z), z^+) - h((Th)(k, z), z^+))$$

$$\cdot [(1 - \tau(y(\cdot))) - \tau'(y(\cdot)) y(\cdot) r^+] Q(z, z^+),$$

where \( y(\cdot) = y((Th)(k, z), z^+) \).

Then it is easy to prove that \( T \) is well defined, continuous and that \( T : F \to F \). From the conditions on function \( h \) and finite return on capital, it follows that \( F \) is an equicontinuous family. That the operator \( T \) has a fixed point in \( F \) follows from the Schauder’s Theorem (see e.g. Theorem 17.4 in Stokey, Lucas, and Prescott (1989)).

The existence of the stationary recursive competitive equilibrium is standard from the monotonicity, Feller and mixing property of \( Q \) and the non-decreasing policy functions (see Chapter 12 in Stokey, Lucas, and Prescott (1989)). Q.E.D.

C Appendix: The Least Squares Projection Method

The optimal income tax policy, \( \pi \), is a solution of the following system of operator equations:

1. FOC for \( \pi \) given by the Euler-Lagrange condition in (19);
2. the Euler equation (10) capturing the individual optimal behavior \( h \);
3. five operator equations (24)-(25) for F-derivatives of \( h \) based on the Euler equation, \( h_\pi, h_{\pi'}, h_K, h_{\pi''}, \) and \( h_{\pi'''} \);
4. the operator equation for distribution function, \( \lambda \), in (11); and
5. five operator equations (26)-(27) for F-derivatives of \( \lambda \) based on the operator equation for distribution function: \( \lambda_\pi, \lambda_{\pi'}, \lambda_K, \lambda_{\pi''}, \) and \( \lambda_{\pi'''} \).
In order to solve the problem numerically, we first approximate all the unknown functions by combinations of polynomials from a polynomial base. Approximated solutions are specified by unknown parameters transforming the original infinitely dimensional problem into a finite dimensional one. After substituting the approximated functions into the original operator equations we construct the residual equations. Ideally, the residual functions should be uniformly equal to zero. In practical situations, however, this is not achievable and we limit the problem to a finite number of conditions, the so called projections, whose satisfaction guarantees a reasonably good approximation. There are many possibilities how to define the projections. We have chosen the least squares projection method for its good convergence properties and advantage in solving systems of nonlinear operator equations. We search for parameters approximating the functional equations that minimize the squared residual functions.

As we specified above, in the system of operator equations given by (10), (11), (19), (24)-(25), and (26)-(27), there are thirteen unknown classes of functions \( \{\pi, h, h_\pi, h_K, h_{\pi'\pi}, \lambda, \lambda_\pi, \lambda_K, \lambda_{\pi'\pi}, \lambda_{\pi'\pi}'\} \). Since we assume that the shocks are discrete, \( z \in Z = \{z_1, z_2, \ldots, z_J\} \) and \( J > 1 \), we define the following family of policy and distribution functions, and their derivatives \( \{h^i(k), \lambda^i(k), h^i_\pi(k), h^i_{\pi'}(k), h^i_K(k), h^i_{\pi'\pi}(k), \lambda^i_\pi(k), \lambda^i_{\pi'}(k), \lambda^i_K(k), \lambda^i_{\pi'\pi}(k), \lambda^i_{\pi'\pi}'(k)\}_{i=1}^J \), for each shock value \( z_1, z_2, \ldots, z_J \). We interpret the policy function \( h^i \) as the next-period capital function of an agent who was hit by a shock level \( z_i \). Analogously, the distribution function \( \lambda^i \) is the distribution of agents with the shock \( z_i \), etc. Similarly, we assign the Euler and distribution function operators to every shock level, \( \mathcal{F}^i \) and \( \mathcal{L}^i \), respectively. We approximate all unknown functions by the orthogonal Chebyshev polynomial base \( \{T_i(x)\}_{i=0}^\infty \) defined for \( x \in [-1, 1] \).

As we have to define our approximation on a finite interval, we set the highest capital level to a value \( \widehat{k} \), greater than the endogenous upper bound on the stationary distribution. Let the interval of approximation be \([k, \widehat{k}]\) and the degrees of approximation for \( \{h^i(k), h^i_\pi(k), h^i_{\pi'}(k), h^i_K(k), h^i_{\pi'\pi}(k), \lambda^i(k), \lambda^i_\pi(k), \lambda^i_{\pi'}(k), \lambda^i_K(k), \lambda^i_{\pi'\pi}(k), \lambda^i_{\pi'\pi}'(k)\} \) be \( M, M_\pi, M_{\pi'}, M_{\pi'\pi}, M_{\pi'\pi'}, N, N_\pi, N_{\pi'}, N_K, N_{\pi'\pi}, N_{\pi'\pi'}, P \geq 2 \), respectively. Thus, we obtain

\[
\hat{h}^i_m(k; a^i_m) = \sum_{j=1}^{M_m} a^i_{m,j} \phi_j(k), \\
\hat{\lambda}^i_m(k; b^i_m) = \sum_{j=1}^{N_m} b^i_{m,j} \phi_j(k),
\]

\(32\)For an excellent survey and description of these methods see Chapter 11 in Judd (1998).

\(33\)The details on Chebyshev polynomials can be found in Judd (1992), Judd (1998) or in any book on numerical mathematics. The linear transformation \( \xi : [k, \widehat{k}] \to [-1, 1] \) is necessary if we want to use the Chebyshev polynomials on the proper domain. It is straightforward to show that \( \xi(k) = 2(k - k)/(\widehat{k} - k) - 1 \).
with \( i \in \{1, 2, \ldots, J \} \) and \( m \in \{0, \pi, \pi', K, \pi' \pi', \pi' \pi' \} \), and

\[
\hat{\pi}(x; c) \equiv \sum_{j=1}^{P} c_j \phi_j(x),
\]

for any \( k \in [k, \bar{k}] \) and \( x \in [r_k + w_Z, r_{\bar{k}} + w_Z] \), where \( \phi_j(k) \equiv T_{j-1}(\xi(k)) \), and \( a' \), \( b' \), and \( c' \)’s are the unknown parameters.

Now we have to define residual functions as approximations to the original operator functions (10), (11), (19), (24)-(25), and (26)-(27). Substituting the above approximations for the unknown functions,

\[
R^L(x; p) = \sum_j \left[ \bar{L}_n \left( \bar{h}, \hat{\Lambda}, \hat{\pi} \right) + \frac{d}{dx} \bar{L}_n \left( \bar{h}, \hat{\Lambda}, \hat{\pi} \right) \right],
\]

(C.1)

\[
R^F_m(k; p) = F^i_m(\hat{h}, \hat{\pi}),
\]

(C.2)

\[
R^L_m(k; p) = L^i_m(\hat{h}, \hat{\Lambda}, \hat{\pi}),
\]

(C.3)

with \( i = 1, \ldots, J \) and \( m \in \{0, \pi, \pi', K, \pi' \pi', \pi' \pi' \} \) where

\[
p \equiv (a, a_\pi, a_{\pi'}, a_K, a_{\pi' \pi'}, a_{\pi' \pi'}, b, b_\pi, b_{\pi'}, b_K, b_{\pi' \pi'}, b_{\pi' \pi'}, c),
\]

\[
a_m \equiv (a^1_m, a^2_m, \ldots, a^J_m),
\]

\[
b_m \equiv (b^1_m, b^2_m, \ldots, b^J_m),
\]

and \( p \) is of a size \( S = J \times (\sum_m (M_m + N_m)) + P \),

\[
\hat{h} = (\hat{h}, \hat{h}_\pi, \hat{h}_{\pi'}, \hat{h}_K, \hat{h}_{\pi' \pi'}, \hat{h}_{\pi' \pi'}),
\]

\[
\hat{h}_m = (\hat{h}_m, \ldots, \hat{h}_m),
\]

\[
\hat{\Lambda} = (\hat{\Lambda}, \hat{\Lambda}_\pi, \hat{\Lambda}_{\pi'}, \hat{\Lambda}_K, \hat{\Lambda}_{\pi' \pi'}, \hat{\Lambda}_{\pi' \pi'}),
\]

\[
\hat{\lambda}_m = (\hat{\lambda}_m, \ldots, \hat{\lambda}_m),
\]

for any \( i = 1, \ldots, J \).

The least squares projection method searches for a vector of parameters \( p \) that minimizes the sum of weighted residuals,

\[
\sum_{i=1}^{J} \int_{k}^{\bar{k}} \left\{ \sum_m \left[ (R^F_m(k; p))^2 + (R^L_m(k; p))^2 \right] \right\} w(k)dk + \int_{k}^{\bar{k}} [R^L(x(k); p)]^2 w(k)dk,
\]

with the weighting function given by \( w(k) \equiv \left( 1 - \left( \frac{2k-k}{\bar{k}-k} \right)^2 \right)^{-1/2} \) and \( i = 1, \ldots, J \). After approximating the integrals by the Gauss-Chebyshev quadrature, we obtain a minimization
problem
\[ \min_{p \in \mathbb{R}^d} \sum_k \left\{ \sum_{i=1}^J \sum_m \left( \left[ R_{m}^{F_i}(k; p) \right]^2 + \left[ R_{m}^{C_i}(k; p) \right]^2 \right) + \left[ R_k^{x}(x(k); p) \right]^2 \right\}, \tag{C.4} \]

with \( \hat{k} \)'s being the zeros of the polynomial \( \phi \) of a degree greater than the biggest degree of the polynomial approximations, i.e. \( \max\{M, M_{\pi'}, M_{\pi''}, M_{\pi'''}, N, N_{\pi}, N_{\pi'}, N_{\pi''}, N_{\pi'''}, N_{\pi''''}, P\} \).

Since the least squares projection method sets up an optimization problem we can use standard methods of numerical optimization, e.g. the Gauss-Newton or the Levenberg-Marquardt methods. Again, the discussion of these methods is not the aim of our paper. However, we found that these traditional methods did not work in our high-dimensional problem mainly due to possible multiple local solutions. We tried several other methods (simulated annealing or genetic algorithm with quantization, for example) and finally succeeded with a genetic algorithm with multiple populations and local search. The used degrees of polynomial approximation for the optimal individual policy functions \( h \), distribution functions, \( \lambda \), the related sensitivity functions \( h_{\pi}, \lambda_{\pi} \), and the optimal government policy function, \( \pi \), where 4, 12, 3, 3, and 4, respectively. The residuals of the related functional equations were of the order \( 10^{-3} \) or \( 10^{-4} \) with the exception of \( h_{\pi} \) which was of the order \( 10^{-2} \).

Figure C.1 shows the sensitivity functions \( h_{\pi} \) and \( \lambda_{\pi} \). The top panel shows the effect of a change in the optimal tax schedule on the savings decision of agents. For the low shock it is close to zero, for the high shock it is negative and monotonically decreasing. The bottom panel displays the same effects on the probability density function of the stationary distribution \( \lambda \), again for each shock. We know from the stationarity condition of the distribution that the integral of these functions must be zero.\(^{34}\)

D  Example: Optimal Tax on Total Income

D.1  Definitions of Terms in Theorem 1

The terms from Theorem 1 for Example are those in Appendix B.1.1 together with

\[
\begin{align*}
    y(z, k(z, x; K)) & = x, \\
    k_x(z, x; K) & = \frac{1}{r(K)}, \\
    k_K(z, x; K) & = -\frac{w_K(K) z r(K) + [x - w(K) z] r_K(K)}{[r(K)]^2},
\end{align*}
\]

and

\[
k_{xK}(z, x; K) = -\frac{r_K}{[r(K)]^2}.
\]

\(^{34}\)Our numerical solution is only very close to zero due to approximation errors.
Figure C.1: Savings and distribution sensitivity functions for high (-) and low (-.-) productivity shocks.

D.2 Detailed Results

Individual welfare and distributional gains for agents with high and low labor productivity shocks are shown in Figure D.1. The top panel compares gains from the optimal tax schedule imposed on the progressive tax steady state. Individual welfare gains are decreasing functions at low levels of wealth. High labor-income agents have larger gains than low labor-income agents. The bottom panel displays welfare and distributional gains from the optimal tax schedule reform of the flat-tax steady state. Both gains are monotonically decreasing at all asset and income levels. Most of the asset-poor agents have both welfare and distributional gains while the rich have losses relative to the steady state with the
Figure D.1: Individual welfare and distributional gains in the optimal tax schedule steady state.

There are two forces present: first is the tax rate (especially for the rich agents in the flat tax steady state) and general equilibrium effects. The huge welfare gains (5-20%) for poor agents are mostly due to the higher wage in the optimal steady state. Note that the big efficiency gain from the optimal tax schedule is not sufficient to compensate all agents for the more unequal distribution (compared to the progressive tax steady state, an agent with a low productivity shock has always a distributional loss in Figure D.1, top panel).

Table D.1 shows the distribution of resources for quintiles of the wealth distribution. Because of the high tax rate on incomes in the bottom quintile, these agents in the optimal tax schedule steady state consume 6.5% less than those of the progressive tax schedule.
Steady State Distribution of Resources

<table>
<thead>
<tr>
<th>Tax Schedule</th>
<th>1st</th>
<th>2nd</th>
<th>3rd</th>
<th>4th</th>
<th>5th</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Average Consumption Level</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Optimal</td>
<td>0.6673</td>
<td>0.8272</td>
<td>0.9171</td>
<td>1.0076</td>
<td>1.1489</td>
</tr>
<tr>
<td>Flat</td>
<td>0.6990</td>
<td>0.8189</td>
<td>0.8904</td>
<td>0.9685</td>
<td>1.1213</td>
</tr>
<tr>
<td>Progressive</td>
<td>0.7120</td>
<td>0.8113</td>
<td>0.8642</td>
<td>0.9237</td>
<td>1.0148</td>
</tr>
<tr>
<td><strong>Average Asset Level</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Optimal</td>
<td>0.9166</td>
<td>2.4051</td>
<td>3.6545</td>
<td>4.9934</td>
<td>7.0381</td>
</tr>
<tr>
<td>Flat</td>
<td>0.6336</td>
<td>1.7708</td>
<td>2.8488</td>
<td>4.1933</td>
<td>6.9921</td>
</tr>
<tr>
<td>Progressive</td>
<td>0.7705</td>
<td>1.8195</td>
<td>2.5666</td>
<td>3.2509</td>
<td>4.2725</td>
</tr>
<tr>
<td><strong>Average Investment/Income Ratio</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Optimal</td>
<td>0.1222</td>
<td>0.2148</td>
<td>0.2935</td>
<td>0.3692</td>
<td>0.4643</td>
</tr>
<tr>
<td>Flat</td>
<td>0.0869</td>
<td>0.1630</td>
<td>0.2363</td>
<td>0.3197</td>
<td>0.4624</td>
</tr>
<tr>
<td>Progressive</td>
<td>0.1058</td>
<td>0.1774</td>
<td>0.2252</td>
<td>0.2592</td>
<td>0.2964</td>
</tr>
<tr>
<td><strong>Average Income Level</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Optimal</td>
<td>1.0141</td>
<td>1.1520</td>
<td>1.2353</td>
<td>1.3221</td>
<td>1.4620</td>
</tr>
<tr>
<td>Flat</td>
<td>0.9644</td>
<td>1.1009</td>
<td>1.1881</td>
<td>1.2883</td>
<td>1.4884</td>
</tr>
<tr>
<td>Progressive</td>
<td>0.9158</td>
<td>1.0565</td>
<td>1.1410</td>
<td>1.2337</td>
<td>1.3739</td>
</tr>
<tr>
<td><strong>Average Tax Contribution Share</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Optimal</td>
<td>0.1944</td>
<td>0.1957</td>
<td>0.1995</td>
<td>0.2014</td>
<td>0.2090</td>
</tr>
<tr>
<td>Flat</td>
<td>0.1588</td>
<td>0.1824</td>
<td>0.1982</td>
<td>0.2133</td>
<td>0.2473</td>
</tr>
<tr>
<td>Progressive</td>
<td>0.1311</td>
<td>0.1702</td>
<td>0.1988</td>
<td>0.2258</td>
<td>0.2741</td>
</tr>
</tbody>
</table>

Table D.1: Distribution of resources in steady states.

However, in all the other quintiles the optimal tax schedule steady state, agents consume on average more than in the other two steady states. Dividing these levels by the average consumption in each steady state, we can calculate average quintile consumption relative to the steady state average. Under the optimal tax schedule, the bottom quintile consumes 73% of the average consumption, in the flat tax it is 77%, in the progressive it is 82%. This shows that incentives of the optimal tax schedule put more emphasis on savings than on redistribution relative to the progressive and the flat tax schedules.

The distribution of capital reveals that the incentives contained in the optimal-tax schedule move the distribution to higher capital levels. The poorest quintile owns on average 17% more assets than in the progressive steady state. This increase is even larger for the other quintiles (40% on the top). Again, the flat-rate steady state leads to a lower level of savings by the bottom two quintiles. These levels are reflected in the shares of the total capital stock. For all steady states the bottom quintile owns only around 5% of the total stock while the top quintile around one third (43% in the flat-tax steady state).

The investment-to-income ratios reveal the agents in the bottom quintile of the optimal schedule invest much more than similar agents in the other two steady states. Agents in the optimal tax schedule steady state invest 30% of their income, more than those in the flat-tax (27%) and progressive (22%) steady states. The investment is also more evenly distributed over the quintiles.

The distribution of income and tax contributions distribution show the differences be-
Reform of the Progressive Tax Steady State
Value in the Initial Steady State (−) and from Transition (−−)

Reform of the Flat Tax Steady State
Value in the Initial Steady State (−) and from Transition (−−)

Figure D.2: Value functions: Steady states and transition from the progressive and the flat tax steady state to the optimal tax schedule steady state.

The top panel in Figure D.2 illustrates the welfare gains presented in Table 2. It shows the expected present discounted values in the progressive-rate steady state and at the moment between the three tax schedules. The progressive tax helps the bottom quintile while the flat tax helps the top quintile. The U-shape of the optimal tax provides savings incentives at the cost of lower after-tax income for the poor agents. The optimal tax equalizes tax contribution share of total tax revenues across the quintiles. Both the flat-tax and progressive-tax steady states put more relative burden on the higher income quintiles.

D.3 Transition to the Optimal Tax Schedule Steady State
The top panel in Figure D.2 illustrates the welfare gains presented in Table 2. It shows the expected present discounted values in the progressive-rate steady state and at the moment
Figure D.3: Individual welfare and distributional gains: Transition from the progressive and the flat tax steady state to the optimal tax schedule steady state.

of the unanticipated reform to the optimal tax schedule. While 73% of the population is better off from the reform, it is not Pareto improving as the poorest 27% of all households are worse off (they are hit by high tax rates from the optimal schedule).

The bottom panel in Figure D.2 shows the expected present discounted values of the flat-rate steady state and of the transition to the optimal tax schedule. Political support is not sufficient, equal only to 33% of the population, mostly agents in the middle of the asset distribution.

Finally, Figure D.3 shows the efficiency and distributional individual gains from tran-
Relative to the steady state analysis, the averages for the progressive steady state reform decline: while the average welfare and efficiency gains remain still positive the distributional loss reaches negative 7%. A reform from the flat rate steady state delivers average welfare and efficiency losses but improves the distribution. Note that due to sizeable general equilibrium effects, welfare gains of poor agents are positive.

\[\text{35 These gains are defined in the same way as in the steady state. A gain from transition is a constant, per-period percentage of consumption in the original steady state that equalizes its corresponding expected present discounted value from the whole transition. For details, see in Domeij and Heathcote (2004).}\]