Appendix:

Does the US Tax Code Favor Automation? Acemoglu, Manera, Restrepo

A.1 Robustness Checks and Additional Figures Discussed in the Main Text

This part of the Appendix presents the following additional results and robustness checks discussed in the main text:

- Figure A.1 provides the time-series of the total value of depreciation allowances by type of capital, α^j, and compares this to the allowance that would result from economic depreciation. Each figure presents a single average across the types of assets included in each category (software, equipment and non-residential structures).
- Figure A.2 provides the time-series of the tax rates on capital income, corporate income and personal income, 1981-2018.
- Figure A.3 presents the evolution of effective taxes on capital when all investment is financed with equity. For comparison, we also show the effective tax on labor.
- In Table A.1 we assume that the wedge ϱ only distorts the extensive margin of labor supply. This reduces employment and welfare gains, but they remain positive.
- In Table A.2 we additionally include the implicit tax on labor implied by means-tested programs. With this higher effective tax on labor (equal to 33.5%), there are greater employment and welfare gains from moving towards optimal taxes and lower estimation.
- Table A.3 considers the possibility that capital directly complements labor at labor-intensive tasks (see footnote 32). In particular, we assume that capital represents 20% of the value added in labor-intensive tasks that are not yet automated.
- Table A.4 is the analogue of Table 1 when the effective tax on capital is based on full equity financing. This leads to somewhat lower employment and welfare gains from moving to optimal taxes.
- Table A.5 presents a version of Table 1 when there is a 15% wedge for capital. This leads to employment and welfare gains that are approximately half as large as those in Table 1.

- In Table A.6 we set $\upsilon = 0$, so that labor has an absolute disadvantage in tasks where it has a comparative advantage. In this case, employment and welfare gains are significantly larger.
- Table A.7 follows our extension in Section 5.1 by adding the endogenous response of human capital to the elasticity of labor supply. This leads to significantly larger employment and welfare gains from moving towards optimal taxes.
- In Table A.8 sets $\varepsilon^k = 1$. This leads to employment and welfare gains that are about half as large as in our baseline in Table 1.

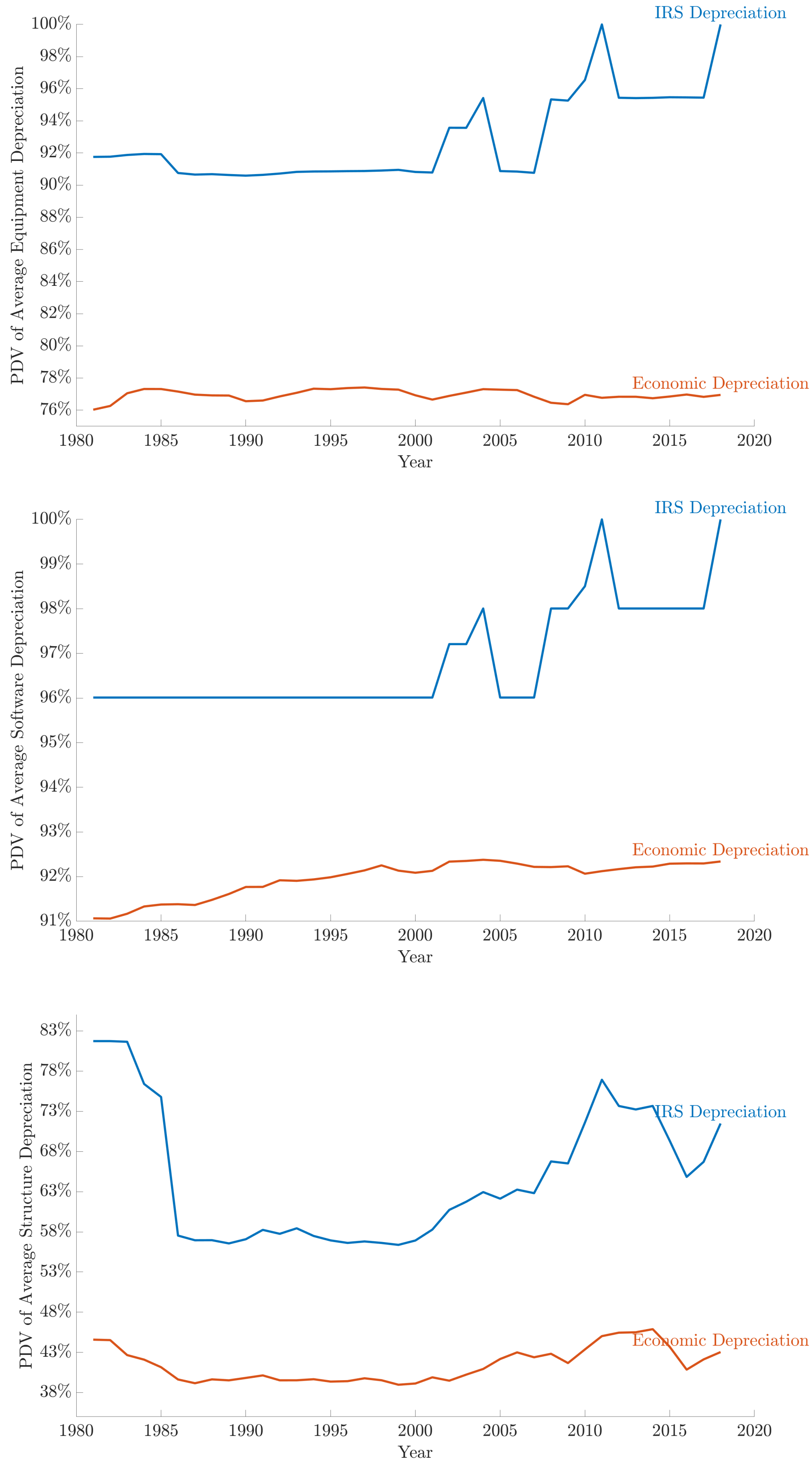

Figure A.1: Estimated depreciation allowances over time for equipment, software and non-residential structures.

Notes: See the text for definitions.

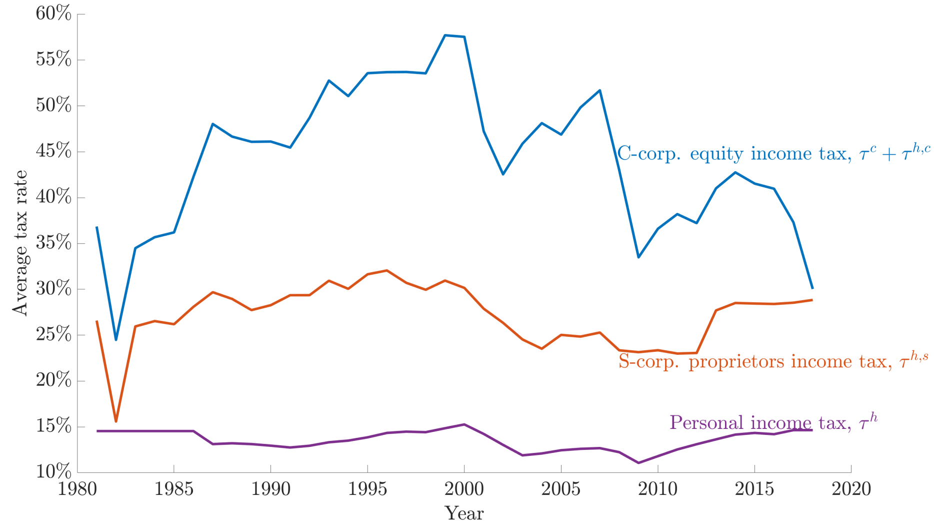

Figure A.2: Average tax rates on capital income, corporate income and personal income, 1981-2018.

Notes: See the text for the definitions and sources.

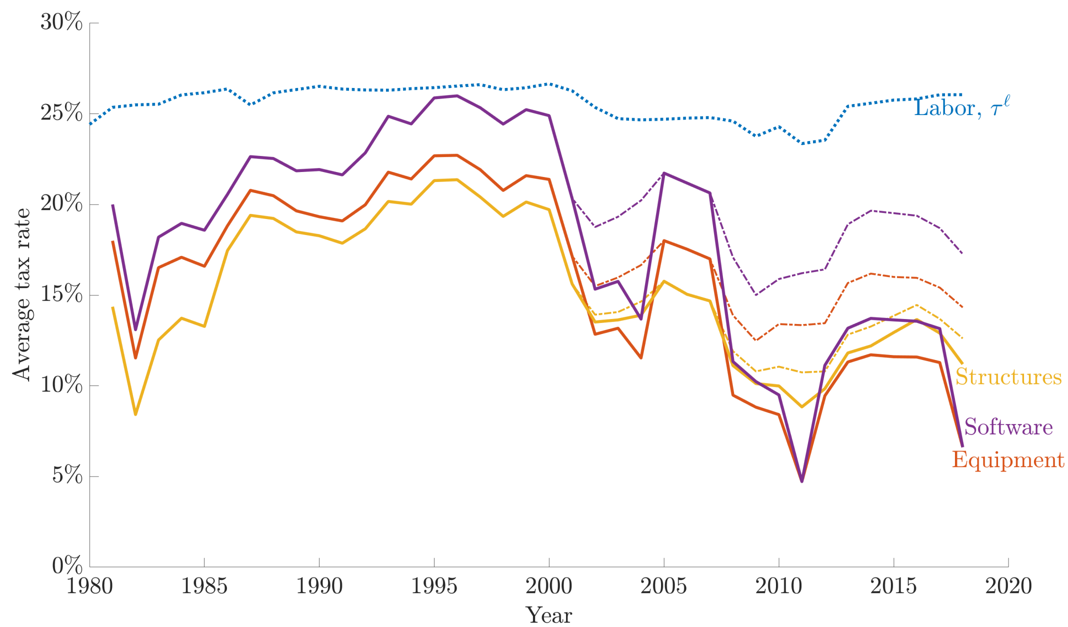

Figure A.3: Effective tax rates on labor, software capital, equipment, and non-residential structures with equity financing.

Notes: The alternative series for the effective tax rate on labor includes the phase out of means tested programs. See the text for definitions and sources.

Table A.1: Robustness: distortions at the extensive margin of labor supply

	Current System	Ramsey Solution	Distorting θ	Distorting θ and changing τ^k	Distorting θ and changing τ^ℓ
	(1)	(2)	(3)	(4)	(5)
Tax system:					
τ^k	10.00%	24.27%	10.00%	8.63%	10.00%
τ^ℓ	25.50%	19.24%	25.50%	25.50%	24.97%
θ	0.276	0.266	0.268	0.267	0.266
τ^A	0.00%	0.00%	8.37%	10.98%	11.18%
Aggregates:					
Employment	.	+3.50%	+0.95%	+1.35%	+1.69%
Labor Share	56.00%	56.66%	57.58%	58.06%	58.15%
Net Output	.	+0.47%	–0.06%	+0.17%	+0.22%
C.E. welfare change	.	0.27%	0.06%	0.10%	0.12%
Revenue	.	0.00%	+1.17%	0.00%	0.00%

Notes: This table shows the effective capital and labor taxes, the level of automation and the automation tax under different scenarios. It also presents the implied changes in employment, output, welfare, and government revenue, and the level of labor share in national income. The first column is for the current US tax system. The second column shows the unconstrained Ramsey solution. Column 3 considers the implications of changing the level of automation, θ via automation taxes (and no other change in policy). Column 4 additionally allows a change in the effective tax on capital, and column 5 considers a change in the effective tax on labor. Change in welfare is in terms of consumption equivalent. See the text for details.

TABLE A.2: Robustness: including the implicit tax on labor from means-tested and disability programs.

	Current System	Ramsey Solution	Distorting θ	Distorting θ and changing τ^k	Distorting θ and changing τ^ℓ
	(1)	(2)	(3)	(4)	(5)
Tax system:					
τ^k	10.00%	31.20%	10.00%	6.44%	10.00%
τ^ℓ	33.50%	23.98%	33.50%	33.50%	32.15%
θ	0.281	0.265	0.270	0.266	0.264
τ^A	0.00%	0.00%	12.13%	17.01%	17.32%
Aggregates:					
Employment	.	+6.07%	+1.37%	+2.23%	+3.18%
Labor Share	56.00%	57.09%	58.36%	59.29%	59.51%
Net Output	.	+1.13%	–0.17%	+0.38%	+0.57%
C.E. welfare change	.	0.81%	0.15%	0.28%	0.41%
Revenue	.	0.00%	+2.03%	0.00%	0.00%

Notes: This table shows the effective capital and labor taxes, the level of automation and the automation tax under different scenarios. It also presents the implied changes in employment, output, welfare, and government revenue, and the level of labor share in national income. The first column is for the current US tax system. The second column shows the unconstrained Ramsey solution. Column 3 considers the implications of changing the level of automation, θ via automation taxes (and no other change in policy). Column 4 additionally allows a change in the effective tax on capital, and column 5 considers a change in the effective tax on labor. Change in welfare is in terms of consumption equivalent. See the text for details.

TABLE A.3: Robustness: allowing for within-task complementarities between capital and labor.

	Current System	Ramsey Solution	Distorting θ	Distorting θ and changing τ^k	Distorting θ and changing τ^ℓ
	(1)	(2)	(3)	(4)	(5)
Tax system:					
τ^k	10.00%	26.65%	10.00%	9.28%	10.00%
τ^ℓ	25.50%	18.22%	25.50%	25.50%	25.27%
θ	0.159	0.150	0.151	0.152	0.151
τ^A	0.00%	0.00%	6.53%	6.35%	6.58%
Aggregates:					
Employment	.	+3.96%	+0.62%	+0.64%	+0.82%
Labor Share	55.98%	56.68%	57.01%	56.96%	57.02%
Net Output	.	+0.43%	−0.05%	+0.09%	+0.12%
C.E. welfare change	.	0.41%	0.05%	0.07%	0.10%
Revenue	.	0.00%	+12.63%	0.00%	0.00%

Notes: This table shows the effective capital and labor taxes, the level of automation and the automation tax under different scenarios. It also presents the implied changes in employment, output, welfare, and government revenue, and the level of labor share in national income. The first column is for the current US tax system. The second column shows the unconstrained Ramsey solution. Column 3 considers the implications of changing the level of automation, θ via automation taxes (and no other change in policy). Column 4 additionally allows a change in the effective tax on capital, and column 5 considers a change in the effective tax on labor. Change in welfare is in terms of consumption equivalent. See the text for details.

Table A.4: Robustness: effective tax on capital for equity financing only

	Current System	Ramsey Solution	Distorting θ	Distorting θ and changing τ^k	Distorting θ and changing τ^ℓ
	(1)	(2)	(3)	(4)	(5)
Tax system:					
τ^k	12.00%	27.17%	12.00%	10.71%	12.00%
τ^ℓ	25.50%	18.88%	25.50%	25.50%	25.00%
θ	0.276	0.266	0.268	0.266	0.265
τ^A	0.00%	0.00%	9.48%	11.96%	12.07%
Aggregates:					
Employment	.	+3.64%	+1.07%	+1.46%	+1.76%
Labor Share	56.00%	56.71%	57.80%	58.26%	58.33%
Net Output	.	+0.35%	−0.08%	+0.13%	+0.16%
C.E. welfare change	.	0.33%	0.08%	0.12%	0.15%
Revenue	.	0.00%	+1.10%	0.00%	0.00%

Notes: This table shows the effective capital and labor taxes, the level of automation and the automation tax under different scenarios. It also presents the implied changes in employment, output, welfare, and government revenue, and the level of labor share in national income. The first column is for the current US tax system. The second column shows the unconstrained Ramsey solution. Column 3 considers the implications of changing the level of automation, θ via automation taxes (and no other change in policy). Column 4 additionally allows a change in the effective tax on capital, and column 5 considers a change in the effective tax on labor. Change in welfare is in terms of consumption equivalent. See the text for details.

TABLE A.5: Robustness: capital wedge of 15%

	Current System	Ramsey Solution	Distorting θ	Distorting θ and changing τ^k	Distorting θ and changing τ^ℓ
	(1)	(2)	(3)	(4)	(5)
Tax system:					
τ^k	10.00%	21.98%	10.00%	8.85%	10.00%
τ^ℓ	25.50%	20.23%	25.50%	25.50%	25.04%
θ	0.272	0.264	0.267	0.265	0.264
τ^A	0.00%	0.00%	5.59%	9.17%	9.36%
Aggregates:					
Employment	.	+2.98%	+0.63%	+1.12%	+1.42%
Labor Share	56.00%	56.53%	57.03%	57.69%	57.77%
Net Output	.	+0.47%	−0.01%	+0.17%	+0.22%
C.E. welfare change	.	0.20%	0.03%	0.07%	0.09%
Revenue	.	0.00%	+0.79%	0.00%	0.00%

Notes: This table shows the effective capital and labor taxes, the level of automation and the automation tax under different scenarios. It also presents the implied changes in employment, output, welfare, and government revenue, and the level of labor share in national income. The first column is for the current US tax system. The second column shows the unconstrained Ramsey solution. Column 3 considers the implications of changing the level of automation, θ via automation taxes (and no other change in policy). Column 4 additionally allows a change in the effective tax on capital, and column 5 considers a change in the effective tax on labor. Change in welfare is in terms of consumption equivalent. See the text for details.

Table A.6: Robustness: labor has an absolute disadvantage at higher-indexed tasks

	Current System	Ramsey Solution	Distorting θ	Distorting θ and changing τ^k	Distorting θ and changing τ^ℓ
	(1)	(2)	(3)	(4)	(5)
Tax system:					
τ^k	10.00%	24.12%	10.00%	8.32%	10.00%
τ^ℓ	25.50%	14.98%	25.50%	25.50%	24.34%
θ	0.637	0.618	0.622	0.619	0.617
τ^A	0.00%	0.00%	12.72%	15.70%	15.89%
Aggregates:					
Employment	.	+7.30%	+2.23%	+2.89%	+3.76%
Labor Share	56.00%	57.96%	59.21%	59.91%	60.17%
Net Output	.	+0.62%	–0.17%	+0.23%	+0.31%
C.E. welfare change	.	0.85%	0.22%	0.32%	0.43%
Revenue	.	0.00%	+2.50%	0.00%	0.00%

Notes: This table shows the effective capital and labor taxes, the level of automation and the automation tax under different scenarios. It also presents the implied changes in employment, output, welfare, and government revenue, and the level of labor share in national income. The first column is for the current US tax system. The second column shows the unconstrained Ramsey solution. Column 3 considers the implications of changing the level of automation, θ via automation taxes (and no other change in policy). Column 4 additionally allows a change in the effective tax on capital, and column 5 considers a change in the effective tax on labor. Change in welfare is in terms of consumption equivalent. See the text for details.

Table A.7: Robustness: accounting for human capital responses

	Current System	Ramsey Solution	Distorting θ	Distorting θ and changing τ^k	Distorting θ and changing τ^ℓ
	(1)	(2)	(3)	(4)	(5)
Tax system:					
τ^k	10.00%	29.21%	10.00%	7.72%	10.00%
τ^ℓ	25.50%	16.90%	25.50%	25.50%	24.63%
θ	0.290	0.275	0.278	0.275	0.274
τ^A	0.00%	0.00%	12.53%	15.39%	15.43%
Aggregates:					
Employment	.	+5.73%	+1.73%	+2.35%	+2.97%
Labor Share	56.00%	57.06%	58.52%	59.07%	59.18%
Net Output	.	+1.23%	+0.03%	+0.48%	+0.62%
C.E. welfare change	.	0.59%	0.14%	0.23%	0.30%
Revenue	.	0.00%	+1.98%	0.00%	0.00%

Notes: This table shows the effective capital and labor taxes, the level of automation and the automation tax under different scenarios. It also presents the implied changes in employment, output, welfare, and government revenue, and the level of labor share in national income. The first column is for the current US tax system. The second column shows the unconstrained Ramsey solution. Column 3 considers the implications of changing the level of automation, θ via automation taxes (and no other change in policy). Column 4 additionally allows a change in the effective tax on capital, and column 5 considers a change in the effective tax on labor. Change in welfare is in terms of consumption equivalent. See the text for details.

Table A.8: Robustness: setting $\epsilon^k = 1$

	Current System	Ramsey Solution	Distorting θ	Distorting θ and changing τ^k	Distorting θ and changing τ^ℓ
	(1)	(2)	(3)	(4)	(5)
Tax system:					
τ^k	10.00%	20.83%	10.00%	9.16%	10.00%
τ^ℓ	25.50%	20.97%	25.50%	25.50%	25.19%
θ	0.256	0.248	0.250	0.248	0.248
τ^A	0.00%	0.00%	6.76%	9.09%	9.09%
Aggregates:					
Employment	.	+2.18%	+0.62%	+0.92%	+1.08%
Labor Share	56.00%	56.48%	57.23%	57.64%	57.68%
Net Output	.	–0.38%	–0.27%	–0.17%	–0.19%
C.E. welfare change	.	0.18%	0.04%	0.07%	0.09%
Revenue	.	0.00%	+0.66%	0.00%	0.00%

Notes: This table shows the effective capital and labor taxes, the level of automation and the automation tax under different scenarios. It also presents the implied changes in employment, output, welfare, and government revenue, and the level of labor share in national income. The first column is for the current US tax system. The second column shows the unconstrained Ramsey solution. Column 3 considers the implications of changing the level of automation, θ via automation taxes (and no other change in policy). Column 4 additionally allows a change in the effective tax on capital, and column 5 considers a change in the effective tax on labor. Change in welfare is in terms of consumption equivalent. See the text for details.

A.2 Derivations and Proofs for the Static Model

This part of the Appendix presents the proofs of the results stated in the text and some additional results briefly mentioned in the text.

Characterization of the Equilibrium and the Ramsey Problem

The next lemma provides the characterization of the competitive equilibrium presented in the text and is the basis of all subsequent proofs.

Lemma A.1 (Equilibrium characterization) *Given a tax system* (τ^k, τ^ℓ) *and a labor wedge* ϱ*, a market equilibrium is given by an allocation* $\{k, \ell\}$ *and a threshold task* θ *such that:*

- *output* y *is given by* $f(k, \ell; \theta)$ *in (2);*
- $\theta = \theta^m(k, \ell)$ *maximizes* $f(k, \ell; \theta)$*;*
- *the capital and labor market-clearing conditions,* (4) *and* (5)*, are satisfied;*
- *tax revenues are given by* (6)*.*

Proof of Lemma A.1. The unit cost of producing task x with labor is

$$p^\ell(x) = \frac{w}{\psi^\ell(x)},$$

whereas the unit cost of producing task x with capital is

$$p^k(x) = \frac{R}{\psi^k(x)}.$$

Because the allocation of tasks to factors is cost-minimizing and because $\psi^\ell(x)/\psi^k(x)$ is (strictly) increasing, there exists a threshold θ such that all tasks below the threshold are produced with capital and those above it will be produced with labor.

The demand for capital in the economy therefore comes from tasks $x \leq \theta$ and satisfies

$$k = \int_0^\theta k(x)dx = \int_0^\theta \frac{y(x)}{\psi^k(x)}dx = \int_0^\theta \frac{y \cdot p^k(x)^{-\lambda}}{\psi^k(x)}dx = y \cdot R^{-\lambda} \cdot \int_0^\theta \psi^k(x)^{\lambda-1}dx,$$

which can be rearranged as

$$R = \left(\frac{y}{k}\right)^{\frac{1}{\lambda}} \cdot \left(\int_0^\theta \psi^k(x)^{\lambda-1}dx\right)^{\frac{1}{\lambda}}. \tag{A.1}$$

Moreover, this equation also implies

$$R = f_k,$$

where $f(k, \ell; \theta)$ is given in (2). Next note that the supply of capital by households is given by the Euler equation,

$$u'(\bar{y} - k) = 1 + (R - \delta) \cdot (1 - \tau^k),$$

which, combined with $R = f_k$, gives the capital market-clearing condition in (4).

Similarly, the demand for labor comes from tasks $x > \theta$ and is given by

$$\ell = \int_\theta^1 \ell(x) dx = \int_\theta^1 \frac{y(x)}{\psi^\ell(x)} dx = \int_\theta^1 \frac{y \cdot p^\ell(x)^{-\lambda}}{\psi^\ell(x)} dx = y \cdot w^{-\lambda} \int_\theta^1 \psi^\ell(x)^{\lambda-1} dx,$$

which can be rearranged as

$$w = \left(\frac{y}{\ell}\right)^{\frac{1}{\lambda}} \cdot \left(\int_\theta^1 \psi^\ell(x)^{\lambda-1} dx\right)^{\frac{1}{\lambda}}. \tag{A.2}$$

This equation implies that

$$w = f_\ell.$$

where $f(k, \ell; \theta)$ is given in (2). Moreover, the supply of labor by households is given by the optimality condition for labor supply,

$$\nu'(\ell) = w \cdot (1 - \tau^\ell),$$

which, combined with $w = f_\ell$, gives the labor market-clearing condition (5).

We next prove that output is given by $f(k, \ell; \theta)$. Since the final good is the numeraire, the ideal price condition is

$$1 = \int_0^\theta p^k(x)^{1-\lambda} dx + \int_\theta^1 p^\ell(x)^{1-\lambda} dx.$$

Substituting task prices in terms of factor prices before taxes, this condition yields

$$1 = R^{1-\lambda} \cdot \int_0^\theta \psi^k(x)^{\lambda-1} dx + w^{1-\lambda} \cdot \int_\theta^1 \psi^\ell(x)^{\lambda-1} dx.$$

Replacing the expressions for R and w from equations (A.1) and (A.2), we obtain the ideal price condition in terms of output, capital, labor, the level of automation and the production parameters:

$$1 = \left(\frac{y}{k}\right)^{\frac{1-\lambda}{\lambda}} \cdot \left(\int_0^\theta \psi^k(x)^{\lambda-1}dx\right)^{\frac{1}{\lambda}} + \left(\frac{y}{\ell}\right)^{\frac{1-\lambda}{\lambda}} \cdot \left(\int_\theta^1 \psi^\ell(x)^{\lambda-1}dx\right)^{\frac{1}{\lambda}}.$$

Solving for y in this equation gives $y = f(k,\ell;\theta)$ as in (2).

We now turn to the determination of θ. Because task allocations are cost-minimizing, the thresholds task θ satisfies

$$\frac{w}{\psi^\ell(x)} = \frac{R}{\psi^k(x)} \Rightarrow \frac{w}{R} = \frac{\psi^\ell(\theta)}{\psi^k(\theta)}.$$

Since $R = f_k$ and $w = f_\ell$, we can rewrite this as

$$\frac{f_\ell}{f_k} = \frac{\psi^\ell(\theta)}{\psi^k(\theta)}. \tag{A.3}$$

This equation has a unique solution $\theta^m(k,\ell)$. Uniqueness is a consequence of the fact that the right-hand side is continuous and increasing in θ (by assumption). The left-hand side, on the other hand, can be written as

$$\frac{f_\ell}{f_k} = \left(\frac{k}{\ell}\frac{\int_\theta^1 \psi^\ell(x)^{\lambda-1}dx}{\int_0^\theta \psi^k(x)^{\lambda-1}dx}\right)^{\frac{1}{\lambda}},$$

and is thus decreasing in θ. This implies that a solution $\theta^m(k,\ell)$ always exists in view of the fact that the left-hand side goes from ∞ (at $\theta = 0$) to 0 (at $\theta = 1$).

We now show that $\theta^m(k,\ell)$ maximizes $f(k,\ell;\theta)$. An infinitesimal change in θ leads to a change in output of

$$f_\theta(k,\ell;\theta) = \frac{y}{1-\lambda}\left(\left(\frac{f_\ell}{\psi^\ell(\theta)}\right)^{1-\lambda} - \left(\frac{f_k}{\psi^k(\theta)}\right)^{1-\lambda}\right). \tag{A.4}$$

This expression follows by totally differentiating (2), which yields

$$\begin{aligned} f_\theta(k,\ell;\theta) =& \frac{1}{1-\lambda}\psi^\ell(\theta)^{\lambda-1} \cdot y^{\frac{1}{\lambda}} \cdot \ell^{\frac{\lambda-1}{\lambda}} \cdot \left(\int_\theta^1 \psi^\ell(x)^{\lambda-1}dx\right)^{\frac{1-\lambda}{\lambda}} \\ &- \frac{1}{1-\lambda}\psi^k(\theta)^{\lambda-1} \cdot y^{\frac{1}{\lambda}} \cdot k^{\frac{\lambda-1}{\lambda}} \cdot \left(\int_0^\theta \psi^k(x)^{\lambda-1}dx\right)^{\frac{1-\lambda}{\lambda}}. \end{aligned}$$

Regrouping terms yields

$$f_\theta(k,\ell;\theta)=\frac{y}{1-\lambda}\left(\psi^\ell(\theta)^{\lambda-1}\cdot\left(\frac{y}{\ell}\cdot\int_\theta^1\psi^\ell(x)^{\lambda-1}dx\right)^{\frac{1-\lambda}{\lambda}}-\psi^k(\theta)^{\lambda-1}\cdot\left(\frac{y}{k}\cdot\int_0^\theta\psi^k(x)^{\lambda-1}dx\right)^{\frac{1-\lambda}{\lambda}}\right).$$

Equation (A.4) follows after substituting in the formulae for f_k and f_ℓ in place of the terms in the inner parentheses.

Equation (A.4) further implies $f_\theta\geq 0$ to the left of $\theta^m(k,\ell)$, since in this region we have

$$\frac{f_\ell}{f_k}>\frac{\psi^\ell(\theta)}{\psi^k(\theta)}.$$

Moreover, $f_\theta<0$ the right of $\theta^m(k,\ell)$, since in this region we have

$$\frac{f_\ell}{f_k}<\frac{\psi^\ell(\theta)}{\psi^k(\theta)}.$$

Thus, $f(k,\ell,\theta)$ is single-peaked with a unique maximum at $\theta^m(k,\ell)$.

Finally, we compute equilibrium tax revenues. Capital taxes, which raise revenue from tasks below θ, generate total revenue:

$$\text{Revenue from capital }=\int_0^\theta\tau^k\cdot(R-\delta)\cdot k(x)dx=\tau^k\cdot(f_k-\delta)\cdot k,$$

where we used the fact that $R=f_k$ (from equation (A.1)). Likewise, labor taxes raise revenue from tasks above θ and thus:

$$\text{Revenue from labor}=\int_0^\theta\tau^\ell\cdot w\cdot\ell(x)dx=\tau^\ell\cdot f_\ell\cdot\ell,$$

where we used the fact that $w=f_\ell$ (from equation (A.2)). ■

The next lemma is straightforward but will be used repeatedly in our proofs.

Lemma A.2 *The production function $f(k,\ell;\theta^m(k,\ell))$ exhibits constant returns to scale and is concave in k and ℓ.*

Proof. We first show that $f(k,\ell;\theta^m(k,\ell))$ exhibits constant returns to scale. Because $f(k,\ell;\theta)$ exhibits constant returns to scale in k and ℓ (which is immediate from (2)), it is sufficient to prove that $\theta^m(k,\ell)$ is homogeneous of degree zero. Equation (A.3) implies that

$\theta^m(k,\ell)$ is the unique solution to

$$\left(\frac{k}{\ell}\frac{\int_\theta^1 \psi^\ell(x)^{\lambda-1}dx}{\int_0^\theta \psi^k(x)^{\lambda-1}dx}\right)^{\frac{1}{\lambda}} = \frac{\psi^\ell(\theta)}{\psi^k(\theta)},$$

which establishes that $\theta^m(k,\ell)$ only depends on k/ℓ and is thus homogeneous of degree zero.

Since $f(k,\ell;\theta^m(k,\ell))$ exhibits constant returns to scale in k and ℓ, it is concave if and only if it is quasi-concave in k and ℓ. Note that $h(k,\ell) = f(k,\ell;\theta^m(k,\ell))$ solves the optimization problem:

$$\text{(A.5)} \qquad f(k,\ell;\theta^m(k,\ell)) = \max_{k(x),\ell(x)\geq 0}\left(\int_0^1 y(x)^{\frac{\lambda-1}{\lambda}}dx\right)^{\frac{\lambda}{\lambda-1}},$$
$$\text{subject to: } y(x) = \psi^k(x)k(x) + \psi^\ell(x)\ell(x), \int_0^1 k(x)dx = k, \int_0^1 \ell(x)dx = \ell.$$

Suppose that $h(k_1,\ell_1) \geq b$ and $h(k_2,\ell_2) \geq b$, and denote by $\{k_1(x),\ell_1(x),y_1(x)\}$ and $\{k_2(x),\ell_2(x),y_2(x)\}$ the solution to (A.5) for $\{k_1,\ell_1\}$ and $\{k_2,\ell_2\}$, respectively. Consider the problem in (A.5) for $\{\alpha k_1 + (1-\alpha)k_2, \alpha\ell_1 + (1-\alpha)\ell_2\}$ for some $\alpha \in [0,1]$. The allocation $\{\alpha k_1(x) + (1-\alpha)k_2(x), \alpha\ell_1(x) + (1-\alpha)\ell_2(x), \alpha y_1(x) + (1-\alpha)y_2(x)\}$ satisfies the constraints in (A.5). Therefore,

$$h(\alpha k_1 + (1-\alpha)k_2, \alpha\ell_1 + (1-\alpha)\ell_2) \geq \left(\int_0^1 (\alpha y_1(x) + (1-\alpha)y_2(x))^{\frac{\lambda-1}{\lambda}}dx\right)^{\frac{\lambda}{\lambda-1}}.$$

Using the concavity of the constant elasticity of substitution function on the right-hand side of the above equation, we get

$$h(\alpha k_1 + (1-\alpha)k_2, \alpha\ell_1 + (1-\alpha)\ell_2) \geq \alpha\left(\int_0^1 y_1(x)^{\frac{\lambda-1}{\lambda}}dx\right)^{\frac{\lambda}{\lambda-1}} + (1-\alpha)\left(\int_0^1 y_2(x)^{\frac{\lambda-1}{\lambda}}dx\right)^{\frac{\lambda}{\lambda-1}} \geq b.$$

It follows that $h(k,\ell) = f(k,\ell;\theta^m(k,\ell))$ is quasi-concave in k and ℓ and hence concave in k and ℓ, completing the proof. ■

Main Proofs

In this section of the Appendix, we provide the proofs of the main results stated in the text. Before presenting the proofs of the results in the text, we provide a derivation of the Implementability Condition (IC) in (7). Exploiting the fact that f has constant returns to scale, we can rewrite the government budget constraint as follows

$$\begin{aligned} g \leq & \tau^k \cdot (f_k - \delta) \cdot k + \tau^\ell \cdot f_\ell \cdot \ell \\ = & f(k,\ell;\theta) + (1-\delta)\cdot k - (1 + (1-\tau^k)\cdot(f_k - \delta))\cdot k - (1-\tau^\ell)\cdot f_\ell \cdot \ell. \end{aligned}$$

Using the capital and labor market-clearing condition in equations (4) and (5), we can substitute out the terms $1 + (1-\tau^k)\cdot(f_k - \delta)$ and $(1-\tau^\ell)\cdot f_\ell$, which gives

$$g \leq f(k,\ell;\theta) + (1-\delta)\cdot k - u'(\bar{y} - k)\cdot k - \frac{\nu'(\ell)\cdot\ell}{1-\varrho},$$

which is the Implementability Condition used in the main text.

We next present the proofs of our main results.

Proof of Proposition 1. We start by solving for the optimal allocation. The utility of the representative household is given by

$$\text{utility} := u(c_0) + c - \nu(\ell) = f(k,\ell;\theta) + (1-\delta)\cdot k + u(\bar{y} - k) - \nu(\ell) - g.$$

The Ramsey problem can therefore be written as

$$\begin{aligned} & \max_{k,\ell,\theta} f(k,\ell;\theta) + (1-\delta)\cdot k + u(\bar{y} - k) - \nu(\ell) \\ & \text{subject to: } g \leq f(k,\ell;\theta) + (1-\delta)\cdot k - u'(\bar{y} - k)\cdot k - \frac{\nu'(\ell)\cdot\ell}{1-\varrho}. \end{aligned}$$

Both the objective function and the right-hand side of the constraint are increasing in θ. Thus, the optimal choice of θ maximizes $f(k,\ell;\theta)$, and this implies that $\theta = \theta^m(k,\ell)$, where $f_\theta(k,\ell;\theta^m(k,\ell)) = 0$, as claimed in the proposition.

With this choice, the problem becomes

$$\begin{aligned} & \max_{k,\ell} f(k,\ell;\theta^m(k,\ell)) + (1-\delta)\cdot k + u(\bar{y} - k) - \nu(\ell) \\ & \text{subject to: } g \leq f(k,\ell;\theta^m(k,\ell)) + (1-\delta)\cdot k - u'(\bar{y} - k)\cdot k - \frac{\nu'(\ell)\cdot\ell}{1-\varrho}. \end{aligned}$$

We next prove that the objective function is concave and the constraint set is convex.

The concavity of the objective function follows from Lemma A.2 and the fact that $u(\bar{y}-k)$ and $\nu(\ell)$ are convex in k and ℓ, respectively. The constraint defines a convex set since Lemma A.2 implies that $f(k,\ell;\theta^m(k,\ell))$ is concave and we assumed that $u'(\bar{y} - k)\cdot k$ and $\nu'(\ell)\cdot\ell$ are convex functions.

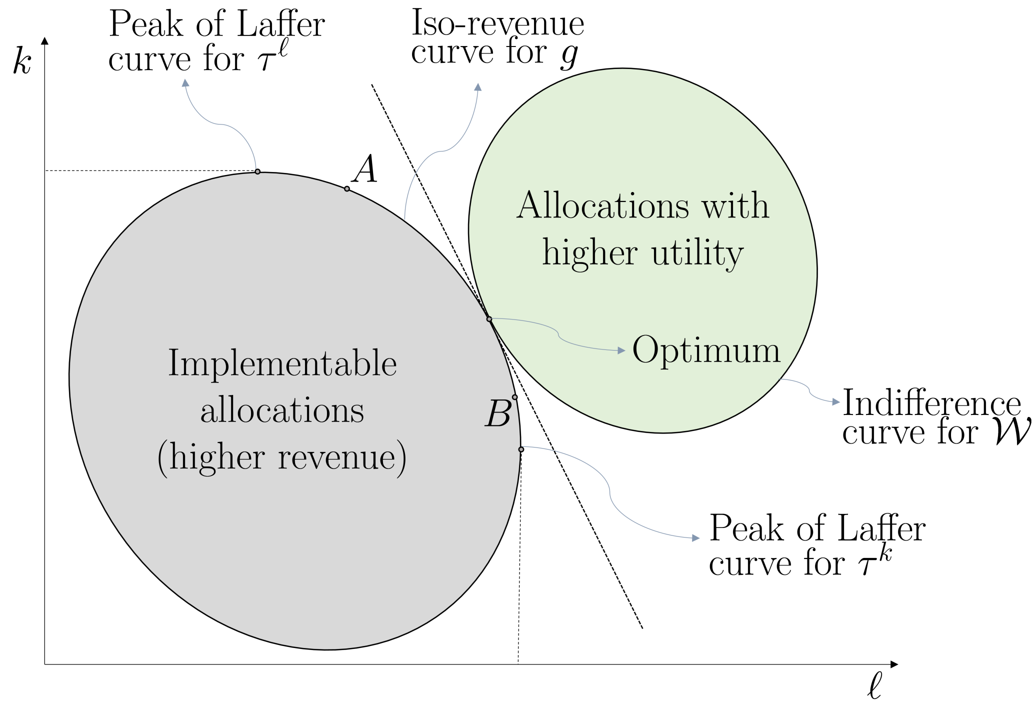

FIGURE A.4: Illustration of optimal policy problem.

Thus, the Ramsey problem is equivalent to the maximization of a concave function over a convex set. This implies that for any $g > 0$, the optimum is unique and yields some utility $\mathcal{W}$. Figure A.4 illustrates this optimum. The figure plots the set of points that satisfies the IC constraint—the points within the iso-revenue curve for g—and further identifies the set of points that yield higher utility than the optimal allocation, which are those inside this contour set of $\mathcal{W}$. The optimal allocation is given by the tangency point between the iso-revenue curve and the contour sets of $\mathcal{W}$.

At this point, the marginal utility per unit of revenue loss from an increase in k (denoted by $U^k(k,\ell)$) equals the marginal utility per unit of revenue loss from a increase in ℓ (denoted by $U^\ell(k,\ell)$), and both are equal to the multiplier μ, which denotes the marginal utility per unit of additional revenue. These marginal utilities can be computed as

$$U^k(k,\ell) = -\frac{\partial \text{utility}}{\partial k}\bigg/\frac{\partial \text{revenue}}{\partial k} = \frac{f_k - \delta - u'(\bar{y}-k) + 1}{-u''(\bar{y}-k)\cdot k + u'(\bar{y}-k) - 1 - f_k + \delta},$$

$$U^\ell(k,\ell) = -\frac{\partial \text{utility}}{\partial \ell}\bigg/\frac{\partial \text{revenue}}{\partial \ell} = \frac{f_\ell - \nu'(\ell)}{\frac{\nu''(\ell)}{1-\varrho}\cdot \ell + \frac{\nu'(\ell)}{1-\varrho} - f_\ell}.$$

The optimum allocation is given by the unique set of points along the iso-revenue curve for g for which

$$U^k(k,\ell) = U^\ell(k,\ell) = \mu.$$

We next prove that this unique optimal allocation can be implemented using the taxes in (8). Starting from $U^k(k,\ell) = \mu$, we obtain

$$U^k(k,\ell) = \frac{f_k - \delta - u'(\bar{y}-k) + 1}{-u''(\bar{y}-k)\cdot k + u'(\bar{y}-k) - (f_k + 1 - \delta)} = \mu.$$

Dividing the numerator and denominator on the left-hand side by $u'(\bar{y}-k) - 1$ (which is positive by assumption) and using (4) to substitute out for $f_k - \delta$, yields

$$\frac{\tau^k}{1-\tau^k} \Big/ \left(\frac{1}{\varepsilon^k(k)} - \frac{\tau^k}{1-\tau^k}\right) = \mu,$$

which can be rearranged to obtain the formula for $\tau^k/(1-\tau^k)$ in (8).

Likewise, starting from $U^\ell(k,\ell) = \mu$, we obtain

$$U^\ell(k,\ell) = \frac{f_\ell - \nu'(\ell)}{\frac{\nu''(\ell)}{1-\varrho}\cdot \ell + \frac{\nu'(\ell)}{1-\varrho} - f_\ell} = \mu.$$

Dividing the numerator and denominator on the left-hand side by $\nu'(\ell)/(1-\varrho)$ and using (5) to substitute out for f_ℓ, we obtain

$$\left(\frac{\tau^\ell}{1-\tau^\ell} + \varrho\right) \Big/ \left(\frac{1}{\varepsilon^\ell(\ell)} - \frac{\tau^\ell}{1-\tau^\ell}\right) = \mu,$$

which gives the formula for $\tau^\ell/(1-\tau^\ell)$ in (8). ∎

Proof of Corollary 1. Obtained by substituting $\varepsilon^k(k) = \varepsilon^\ell(\ell)$ and $\varrho = 0$ in (8). ∎

Proof of Corollary 2. First, note that the function $U^k(k,\ell)$ is decreasing in k and increasing in ℓ. This follows from our assumptions that $u'(\bar{y}-k)\cdot k$ is convex (which implies that $-u''(\bar{y}-k)\cdot k + u'(\bar{y}-k)$ is increasing in k) and u is a concave function (which implies that $u'(\bar{y}-k)$ is increasing in k), and the fact that Lemma A.2 implies that f_k is decreasing in k and increasing in ℓ.

Likewise, the function $U^\ell(k,\ell)$ is increasing in k and decreasing in ℓ. This follows from our assumptions that $\nu'(\ell)\cdot\ell$ is convex (which implies that $\nu'(\ell) + \nu''(\ell)\cdot\ell$ is increasing in ℓ) and ν is a convex function (which implies that $\nu'(\ell)$ is increasing in ℓ), and the fact that Lemma A.2 implies that f_ℓ is decreasing in ℓ and increasing in k.

Consider a suboptimal tax system $\{\tau^k, \tau^\ell\}$ implementing an allocation along the iso-revenue curve for g in Figure A.4. There are three possibilities for this allocation. This allocation is either in the segment between the optimum and the peak of the Laffer curve for τ^ℓ (point A in Figure A.4); or between the optimum and the peak of the Laffer curve for τ^k (point B in Figure A.4); or it is beyond the peak of the Laffer curve (meaning that k and ℓ are both too low, and both taxes are too high and they can both be decreased to increase revenue). The corollary assumes that the tax system is not beyond the peak of the Laffer curve.

At point A, capital is above the optimum and employment is below the optimum. Therefore,

$$U^\ell(k,\ell) > \mu^* > U^k(k,\ell),$$

where μ^* is the Lagrange multiplier at the optimum allocation. The inequality $U^\ell(k,\ell) > U^k(k,\ell)$ implies

$$\frac{f_\ell - \nu'(\ell)}{\dfrac{\nu''(\ell)}{1-\varrho}\cdot \ell + \dfrac{\nu'(\ell)}{1-\varrho} - f_\ell} > \frac{f_k + 1 - \delta - u'(\bar{y}-k)}{-u''(\bar{y}-k)\cdot k + u'(\bar{y}-k) - (f_k + 1 - \delta)}.$$

Dividing the numerator and the denominator on the left-hand side by $\nu'(\ell)/(1-\varrho)$, and the numerator and the denominator on the right-hand side by $u'(\bar{y}-k) - 1$, and using the definition of $\varepsilon^\ell(\ell)$ and $\varepsilon^k(k)$ yields (9).

Finally, we prove that τ^k and τ^ℓ satisfy $\tau^\ell > \tau^{\ell,r}$ and $\tau^{k,r} > \tau^k$. In particular, observe that the market-clearing condition for capital is

$$1 - \tau^k = \frac{u'(\bar{y}-k) - 1}{f_k - \delta}.$$

The numerator on the right-hand side increases with k, and the denominator decreases in k and increases in ℓ (this is due to the concavity of f by Lemma A.2 and the fact that f exhibits constant returns to scale). Thus, the right-hand side of this equation increases as we move from the optimal allocation to the current allocation, which implies $\tau^{k,r} > \tau^k$. Likewise,

$$1 - \tau^\ell = \frac{\nu'(\ell)}{(1-\varrho)\cdot f_\ell}.$$

The numerator on the right-hand side increases with ℓ, and the denominator decreases in ℓ and increases in k (this is due to the concavity of f by Lemma A.2 and the fact that f exhibits constant returns to scale). Therefore, the right-hand side of this equation decreases as we move from the optimal allocation to the current one, which implies $\tau^{\ell,r} < \tau^{\ell}$.

Conversely, the same argument implies that at point B the opposite of (9) holds and thus in this region $\tau^{\ell,r} > \tau^{\ell}$ and $\tau^{k,r} < \tau^{k}$. Hence, this region is ruled out by (9).

Therefore, (9) is a necessary and sufficient condition for the tax system to be biased against labor and in favor of capital (and to lead to an equilibrium with employment below the optimum and the capital stock above the optimum). ■

Proof of Proposition 2. We can write the equilibrium quantities of capital and labor as $k(\theta)$ and $\ell(\theta)$, which are implicitly determined by (4) and (5).

Differentiating (4) and (5), we obtain that after an infinitesimal change in θ, the change in employment and capital are given by the solution to the system of equations:

$$-f_{k\ell} \cdot \ell_\theta + \left(-\frac{u''(\bar{y}-k)}{1-\tau^k} - f_{kk}\right) \cdot k_\theta = f_{k\theta} \qquad \left(\frac{\nu''(\ell)}{(1-\varrho)\cdot(1-\tau^\ell)} - f_{\ell\ell}\right) \cdot \ell_\theta - f_{\ell k} \cdot k_\theta = f_{\ell\theta},$$

which has a unique solution given by

$$\ell_\theta = \frac{f_{\ell\theta} \cdot \left(-\frac{u''(\bar{y}-k)}{1-\tau^k} - f_{kk}\right) + f_{k\theta} \cdot f_{\ell k}}{\left(\frac{\nu''(\ell)}{(1-\varrho)\cdot(1-\tau^\ell)} - f_{\ell\ell}\right) \cdot \left(-\frac{u''(\bar{y}-k)}{1-\tau^k} - f_{kk}\right) - f_{k\ell} \cdot f_{\ell k}}$$

$$k_\theta = \frac{f_{k\theta} \cdot \left(\frac{\nu''(\ell)}{(1-\varrho)\cdot(1-\tau^\ell)} - f_{\ell\ell}\right) + f_{\ell\theta} \cdot f_{k\ell}}{\left(\frac{\nu''(\ell)}{(1-\varrho)\cdot(1-\tau^\ell)} - f_{\ell\ell}\right) \cdot \left(-\frac{u''(\bar{y}-k)}{1-\tau^k} - f_{kk}\right) - f_{k\ell} \cdot f_{\ell k}}$$

Note that $f_\theta(k, \ell; \theta)$ has constant returns to scale in k and ℓ. Moreover, at $\theta = \theta^m(k, \ell)$, we have $f_\theta = 0$. The Euler theorem implies that $kf_{\theta k} + \ell f_{\theta\ell} = 0$. Then $f_{\theta k} > 0 > f_{\theta\ell}$. A second application of Euler's theorem yields $kf_{kk} + \ell f_{k\ell} = 0$; and a third application gives $kf_{\ell k} + \ell f_{\ell\ell} = 0$. It follows that, at $\theta = \theta^m(k, \ell)$, the following identities hold

$$f_{\ell\theta} \cdot f_{kk} = f_{k\theta} \cdot f_{\ell k} \qquad f_{k\theta} \cdot f_{\ell\ell} = f_{\ell\theta} \cdot f_{k\ell} \qquad f_{\ell\ell} \cdot f_{kk} = f_{k\ell} \cdot f_{\ell k}.$$

Using these identities, we can simplify the formulae for ℓ_θ and k_θ above as

$$\ell_\theta = \frac{-f_{\ell\theta}\frac{u''(\bar{y}-k)}{1-\tau^k}}{\Lambda} < 0 \qquad k_\theta = \frac{f_{k\theta}\frac{\nu''(\ell)}{(1-\varrho)\cdot(1-\tau^\ell)}}{\Lambda} > 0,$$

where $\Lambda = -\frac{\nu''(\ell)}{(1-\varrho)\cdot(1-\tau^\ell)}\frac{u''(\bar{y}-k)}{1-\tau^k} - f_{kk}\frac{\nu''(\ell)}{(1-\varrho)\cdot(1-\tau^\ell)} + f_{\ell\ell}\frac{u''(\bar{y}-k)}{1-\tau^k} > 0$. This establishes that reducing θ on the margin below $\theta^m(k, \ell)$ will always result in an increase in employment and a reduction in capital.

To complete the proof of the proposition, note that welfare (inclusive of the value of public funds) is given by

$$\begin{aligned}\mathcal{W} =& f(k,\ell;\theta) + (1-\delta)k + u(\bar{y}-k) - \nu(\ell) \\ &+ \mu^* \cdot \left(f(k,\ell;\theta) + (1-\delta)k - u'(\bar{y}-k)\cdot k - \frac{\nu'(\ell)\cdot \ell}{1-\varrho}\right),\end{aligned}$$

where μ^* denotes the Lagrange multiplier at the optimum allocation. Therefore, following an infinitesimal change in θ, welfare changes by

$$\frac{d\mathcal{W}}{d\theta} = \mathcal{W}_\ell \cdot \ell_\theta + \mathcal{W}_k \cdot k_\theta + (1+\mu^*)\cdot f_\theta,$$

where $\mathcal{W}_\ell$ and $\mathcal{W}_k$ denote the changes in welfare arising from improvements in allocative efficiency and $(1+\mu^*)\cdot f_\theta$ accounts for changes in productive efficiency.

Suppose that the current tax system satisfies (9). Corollary 2 implies that $U^\ell(k,\ell) > \mu^* > U^k(k,\ell)$—that is, employment is too low and capital too high. It follows that

$$\begin{aligned}\mathcal{W}_\ell =& f_\ell - \nu'(\ell) + \mu^* \cdot \left(\frac{\nu''(\ell)}{1-\varrho}\cdot \ell + \frac{\nu'(\ell)}{1-\varrho} - f_\ell\right) > 0 \Leftrightarrow U^\ell(k,\ell) > \mu^* \\ \mathcal{W}_k =& f_k - \delta - u'(\bar{y}-k) + 1 + \mu^* \cdot (-u''(\bar{y}-k)\cdot k + u'(\bar{y}-k) - 1 - f_k + \delta) < 0 \Leftrightarrow U^\ell(k,\ell) < \mu^*,\end{aligned}$$

so that welfare increases as employment increases and capital is reduced.

Moreover, starting from $\theta = \theta^m(k,\ell)$, we have $f_\theta = 0$, $\ell_\theta > 0$ and $k_\theta > 0$. Therefore,

$$\frac{d\mathcal{W}}{d\theta} < 0,$$

and welfare (inclusive of the value of public funds) increases following an infinitesimal reduction in θ.

We now turn to the implications of a reduction in θ for output and for revenue. The change in net output at $\theta^m(k,\ell)$ is

$$\frac{d \text{ net output}}{d\theta} = f_\ell \cdot \ell_\theta + (f_k - \delta)\cdot k_\theta,$$

which can be written as

$$\frac{d \text{ net output}}{d\theta} = -\frac{f_\ell \cdot (f_k-\delta)\cdot f_{k\theta}}{\ell \cdot \Lambda}\left(-\frac{u''(\bar{y}-k)\cdot k}{u'(\bar{y}-k)-1} - \frac{\nu''(\ell)\cdot \ell}{\nu'(\ell)}\right).$$

Thus, an infinitesimal reduction in θ will also expand net output provided that $\varepsilon^\ell(\ell) > \varepsilon^k(k)$, as claimed in the proposition.

Finally, the change in revenue near $\theta^m(k,\ell)$ is

$$\frac{d\text{ revenue}}{d\theta} = \left(f_\ell - \frac{\nu'(\ell)}{1-\varrho} - \frac{\nu''(\ell)\cdot\ell}{1-\varrho}\right)\cdot \ell_\theta + \left(f_k - \delta - u'(\bar{y}-k) + 1 + u''(\bar{y}-k)\cdot k\right)\cdot k_\theta,$$

which can be written as

$$\frac{d\text{ revenue}}{d\theta} = \frac{\nu'(\ell)\cdot(u'(\bar{y}-k)-1)\cdot f_{k\theta}}{\ell\cdot(1-\varrho)\cdot\Lambda}\cdot\frac{\tau^k\cdot(1+\varepsilon^k)-\tau^\ell\cdot(1+\varepsilon^\ell)}{\varepsilon^k\cdot\varepsilon^\ell\cdot(1-\tau^k)\cdot(1-\tau^\ell)}.$$

Thus, an infinitesimal reduction in θ will also expand revenue if $\tau^\ell\cdot(1+\varepsilon^\ell(\ell)) > \tau^k\cdot(1+\varepsilon^k(k))$, as claimed (recall that $u'(\bar{y}-k) > 1$ by assumption). ■

Proof of Proposition 3. The constrained Ramsey problem can be written as

$$\max_{k,\ell,\theta} f(k,\ell;\theta) + (1-\delta)\cdot k + u(\bar{y}-k) - \nu(\ell)$$
$$\text{subject to: } g \le f(k,\ell;\theta) + (1-\delta)\cdot k - u'(\bar{y}-k)\cdot k - \frac{\nu'(\ell)\cdot\ell}{1-\varrho}$$
$$\nu'(\ell)\cdot\ell \le (1-\bar{\tau}^\ell)\cdot(1-\varrho)\cdot f_\ell\cdot\ell.$$

Let $\mu > 0$ and $\gamma^\ell \ge 0$ denote the multipliers on the IC constraint and the constraint on labor taxes, respectively. We assume throughout that the constraint on labor taxes binds, so that $\gamma^\ell > 0$.

The first-order condition with respect to capital is given by

$$f_k - \delta - u'(\bar{y}-k) + 1 - \mu\cdot(-u''(\bar{y}-k)\cdot k + u'(\bar{y}-k) - 1 - f_k + \delta) + \gamma^\ell\cdot(1-\bar{\tau}^\ell)\cdot(1-\varrho)\cdot f_{\ell k}\cdot\ell = 0.$$

Dividing by $u'(\bar{y}-k)-1$, using the capital market-clearing condition (4) to substitute for $f_k - \delta$, and rearranging yields (11).

Note next that the choice of θ^c maximizes the Lagrangean of the constrained Ramsey problem. Thus, we have

$$\theta^c = \arg\max_{\theta\in[0,1]} (1+\mu)\cdot f(k,\ell;\theta) + \gamma^\ell\cdot(1-\bar{\tau}^\ell)\cdot(1-\varrho)\cdot f_\ell(k,\ell;\theta)\cdot\ell.$$

Denote by $g(\theta)$ the right-hand of this equation. We now show that $g(\theta)$ is strictly decreasing for $\theta \ge \theta^m(k,\ell)$. To prove this, note that

$$f_{\ell\theta}(k,\ell;\theta) = \frac{1}{\lambda}f_\theta(k,\ell;\theta)\frac{1}{\ell}\cdot\left(\frac{y}{\ell}\right)^{\frac{1}{\lambda}-1}\cdot\left(\int_\theta^1 \psi^\ell(x)^{\lambda-1}dx\right)^{\frac{1}{\lambda}} - \frac{1}{\lambda}\psi^\ell(\theta)^{\lambda-1}\cdot\left(\frac{y}{\ell}\right)^{\frac{1}{\lambda}-1}\cdot\left(\int_\theta^1 \psi^\ell(x)^{\lambda-1}dx\right)^{\frac{1}{\lambda}},$$

which is negative for $\theta \geq \theta^m(k,\ell)$. Moreover, $f_\theta(k,\ell;\theta)$ is zero at $\theta^m(k,\ell)$ and negative for all $\theta > \theta^m(k,\ell)$. Therefore, $g(\theta)$ is strictly decreasing for $\theta \geq \theta^m(k,\ell)$ (note that if we did not have $\gamma^\ell > 0$, $g(\theta)$ could not be strictly decreasing at $\theta^m(k,\ell)$).

Finally, because $g(\theta)$ is strictly decreasing for $\theta \geq \theta^m(k,\ell)$, we must have $\theta^c < \theta^m(k,\ell)$ as claimed. ∎

Proof of Proposition 4. See next section. ∎

Additional Results

The next proposition provides four alternative ways of implementing the desired level of automation via taxes and subsidies, and the first of these is the scheme presented in Proposition 4 in the text.

PROPOSITION A.1 ***(Implementation of a reduction in θ via task-specific taxes and subsidies)*** *Consider any allocation $\{k_p, \ell_p, \theta_p\}$ that satisfies the implementability condition, and where $\theta_p \leq \theta^m(k_p,\ell_p)$. Let τ^k and τ^ℓ be given by*

$$1-\tau^k = \frac{u'(\bar{y}-k)-1}{f_k-\delta} \qquad\qquad 1-\tau^\ell = \frac{\nu'(\ell)}{(1-\varrho)\cdot f_\ell}.$$

Moreover, define

$$\tau^{A,gross} = 1 - \frac{f_k}{f_\ell}\frac{\psi^\ell(\theta_p)}{\psi^k(\theta_p)} \geq 0,$$

and define $\tau^A \geq 0$ implicitly as

$$\frac{1}{1-\tau^{A,gross}} = \frac{r}{r+\delta}\frac{1-\tau^k}{1-\tau^k-\tau^A} + \delta.$$

The allocation $\{k,\ell,\theta\}$ can be implemented in any of the following ways:

1. *A uniform tax τ^ℓ on labor and the following tax schedule ("automation tax") on capital:*

$$\tau^k(x) = \begin{cases} \tau^k & \text{for } x \leq \theta_p \\ \tau^k+\tau^A & \text{for } x > \theta_p \end{cases}$$

2. *A uniform tax τ^ℓ on labor, a uniform tax τ^k on net capital income, and a gross automation tax $\tau^{A,gross}$ on the use of capital at tasks $x > \theta_p$.*

3. *A uniform tax τ^ℓ on labor, a uniform net tax $\tau^k+\tau^A$ on capital, and a subsidy of $\tau^{A,gross}$ to tasks below θ_p.*

4. *A uniform tax/subsidy* $\frac{1-\tau^\ell}{1-\tau^A}$ *on labor, a uniform tax* τ^k *on capital, and a tax* $\tau^{A,gross}$ *on the output of tasks above* θ_p.

Proof. Consider the first implementation. We show that these taxes generate a competitive equilibrium with factor prices $R = f_k(k_p, \ell_p; \theta_p)$ and $w = f_\ell(k_p, \ell_p; \theta_p)$, and where all tasks below θ_p are produced by capital.

Take the factor prices $R = f_k(k_p, \ell_p; \theta_p)$ and $w = f_\ell(k_p, \ell_p; \theta_p)$ as given. The unit cost of producing task x with labor is

$$p^\ell(x) = \frac{w}{\psi^\ell(x)},$$

whereas the unit cost of producing task x with capital is

$$p^k(x) = \frac{R(x)}{\psi^k(x)},$$

where $R(x)$ is the pre-tax rental rate paid for the use of capital in task x. Because after-tax net returns must be equal across tasks, we have

$$(R(x) - \delta)(1 - \tau^k(x)) = (R - \delta) \cdot (1 - \tau^k)$$

for all x. The definitions of $\tau^k(x)$ and τ^A then imply

$$R(x) = \begin{cases} R & \text{if } x \leq \theta_p \\ \dfrac{R}{1 - \tau^{A,\text{gross}}} & \text{if } x > \theta_p, \end{cases}$$

and the unit cost of producing task x with capital becomes

$$p^k(x) = \begin{cases} \dfrac{R}{\psi^k(x)} & \text{if } x \leq \theta_p \\ \dfrac{R}{\psi^k(x) \cdot (1 - \tau^{A,\text{gross}})} & \text{if } x > \theta_p. \end{cases}$$

Because $\theta_p < \theta^m(k_p, \ell_p)$, we have that for all all tasks $x \in [0, \theta_p]$:

$$\frac{\psi^\ell(x)}{\psi^k(x)} < \frac{\psi^\ell(\theta^m)}{\psi^k(\theta^m)} = \frac{f_\ell}{f_k},$$

which implies

$$p^k(x) = \frac{f_k}{\psi^k(x)} < \frac{f_\ell}{\psi^\ell(x)} = p^\ell(x),$$

and all these tasks are produced by capital.

On the other hand, for all tasks $x \in (\theta_p, 1]$, we have

$$(1 - \tau^{A,\text{gross}}) \frac{f_\ell}{f_k} = \frac{\psi^\ell(\theta_p)}{\psi^k(\theta_p)} < \frac{\psi^\ell(x)}{\psi^k(x)},$$

which implies

$$p^\ell(x) = \frac{f_\ell}{\psi^\ell(x)} < \frac{f_k}{\psi^k(x) \cdot (1 - \tau^{A,\text{gross}})} = p^k(x),$$

and all these tasks are produced by labor.

We now compute the market-clearing conditions and show that markets clear at the stipulated factor prices. The market-clearing condition for capital is

$$k = \int_0^{\theta_p} k(x)dx = y \cdot \int_0^{\theta_p} \frac{p^k(x)^{-\lambda}}{\psi^k(x)} dx = y \cdot R^{-\lambda} \cdot \int_0^{\theta_p} \psi^k(x)^{\lambda-1} dx,$$

which holds with equality when $R = f_k$.

Likewise, the market-clearing condition for labor is

$$\ell = \int_{\theta_p}^1 \ell(x)dx = y \cdot \int_{\theta_p}^1 \frac{p^\ell(x)^{-\lambda}}{\psi^k(x)} dx = y \cdot w^{-\lambda} \cdot \int_{\theta_p}^1 \psi^\ell(x)^{\lambda-1} dx,$$

which holds with equality when $w = f_\ell$.

Finally, note that revenue remains as in equation (6) since capital is not used in tasks where it is subject to the higher automation tax.

The argument for the second implementation strategy is essentially identical, but with the difference that the gross tax on the use of capital directly implies that

$$R(x) = \begin{cases} R & \text{if } x \leq \theta_p \\ \frac{R}{1 - \tau^{A,\text{gross}}} & \text{if } x > \theta_p, \end{cases}$$

We now turn to the third implementation strategy. The definition of τ^A implies that the pre-tax gross return required by households is given by

$$R = \frac{\frac{u'(\bar{y}-k_p)-1}{1-\tau^k} + \delta}{1 - \tau^{A,\text{gross}}};$$

whereas the pre-tax wage required by households is given by

$$w = \frac{\nu'(\ell_p)}{(1-\varrho)\cdot(1-\tau^\ell)}.$$

The definition of τ^k implies

$$R = \frac{f_k}{1 - \tau^{A,\text{gross}}};$$

and the definition of τ^ℓ implies $w = f_\ell$.

We next show that at these factor prices, all tasks below θ_p are produced by capital and all tasks above θ_p are produced by labor. For $x \leq \theta_p$ we have

$$\frac{\psi^\ell(x)}{\psi^k(x)} \leq \frac{\psi^\ell(\theta_p)}{\psi^k(\theta_p)} = \frac{f_\ell \cdot (1 - \tau^{A,\text{gross}})}{f_k},$$

which implies

$$p^k(x) = \frac{f_k}{\psi^k(x)\cdot(1-\tau^{A,\text{gross}})} \leq \frac{f_\ell}{\psi^\ell(x)} = p^\ell(x),$$

and all these tasks are produced by capital.

On the other hand, for all tasks $x \in (\theta_p, 1]$, we have

$$\frac{\psi^\ell(x)}{\psi^k(x)} > \frac{\psi^\ell(\theta_p)}{\psi^k(\theta_p)} = \frac{f_\ell \cdot (1 - \tau^{A,\text{gross}})}{f_k},$$

which implies

$$p^\ell(x) = \frac{f_\ell}{\psi^\ell(x)} < \frac{f_k}{\psi^k(x)\cdot(1-\tau^{A,\text{gross}})} = p^k(x),$$

and all these tasks are produced by labor.

We now show that markets clear for $R = \frac{f_k}{1-\tau^{A,\text{gross}}}$ and $w = f_\ell$. The market-clearing condition for capital is

$$k = \int_0^{\theta_p} k(x)dx = y \cdot \int_0^{\theta_p} \frac{((1-\tau^{A,\text{gross}})\cdot p^k(x))^{-\lambda}}{\psi^k(x)} dx = y \cdot f_k^{-\lambda} \cdot \int_0^{\theta_p} \psi^k(x)^{\lambda-1} dx,$$

which holds with equality. Note that here we used the fact that all tasks below θ_p receive a subsidy of $\tau^{A,\text{gross}}$. Likewise, the market-clearing condition for labor is

$$\ell = \int_{\theta_p}^{1} \ell(x)dx = y \cdot \int_{\theta_p}^{1} \frac{p^{\ell}(x)^{-\lambda}}{\psi^{\ell}(x)}dx = y \cdot f_{\ell}^{-\lambda} \cdot \int_{\theta_p}^{1} \psi^{k}(x)^{\lambda-1}dx,$$

which holds with equality.

Finally, note that revenue remains as in equation (6) since the tax τ^A on capital raises revenue $\tau^{A,\text{gross}} f_k \cdot k$, but this coincides with the cost of subsidizing all tasks below θ_p at a rate $\tau^{A,\text{gross}}$, since the total value of these tasks is $f_k \cdot k$.

We conclude with the fourth implementation strategy. The pre-tax gross return required by households is given by

$$R = \frac{u'(\bar{y} - k_p) - 1}{1 - \tau^k} + \delta;$$

whereas the pre-tax wage required by households is given by

$$w = \frac{\nu'(\ell_p) \cdot (1 - \tau^{A,\text{gross}})}{(1 - \varrho) \cdot (1 - \tau^{\ell})}.$$

The definition of τ^k yields $R = f_k$, and from the definition of τ^{ℓ} we have

$$w = f_{\ell} \cdot (1 - \tau^{A,\text{gross}}).$$

We now show that at these factor prices, all tasks below θ_p are produced by capital and all tasks above θ_p are produced by labor. For $x \leq \theta_p$ we have

$$\frac{\psi^{\ell}(x)}{\psi^{k}(x)} \leq \frac{\psi^{\ell}(\theta_p)}{\psi^{k}(\theta_p)} = \frac{f_{\ell} \cdot (1 - \tau^{A,\text{gross}})}{f_k},$$

which implies

$$p^{k}(x) = \frac{f_k}{\psi^{k}(x)} \leq \frac{f_{\ell} \cdot (1 - \tau^{A,\text{gross}})}{\psi^{\ell}(x)} = p^{\ell}(x),$$

and all these tasks are produced by capital.

For all tasks $x \in (\theta_p, 1]$, we have

$$\frac{\psi^{\ell}(x)}{\psi^{k}(x)} > \frac{\psi^{\ell}(\theta_p)}{\psi^{k}(\theta_p)} = \frac{f_{\ell} \cdot (1 - \tau^{A,\text{gross}})}{f_k},$$

which implies

$$p^{\ell}(x) = \frac{f_{\ell} \cdot (1 - \tau^{A,\text{gross}})}{\psi^{\ell}(x)} < \frac{f_k}{\psi^{k}(x)} = p^{k}(x),$$

and all these tasks are produced by labor.

We now show that markets clear for $R = f_k$ and $w = f_\ell \cdot (1-\tau^\ell)$. The market-clearing condition for capital is

$$k = \int_0^{\theta_p} k(x)dx = y \cdot \int_0^{\theta_p} \frac{p^k(x)^{-\lambda}}{\psi^k(x)} dx = y \cdot f_k^{-\lambda} \cdot \int_0^{\theta_p} \psi^k(x)^{\lambda-1} dx,$$

which holds with equality. Likewise, the market-clearing condition for labor is

$$\ell = \int_{\theta_p}^1 \ell(x)dx = y \cdot \int_{\theta_p}^1 \frac{(p^\ell(x)/(1-\tau^{A,\text{gross}}))^{-\lambda}}{\psi^\ell(x)} dx = y \cdot f_\ell^{-\lambda} \cdot \int_{\theta_p}^1 \psi^k(x)^{\lambda-1} dx,$$

which holds with equality. Note that here we used the fact that all tasks above θ_p are taxed at the rate $\tau^{A,\text{gross}}$.

Finally, note that revenue remains as in equation (6) since the subsidy $\tau^{A,\text{gross}}$ on labor costs $\tau^{A,\text{gross}} f_\ell \cdot \ell$, but this coincides with the taxes raised on the production of all tasks above θ_p (since the total value of these tasks is $f_\ell \cdot \ell$). ■

The next proposition presents the analogue to Proposition 3, where there is an upper bound on capital taxes (rather than a lower bound on labor taxes).

Proposition A.2 ***(Excessive automation when capital taxes are constrained)*** *Consider the constrained Ramsey problem of maximizing (7) subject to the additional constraint $\tau^k \leq \bar{\tau}^k$, which is equivalent to*

$$u'(\bar{y} - k) - 1 \geq (1 - \bar{\tau}^k) \cdot (f_k - \delta), \tag{A.6}$$

and suppose that in the solution to this problem (A.6) binds and has multiplier $\gamma^k \cdot k > 0$. Suppose also that this multiplier satisfies

$$1 + \mu > \gamma^k \cdot (1 - \bar{\tau}^k),$$

so that an increase in capital income holding labor income constant is socially beneficial (see the proof).

Then the constrained optimal taxes and allocation are given by

- *a capital tax of $\tau^{k,c} = \bar{\tau}^k$ and a tax/subsidy on labor $\tau^{\ell,c}$ that satisfies*

$$\frac{\tau^{\ell,c}}{1-\tau^{\ell,c}} = \frac{\mu}{1+\mu} \frac{1}{\varepsilon^\ell(\ell)} - \frac{\varrho}{1+\mu} + \frac{\gamma^k}{1+\mu} \cdot (1 - \bar{\tau}^k) \cdot (1 - \varrho) \cdot \frac{f_{k\ell} \cdot k}{\nu'(\ell)}, \tag{A.7}$$

- *a level of automation $\theta^c < \theta^m(k, \ell)$.*

Moreover, the level of automation θ^c can be implemented through an extra subsidy to labor and a tax of the same magnitude on the output of tasks above θ^c (so that capital taxes still remain no greater than $\bar{\tau}^k$).

Proof. The constrained Ramsey problem can be written as

$$\max_{k,\ell,\theta} f(k,\ell;\theta) + (1-\delta)\cdot k + u(\bar{y}-k) - \nu(\ell)$$

$$\text{subject to: } g \le f(k,\ell;\theta) + (1-\delta)\cdot k - u'(\bar{y}-k)\cdot k - \frac{\nu'(\ell)\cdot \ell}{1-\varrho}$$

$$(1+(1-\bar{\tau}^k)\cdot(f_k-\delta))\cdot k \le u'(\bar{y}-k)\cdot k.$$

Let $\mu > 0$ and $\gamma^k \ge 0$ denote the multipliers on the IC constraint and the constraint on labor taxes, respectively. We assume throughout that the constraint on capital taxes binds, so that $\gamma^k > 0$.

The first-order condition with respect to labor is given by

$$f_\ell - \nu'(\ell) - \mu\cdot\left(\frac{\nu''(\ell)\cdot\ell}{1-\varrho} + \frac{\nu'(\ell)}{1-\varrho} - f_\ell\right) - \gamma^k\cdot(1-\bar{\tau}^k)\cdot f_{k\ell}\cdot k = 0.$$

Dividing by $\nu'(\ell)/(1-\varrho)$, using the labor market-clearing condition (5) to substitute for f_ℓ, and rearranging yields (A.7).

Note next that the choice of θ^c maximizes the Lagrangean of the constrained Ramsey problem. Therefore,

$$\theta^c = \arg\max_{\theta\in[0,1]} (1+\mu)\cdot f(k,\ell;\theta) - \gamma^k\cdot(1-\bar{\tau}^k)\cdot f_k(k,\ell;\theta)\cdot k.$$

Using the fact that f has constant returns to scale, we can rewrite this maximization problem as

$$\theta^c = \arg\max_{\theta\in[0,1]} \left(1+\mu-\gamma^k\cdot(1-\bar{\tau}^k)\right)\cdot f(k,\ell;\theta) + \gamma^k\cdot(1-\bar{\tau}^k)\cdot f_\ell(k,\ell;\theta)\cdot\ell.$$

Since, by assumption, $1+\mu-\gamma^k\cdot(1-\bar{\tau}^k) > 0$, the argument outlined in the proof of Proposition 3 can be applied to prove that the objective is strictly decreasing in θ for $\theta \ge \theta^m(k,\ell)$, so that $\theta^c < \theta^m(k,\ell)$. (Note that the inequality $1+\mu-\gamma^k\cdot(1-\bar{\tau}^k) \le 0$ implies that welfare would decline if capital income increased and labor income remained constant—i.e., an increase in f leaving $f_\ell\cdot\ell$ constant. Alternatively, our assumption implies that distortions are not too large, so that increases in income always raise welfare.)

The fact that this allocation can be implemented via a subsidy to labor and a tax on the production of tasks above θ^c follows from Proposition A.1. ■

Proofs of Extension Propositions in Section 5

Proof of Proposition 5. Define

$$\nu_h(\ell_h) := \min_{\ell,h} \frac{\ell^{1+1/\varepsilon^\ell}}{1+1/\varepsilon^\ell} + \frac{\ell \cdot h^{1+1/\varepsilon^h}}{1+1/\varepsilon^h} \text{ subject to: } h \cdot \ell \geq \ell_h$$

as the disutility of supplying ℓ_h efficiency units of labor.

The solution to this minimization problem satisfies

$$\nu_h(\ell_h) = \left(\frac{1}{1+1/\varepsilon^\ell} + \frac{1}{1+1/\varepsilon^h}\right) \cdot \ell_h^{1+1/(\varepsilon^\ell+\varepsilon^h+\varepsilon^\ell \cdot \varepsilon^h)}.$$

The Ramsey problem is analogous to the one studied in Proposition 1 but with $\nu_h(\ell_h)$ in place of $\nu(\ell)$ and ℓ_h in place of ℓ. Thus, the same formulae in equation (8) apply, but with $\varepsilon^\ell + \varepsilon^h + \varepsilon^\ell \cdot \varepsilon^h$ in place of $\varepsilon^\ell(\ell)$. ■

Proof of Proposition 6. The planner can undo the effects of the aggregate markup, κ, introduced by the technology sector by using a production subsidy to the final good sector, at the cost of $\kappa \cdot f$. With this subsidy in place, the market equilibrium is an allocation $\{k, \ell\}$, a level of automation adoption θ, and a state of automation technology Θ such that:

- the capital and labor market clear

$$(1-\tau^k) \cdot (f_k - \delta) = u'(\bar{y} - k) \qquad (1-\varrho) \cdot (1-\tau^\ell) \cdot f_\ell = \nu'(\ell);$$

- adoption decisions maximize output and are given by $\theta^m(k, \ell, \Theta)$ and $\omega^m(k, \ell, \Theta)$, where

$$\{\theta^m(k, \ell, \Theta), \omega^m(k, \ell, \Theta)\} = \arg\max_{G(\theta,\omega;\Theta) \leq 0} f(k, \ell; \theta, \omega);$$

- automation technology Θ maximizes monopolists' profits in (20).

Define

$$\tilde{\Theta}(k, \ell, \kappa) = \arg\max_{\Theta} \kappa \cdot f(k, \ell; \theta^m(k, \ell, \Theta), \omega^m(k, \ell, \Theta)) - \Gamma(\Theta),$$

which determines the optimal choice of technology given k, ℓ and some profit rate κ. Also, let $k(\Theta)$ and $\ell(\Theta)$ denote the level of capital and labor resulting in the market equilibrium when the bias of technology is Θ. The market equilibrium is characterized by a bias of technology Θ^m such that

$$\Theta^m = \tilde{\Theta}(k(\Theta^m), \ell(\Theta^m); \kappa),$$

which, by assumption, exists and is uniquely defined for every $\kappa \in (0,1)$. Moreover, because we assumed that the equilibrium is unique and that $\tilde{\Theta} > 0$, we have that in this equilibrium the curve $\tilde{\Theta}(k(\Theta), \ell(\Theta); \kappa)$ cuts the 45 degree line Θ^m from above.

We now turn to the Ramsey problem. To derive an IC constraint, we start from the government budget constraint, which in this context is given by

$$g + \kappa \cdot f(k, \ell; \theta, \omega) \leq \tau^k \cdot (f_k - \delta) \cdot k + \tau^\ell \cdot f_\ell \cdot \ell + \kappa \cdot f(k, \ell; \theta) - \Gamma(\Theta).$$

Here, the term $\kappa \cdot f(k, \ell; \theta, \omega)$ on the left-hand side accounts for the subsidy on production required to undo markups. The term $\kappa \cdot f(k, \ell; \theta, \omega) - \Gamma(\Theta)$ accounts for profits the taxation of profits in the technology sector, which we have assumed the government can (and will) fully tax.

Using the market-clearing conditions for capital and labor and the fact that f has constant returns to scale, we can rewrite the IC in terms of the allocation as

$$g \leq f(k, \ell; \theta, \omega) + (1 - \delta)k - u'(\bar{y} - k) \cdot k - \frac{\nu'(\ell) \cdot \ell}{1 - \varrho} - \Gamma(\Theta).$$

Thus, the Ramsey problem can be expressed as

$$\max_{k, \ell, \theta, \omega, \Theta} \; f(k, \ell; \theta, \omega) + (1 - \delta) \cdot k + u(\bar{y} - k) - \nu(\ell) - \Gamma(\Theta)$$

$$\text{subject to: } g \leq f(k, \ell; \theta, \omega) + (1 - \delta) \cdot k - u'(\bar{y} - k) \cdot k - \frac{\nu'(\ell) \cdot \ell}{1 - \varrho} - \Gamma(\Theta)$$

$$G(\theta, \omega; \Theta) \leq 0.$$

It follows that the optimal choice of k and ℓ is identical to the one in Proposition 1, and therefore optimal taxes on capital and labor are given by equation (8). Likewise, optimal adoption decisions maximize output subject to $G(\theta, \omega; \Theta)$, and are therefore given by $\theta^m(k, \ell, \Theta)$ and $\omega^m(k, \ell, \Theta)$.

Turning to the optimal bias of technology, it is straightforward to see that, given an allocation for capital and employment, the optimal bias of technology maximizes

$$\max_{\Theta} f(k, \ell; \theta^m(k, \ell, \Theta), \omega^m(k, \ell, \Theta)) - \Gamma(\Theta).$$

It follows that the Ramsey solution involves a bias of technology given by Θ^r, which is the (unique) solution to

$$\Theta^r = \tilde{\Theta}(k(\Theta^r), \ell(\Theta^r); 1).$$

The claim in the proposition is equivalent to $\Theta^r < \Theta^m$ if $\Theta^r < \bar{\Theta}$—that is, the state of automation technology is too high relative to the Ramsey solution when the status quo is above the Ramsey solution. On the other hand, $\Theta^r > \Theta^m$ if $\Theta^r > \bar{\Theta}$—that is, the state of automation technology is too low relative to the Ramsey solution when the status quo is below the Ramsey solution. (And $\Theta^r = \Theta^m$ if $\Theta^r = \bar{\Theta}$).

To establish this result, denote the resulting output when technology is Θ by $F(k, \ell; \Theta) = f(k, \ell; \theta^m(k, \ell, \Theta), \omega^m(k, \ell, \Theta))$ and suppose that $\Theta^r < \bar{\Theta}$. Because for $\Theta < \bar{\Theta}$, $\Gamma(\Theta)$ is decreasing, we have that the maximization problem defining $\tilde{\Theta}(k(\Theta^r), \ell(\Theta^r); \kappa)$ can be rewritten as

$$\max_{\Theta} F(k, \ell; \Theta) - \frac{\Gamma(\Theta)}{\kappa},$$

which has decreasing differences in κ and θ. Thus, as we move from the optimal allocation (which obtains when $\kappa = 1$) to the market equilibrium with $\kappa < 1$, we get

$$\tilde{\Theta}(k(\Theta^r), \ell(\Theta^r); \kappa) > \tilde{\Theta}(k(\Theta^r), \ell(\Theta^r); 1) = \Theta^r.$$

This implies that the unique competitive equilibrium lies to the right of Θ^r as claimed in the proposition (recall that in this equilibrium, the curve $\tilde{\Theta}(k(\Theta), \ell(\Theta); \kappa)$ must cut the 45 degree line at a unique point Θ^m from above).

Suppose next that $\Theta^r > \bar{\Theta}$. Because for $\Theta > \bar{\Theta}$, $\Gamma(\Theta)$ is increasing, we now have that the maximization problem defining $\tilde{\Theta}(k(\Theta^r), \ell(\Theta^r); \kappa)$ can be rewritten as

$$\max_{\Theta} F(k, \ell; \Theta) - \frac{\Gamma(\Theta)}{\kappa},$$

which has increasing differences in κ and θ. Thus, as we move from the optimal allocation (which obtains when $\kappa = 1$) to the market equilibrium with $\kappa < 1$, we get

$$\tilde{\Theta}(k(\Theta^r), \ell(\Theta^r); \kappa) < \tilde{\Theta}(k(\Theta^r), \ell(\Theta^r); 1) = \Theta^r.$$

This implies that the unique competitive equilibrium lies to the left of Θ^r as claimed in the proposition (recall that in this equilibrium, the curve $\tilde{\Theta}(k(\Theta), \ell(\Theta); \kappa)$ must cut the 45 degree line at a unique point Θ^m from above). Finally, these two arguments together imply that $\Theta^r = \Theta^m$ if $\Theta^r = \bar{\Theta}$, completing the proof of the proposition. ■

A.3 Infinite-Horizon Model

This section presents an infinite-horizon version of our model and derives two results. The first one shows that, in the presence of a labor market wedge, if long-run capital taxes converge to zero, labor should be subsidized in order to (completely) undo this wedge. The second one shows that if there is an upper bound to the government budget (for example, for political economy reasons), which implies that the government cannot accumulate asset, then both capital and labor taxes converge to finite values and these values depend on the supply elasticities of these factors as in Proposition 1 in the text.

Environment

As in the text, we work with a representative household economy. Preferences over sequences of consumption and work $\{(c_0, \ell_0), (c_1, \ell_1), \dots\}$ are defined recursively as

$$V_t = \mathcal{W}(u(c_t, \ell_t), V_{t+1}). \tag{A.8}$$

From this recursion, we can compute lifetime utility as a function of the time-paths of consumption and labor as

$$V_0 = \mathcal{V}(u(c_0, \ell_0), u(c_1, \ell_1), \dots).$$

The aggregator $\mathcal{W}$ satisfies the following properties:

(W1) $\mathcal{W}$ is a continuous and increasing function from $\mathbb{R}^2$ to $\mathbb{R}$.

(W2) Its partial derivative with respect to V satisfies $\mathcal{W}_V \in (0, 1)$.

(W3) Denote by $\mathcal{V}^N(u_0, u_1, \dots, u_{N-1}; y)$ the value of receiving stage utility u_t for $t = 0, \dots, N-1$ and a continuation value of y at time N. The aggregator $\mathcal{W}$ satisfies that, for all $N \geq 1$ and y, the function $\mathcal{V}^N(u_0, u_1, \dots, u_{N-1}; y)$ is concave in $\{u_0, \dots u_{N-1}\}$.

The stage utility function $u(c, \ell)$ satisfies

$$\frac{u_{cc}}{u_c} - \frac{u_{\ell c}}{u_\ell} \leq 0 \qquad \frac{u_{c\ell}}{u_c} - \frac{u_{\ell\ell}}{u_\ell} \leq 0. \tag{A.9}$$

These two assumptions imply that consumption and leisure are normal goods. They are satisfied when u is quasi-linear as in the main text and as we impose for our second main result in this Appendix.

The notation in this section follows Straub & Werning (2020). We denote the derivative of X with respect to z at time t by X_{zt}. Also, it will be useful to define $\beta_t = \prod_{s=0}^{t-1} \mathcal{W}_{Vs}$. With this notation, the derivatives of $\mathcal{V}$ are

$$\mathcal{V}_{ct} = \beta_t \cdot \mathcal{W}_{ut} \cdot u_{ct} \qquad\qquad \mathcal{V}_{\ell t} = \beta_t \cdot \mathcal{W}_{ut} \cdot u_{\ell t}.$$

We also use M_t to denote the marginal rate of substitution between consumption in periods $t-1$ and t, given by

$$M_t = \frac{\mathcal{V}_{ct-1}}{\mathcal{V}_{ct}} = \frac{1}{\mathcal{W}_{Vt-1}} \frac{\mathcal{W}_{ut-1}}{\mathcal{W}_{ut}} \frac{u_{ct-1}}{u_{ct}}.$$

Consider a constant path of consumption and labor yielding stage utility u and generating lifetime utility V. Let us then define the function $\bar{M}(V) = 1/\mathcal{W}_V(u, V) \in (1, \infty)$, where u satisfies $V = \mathcal{W}(u, V)$. When preferences are time-separable, we have $V = u + \beta \cdot V$ and $\bar{M}(V) = 1/\beta$. However, when preferences are not time-separable, we have $\bar{M}'(V) \neq 0$.

Starting from a given $k_0 > 0$, and given sequences of effective taxes on capital and labor $\{\tau_t^k\}$ and $\{\tau_t^\ell\}$, a competitive equilibrium is given by a sequence of consumption, labor, capital, and automation levels, $\{(c_0, \ell_0, k_0, \theta_0), (c_1, \ell_1, k_1, \theta_1), \ldots\}$, such that:

- production is given by $y_t = f(k_t, \ell_t; \theta_t)$, where $\theta_t = \theta^m(k_t, \ell_t)$;
- the representative household's Euler equation holds:

(A.10) $$M_t = 1 + (f_{kt} - \delta) \cdot (1 - \tau_t^k);$$

- the labor market clears:

(A.11) $$-\frac{u_{\ell t}}{u_{ct}} = f_{\ell t} \cdot (1 - \varrho) \cdot (1 - \tau_t^\ell);$$

- the resource constraint holds:

(A.12) $$c_t + k_{t+1} + g \leq f(k_t, \ell_t; \theta_t) + (1 - \delta) k_t.$$

Optimal Policy with an Intertemporal Government Budget

Optimal policy maximizes V_0 subject to the recursion (A.8), the Euler equation (A.10), the labor market-clearing condition (A.11), the resource constraint (A.12) and a government budget restriction.

We first study an intertemporal budget restriction of the form

$$0 \leq \sum_{t=0}^{\infty} \mathcal{V}_{ct} \cdot (\tau_t^k \cdot (f_{kt} - \delta) \cdot k_t + \tau_t^\ell \cdot f_{\ell t} \cdot \ell_t - g).$$

The assumption here is that government can issue debt or accumulate assets that yield a return equal to $\mathcal{V}_{ct-1}/\mathcal{V}_{ct}$, which is the gross rate of return required by the representative household. We will also study a different version of this problem where the government must keep a balanced budget every period.

Following Straub & Werning (2020), the Ramsey problem boils down to choosing a sequence of consumption, labor, capital, and automation, $\{(c_0, \ell_0, k_0, \theta_0), (c_1, \ell_1, k_1, \theta_1), \ldots\}$ that maximizes V_0 subject to the recursion (A.8), the resource constraint (A.12), and an *Implementability Constraint* (IC) that ensures that the taxes needed to implement that allocation are sufficient to cover government expenditure:

$$\mathcal{V}_{c0} \cdot M_0 \cdot k_0 \leq \sum_{t=0}^{\infty} \left(\mathcal{V}_{ct} \cdot c_t + \mathcal{V}_{\ell t} \frac{\ell_t}{1 - \varrho} \right). \tag{A.13}$$

As is common in these problems, we assume that τ_t^k is bounded from above, so that the government cannot expropriate the entire capital stock at time 0 to satisfy the IC. In particular, following Straub & Werning (2020), we assume that capital taxation is constrained and one most have $M_0 \geq 1$.

Our first proposition shows that, as in our static model, when optimal (unconstrained) taxes are in place, the planner will not distort automation decisions.

Proposition A.3 *Suppose taxes are unconstrained. The solution to the Ramsey problem always involves setting* $\theta_t^r = \theta^m(k_t, \ell_t)$.

Proof. In this problem, θ_t only appears in the term $f(k_t, \ell_t; \theta_t)$ in the resource constraint (A.12). Thus, the optimal θ maximizes $f(k_t, \ell_t; \theta_t)$ and coincides with $\theta^m(k_t, \ell_t)$. ■

Our second result in this Appendix, presented in the next proposition, generalizes Proposition 6 in Straub & Werning (2020) to the case with labor market imperfections.

Proposition A.4 *Suppose that the Ramsey problem yields a solution where the resulting allocation converges to an interior steady state with non-zero private wealth and optimal taxes* $\tau^{k,r}$ *and* $\tau^{\ell,r}$. *If* $\bar{M}'(V) \neq 0$, *optimal policy in the long run involves a zero tax on capital and a subsidy to labor that corrects for the labor market distortion introduced by* ϱ, *i.e.,* $\tau^{k,r} = 0$ *and* $\tau^{\ell,r} = 1 - 1/(1 - \varrho)$.

PROOF. Exploiting the recursive formulation of preferences, we can write the Ramsey problem as maximizing V_0 subject to $V_t = \mathcal{W}(u(c_t, \ell_t), V_{t+1})$, (A.12) and the IC constraint in equation (A.13), which can be rewritten as

$$\mathcal{W}_{u0} \cdot u_{c0} \cdot M_0 \cdot k_0 \leq \sum_{t=0}^{\infty} \beta_t \cdot \mathcal{W}_{ut} \cdot \left(u_{ct} \cdot c_t + \frac{u_{\ell t} \cdot \ell_t}{1-\varrho} \right). \tag{A.14}$$

Using the same notation as in Straub & Werning (2020), let us define

$$A_{t+1} = \frac{1}{\beta_{t+1}} \frac{\partial}{\partial V_{t+1}} \sum_{s=0}^{\infty} \beta_s \cdot \mathcal{W}_{us} \cdot \left(u_{cs} \cdot c_s + \frac{u_{\ell s} \cdot \ell_s}{1-\varrho} \right) \quad B_t = \frac{1}{\beta_t} \sum_{s=0}^{\infty} \frac{\partial(\beta_s \cdot \mathcal{W}_{us})}{\partial u_t} \cdot \left(u_{cs} \cdot c_s + \frac{u_{\ell s} \cdot \ell_s}{1-\varrho} \right).$$

Because these objects depend only on allocations, asymptotically they converge to limiting values, which we denote by A^{st} and B^{st}. The same holds for all the derivatives of components of the utility function or the production function with respect to changes in the allocation. In what follows, we use the superscript st to denote steady-state values.

Moreover, as shown in Straub & Werning (2020), A^{st} satisfies

$$A^{\text{st}} = -\frac{\bar{M}'(V^{\text{st}})}{\bar{M}(V^{\text{st}})} \cdot W_u^{\text{st}} \cdot u_c^{\text{st}} \cdot (1 + (f_k^{\text{st}} - \delta) \cdot (1 - \tau^{k,r})) \cdot a^{\text{st}},$$

where $a^{\text{st}} \neq 0$ (by assumption) is the representative household's wealth. Therefore, when $\bar{M}'(V^{\text{st}}) \neq 0$, we have $A^{\text{st}} \neq 0$.

Denote by $\beta_t \cdot \eta_t$ the multiplier on $V_t = \mathcal{W}(u(c_t, \ell_t), V_{t+1})$; by $\beta_t \cdot \vartheta_t$ the multiplier on the resource constraint (A.12); and μ the multiplier on the Implementability Constraint, IC, (A.14). We can write the limit of the first-order conditions for the Ramsey problem as

$$\begin{aligned}
-\eta_t + \eta_{t+1} + \mu \cdot A^{\text{st}} &= 0 \\
-\eta_t \cdot W_u^{\text{st}} \cdot u_c^{\text{st}} + \mu \cdot W_u^{\text{st}} \cdot \left(u_c^{\text{st}} + u_{cc}^{\text{st}} \cdot c^{\text{st}} + \frac{u_{\ell c}^{\text{st}} \cdot \ell^{\text{st}}}{1-\varrho} \right) + \mu \cdot B^{\text{st}} \cdot u_c^{\text{st}} &= \vartheta_t \\
\eta_t \cdot W_u^{\text{st}} \cdot u_\ell^{\text{st}} - \mu \cdot W_u^{\text{st}} \cdot \left(u_{c\ell}^{\text{st}} \cdot c^{\text{st}} + \frac{u_\ell^{\text{st}}}{1-\varrho} + \frac{u_{\ell\ell}^{\text{st}} \cdot \ell^{\text{st}}}{1-\varrho} \right) - \mu \cdot B^{\text{st}} \cdot u_\ell^{\text{st}} &= \vartheta_t \cdot f_\ell^{\text{st}} \\
-\vartheta_t + \vartheta_{t+1} \cdot W_V^{\text{st}} \cdot (1 + f_k^{\text{st}} - \delta) &= 0
\end{aligned}$$

Subtracting the second equation at time $t+1$ from the same equation at time t, and substituting $-\eta_t + \eta_{t+1}$ from the first equation, we obtain

$$\vartheta_t - \vartheta_{t+1} = -W_u^{\text{st}} \cdot u_c^{\text{st}} \cdot \mu \cdot A^{\text{st}}. \tag{A.15}$$

Likewise, eliminating η_t from the first-order conditions for consumption and capital (the second and third first-order conditions above), we obtain

$$(\text{A.16}) \quad \vartheta_t \cdot (f_\ell^{\text{st}} \cdot u_c^{\text{st}} + u_\ell^{\text{st}}) = \mu \cdot W_u^{\text{st}} \cdot u_c^{\text{st}} \cdot u_\ell^{\text{st}} \cdot \left(-\frac{\varrho}{1-\varrho} + \left(\frac{u_{cc}^{\text{st}}}{u_c^{\text{st}}} - \frac{u_{c\ell}^{\text{st}}}{u_\ell^{\text{st}}} \right) \cdot c^{\text{st}} + \left(\frac{u_{\ell c}^{\text{st}}}{u_c^{\text{st}}} - \frac{u_{\ell\ell}^{\text{st}}}{u_\ell^{\text{st}}} \right) \frac{\ell^{\text{st}}}{1-\varrho} \right).$$

Equation (A.9) ensures that the term in brackets is strictly negative.

We next use (A.15) and (A.16) to prove the claims in the proposition.

Suppose first that $\mu = 0$. Then equation (A.15) implies that $\vartheta_t = \vartheta_{t+1}$. The first-order condition for capital (the fourth equation of the block) then gives

$$1 + f_k^{\text{st}} - \delta = \frac{1}{W_V^{\text{st}}} = \bar{M}(V^{\text{st}}),$$

which is equivalent to having zero taxes on capital. Likewise, equation (A.16) yields $f_\ell^{\text{st}} \cdot u_c^{\text{st}} + u_\ell^{\text{st}} = 0$. From equation (A.11), this is only possible if $1 - \tau^{\ell,r} = 1/(1-\varrho)$, or in other words if there is a labor subsidy fully offsetting the distortion introduced by ϱ. Thus, when $\mu = 0$, the desired result is established.

Now suppose that $\mu \neq 0$. Then equation (A.15) implies that ϑ_t diverges to $-\infty$ or ∞ (recall that $\mu \cdot A^{\text{st}} \neq 0$). In this case, (A.16) requires that $f_\ell^{\text{st}} \cdot u_c^{\text{st}} + u_\ell^{\text{st}}$ converge to zero. This again implies from equation (A.11) that $1 - \tau^{\ell,r} = 1/(1-\varrho)$, as desired. Likewise, the first-order condition for capital (the fourth equation of the block) implies that

$$1 + f_k^{\text{st}} - \delta = \frac{\vartheta_t}{\vartheta_{t+1}} \bar{M}(V^{\text{st}}).$$

Because ϑ_t is an arithmetic series, the right-hand side in this equation converges to $\bar{M}(V^{\text{st}})$, which implies a zero tax on capital. ■

This proposition implies that when capital is not perfectly elastic (that is, $\bar{M}'(V) \neq 0$) and the government can build as much of a positive asset position as it likes, optimal policy involves zero capital taxation and a subsidy to labor in the long run, financed by (relatively heavy) taxation of capital and labor along the transition.

Optimal Policy with a Balanced Budget

While Proposition A.4 is conceptually interesting, the government building a very large asset position is unrealistic for various reasons. Most importantly, political economy considerations would make it infeasible for the government to have a huge surplus and accumulate vast amounts of assets. In this part of the Appendix, we explore the implications of limiting the ability of the government to build vast asset positions. To do this in the simplest possible way, we impose a balanced budget for the government in each period, so that its budget constraint now becomes:[39]

$$g \le \tau_t^k \cdot (f_{kt} - \delta) \cdot k_t + \tau_t^\ell \cdot f_{\ell t} \cdot \ell_t.$$

With these series of budget constraints, the Ramsey problem is now to maximize V_0 subject to the recursion in (A.8), the resource constraint (A.12), and the series of IC constraints

$$g \le f(k_t, \ell_t; \theta_t) + (1-\delta)k_t - M_t \cdot k_t - \frac{\nu'(\ell_t) \cdot \ell_t}{1-\varrho}. \tag{A.17}$$

This IC is very similar to that in our static model, with the only difference that the intertemporal marginal rate of substitution is now M_t (rather than $u'(\bar{y} - k)$ as in the static model).

To simplify the analysis and maximize the similarity with our static model, we now assume that the stage utility function takes a quasi-linear form: $u(c, \ell) = c - \nu(\ell)$ (see the next section of the Appendix for the implications of more general preferences). Finally, throughout this section we assume that, for a given path of future consumption and labor $\{c_{t+s}, \ell_{t+s}\}_{s=0}^{\infty}$, the intertemporal marginal rate of substitution M_t is decreasing in c_{t-1}. In economic terms, this requirement makes intuitive sense and holds even for the usual time-additive separable aggregator $\mathcal{W}(u, v) = u + \beta v$. This additional assumption ensures that the solution to the savings problem faced by households has a well defined limit, with assets converging to a fixed amount that could be infinite (see the Turnpike and Monotonicity Theorems in Section 4 of Becker and Boyd, 1993).[40]

Before providing our characterization of optimal policy in this environment, it is useful to define the relevant capital and labor supply elasticities that will play a key role in shaping optimal policy. We define the *Hicksian elasticity of capital supply* as the percent increase in savings of a given household in response to a compensated change in net capital taxes. This is analogous to the standard definition of the Hicksian elasticity of labor supply. Consider a household that faces a constant after-tax net return $r \cdot (1 - \tau^k)$ and an after-tax wage rate $w \cdot (1 - \tau^\ell)$. In addition, the household receives a government transfer T, so that household

[39] More generally, we may impose the constraint that the government's assets should not exceed a certain amount. In that case, a similar constraint would apply with g denoting the expenditures that cannot be covered by interest payments on the long-run assets of the government. See also the next section.

[40] A necessary and sufficient condition for this is that $\mathcal{W}_{uu} \cdot \mathcal{W}_V - \mathcal{W}_{uV} \cdot \mathcal{W}_u < 0$. Property W3 of aggregators introduced above implies that $\mathcal{W}_{uu} \le 0$. Thus, all aggregators with $\mathcal{W}_{uV} \ge 0$ (including the usual time-additive separable aggregator $\mathcal{W}(u, v) = u + \beta v$) satisfy this property.

consumption is given by $r \cdot (1-\tau^k) \cdot k + w \cdot (1-\tau^\ell) \cdot \ell + T$. The long-run choice of capital and labor by this household converges to some level k^{st} and ℓ^{st} pinned down by the optimality conditions:

$$1 + r^{\text{st}} \cdot (1-\tau^{k,\text{st}}) = \bar{M}(V^{\text{st}}) \qquad\qquad w^{\text{st}} \cdot (1-\tau^{\ell,\text{st}}) = \nu'(\ell^{\text{st}})$$

where in addition, the utility level V is a fixed point of (A.8):

$$V = \mathcal{W}(r \cdot (1-\tau^k) \cdot k + w \cdot (1-\tau^\ell) \cdot \ell + T - \nu(\ell), V).$$

The Hicksian elasticity of capital supply is given by the change in k following a permanent increase in τ^k, where households get a rebate of $dT = r \cdot k \cdot d\tau^k$. This is the transfer required to compensate households for the change in after-tax returns, so that if the household in question did not change its plans, it would achieve the exact same utility as before. The optimality condition for capital implies

$$r \cdot (1-\tau^k) \cdot d\ln(1-\tau^k) = \bar{M}'(V) \cdot dV.$$

Moreover, the definition of V implies

$$(1-\mathcal{W}_V) \cdot dV = \mathcal{W}_u \cdot \left(r \cdot (1-\tau^k) \cdot dk - r \cdot k \cdot d\tau^k + dT\right) = \mathcal{W}_u \cdot r \cdot (1-\tau^k) \cdot dk.$$

These two equations together imply that the Hicksian elasticity of capital is

$$\varepsilon^k = \frac{d\ln k}{d\ln(1-\tau^k)} = \frac{1-\mathcal{W}_V}{\mathcal{W}_u \cdot \bar{M}'(V) \cdot k}.$$

On the other hand, the Hicksian elasticity of labor supply is given by the change in ℓ following a change in τ^ℓ. Using the optimality condition for labor, we obtain a Hicksian elasticity given by

$$\varepsilon^\ell = \frac{\partial \ln \ell}{\partial \ln(1-\tau^\ell)} = \frac{\nu'(\ell)}{\nu''(\ell) \cdot \ell}.$$

Because the stage utility function is quasi-linear, this elasticity is independent of whether the tax change is compensated or not. We are now in a position to state and prove our second main result in this Appendix.

PROPOSITION A.5 *Consider the Ramsey problem of maximizing V_0 subject to the recursion in (A.8), the resource constraint (A.12), and the sequence of ICs in (A.17).*

- *Optimal policy leaves automation undistorted at $\theta_t^r = \theta^m(k_t, \ell_t)$.*

- *If the optimal allocation converges, optimal taxes are given by*

$$\text{(A.18)} \qquad \frac{\tau^{k,r}}{1-\tau^{k,r}} = \frac{\tilde{\mu}^{st}}{1+\tilde{\mu}^{st}}\frac{1}{\varepsilon^k} + \mathcal{O}(\tau^{k,r^2}) \qquad \frac{\tau^{\ell,r}}{1-\tau^{\ell,r}} = \frac{\tilde{\mu}^{st}}{1+\tilde{\mu}^{st}}\frac{1}{\varepsilon^\ell} - \frac{1}{1+\tilde{\mu}^{st}}\varrho,$$

where $\tilde{\mu}^{st} > 0$ which gives the long-run social value of government funds. Moreover, if $\varepsilon^k = \infty$ (or $\bar{M}'(V) = 0$), we have $\tau^{k,r} = 0$; whereas if $\varepsilon^k \in (0,\infty)$ (or $\bar{M}'(V) > 0$), we have $\tau^{k,r} > 0$.

Proof. The first part of the proposition—that $\theta^r_t = \theta^m(k_t,\ell_t)$—follows from the fact that θ_t only shows up in the term $f(k_t,\ell_t;\theta_t)$ in the resource constraint (A.12) and the right-hand side of the ICs in (A.17). Thus, the optimal choice of θ maximizes $f(k_t,\ell_t;\theta_t)$ and coincides with $\theta^m(k_t,\ell_t)$.

The rest of the proof establishes the second part of the proposition. We write M_t as a function of V_{t-1}, V_t, and V_{t+1}. This can be done without any loss of generality, since the recursive formulation of preferences implies

$$M_t = \frac{1}{\mathcal{W}_V(u_{t-1},V_t)}\frac{\mathcal{W}_u(u_{t-1},V_t)}{\mathcal{W}_u(u_t,V_{t+1})}.$$

In addition, u_{t-1} and u_t can be obtained implicitly as functions of V_{t-1}, V_t, and V_{t+1} using (A.8). Thus, we write $M_t = M(V_{t-1},V_t,V_{t+1})$, and denote the partial derivatives of M_t with respect to V_{t-1},V_t,V_{t+1} by M_{1t}, M_{2t}, and M_{3t}, respectively. These definitions imply $\bar{M}(V) = M(V,V,V)$.

Denote by $\beta_t \cdot \eta_t$ the multiplier on $V_t = \mathcal{W}(u(c_t,\ell_t),V_{t+1})$; by $\beta_t \cdot \vartheta_t$ the multiplier on the resource constraint (A.12); and $\beta_t \cdot \mu_t$ the multiplier on the IC in (A.17).

The first-order condition for consumption is:

$$\text{(A.19)} \qquad \vartheta_t = \eta_t \cdot W_{ut};$$

and the first-order condition for V_t is given by:

$$\eta_{t-1} = \eta_t + M_{1t+1} \cdot \mathcal{W}_{Vt} \cdot \mu_{t+1} \cdot k_{t+1} + M_{2t} \cdot \mu_t \cdot k_t + M_{3t-1} \cdot \frac{1}{\mathcal{W}_{Vt-1}} \cdot \mu_{t-1} \cdot k_{t-1}.$$

Combining these two equations yields a single first-order condition for consumption:

$$\text{(A.20)} \quad \frac{1}{\mathcal{W}_{Vt-1}}\vartheta_{t-1} = M_t \cdot \vartheta_t + M_t \cdot \mathcal{W}_{ut} \cdot \left(M_{1t+1} \cdot \mathcal{W}_{Vt} \cdot \mu_{t+1} \cdot k_{t+1} + M_{2t} \cdot \mu_t \cdot k_t + M_{3t-1} \cdot \frac{1}{\mathcal{W}_{Vt-1}} \cdot \mu_{t-1} \cdot k_{t-1}\right).$$

The first-order condition for labor is:

$$\eta_t \cdot W_{ut} \cdot \nu'(\ell_t) = \vartheta_t \cdot f_{\ell t} + \mu_t \cdot \left(f_{\ell t} - \frac{\nu'(\ell_t)}{1-\varrho} - \frac{\nu''(\ell_t) \cdot \ell_t}{1-\varrho} \right),$$

which can be combined with the first-order condition for consumption:

$$0 = \vartheta_t \cdot (f_{\ell t} - \nu'(\ell_t)) + \mu_t \cdot \left(f_{\ell t} - \frac{\nu'(\ell_t)}{1-\varrho} - \frac{\nu''(\ell_t) \cdot \ell_t}{1-\varrho} \right). \tag{A.21}$$

Finally, the first-order condition for capital is given by

$$\frac{1}{\mathcal{W}_{Vt-1}} \vartheta_{t-1} = \vartheta_t \cdot (f_{kt} + 1 - \delta) + \mu_t \cdot (f_{kt} + 1 - \delta - M_t). \tag{A.22}$$

Suppose that the optimal allocation converges, as assumed in the proposition. In what follows, we again use the superscript *st* to denote the steady-state value of different quantities. As before, because they only depend on allocations, the derivatives of the preference aggregator converge to $\mathcal{W}_{Vt} \to \mathcal{W}_V^{\mathrm{st}}$ and $\mathcal{W}_{ut} \to \mathcal{W}_u^{\mathrm{st}}$, and the derivatives of the marginal rate of substitution M also converge to $M_{1t+1} \to M_1^{\mathrm{st}}$, $M_{2t} \to M_2^{\mathrm{st}}$ and $M_{3t-1} \to M_3^{\mathrm{st}}$.

The first-order condition for labor in equation (A.21) implies that $\mu_t/\vartheta_t \to \tilde{\mu}^{\mathrm{st}}$, where $\tilde{\mu}^{\mathrm{st}}$ denotes the steady-state value of government funds. Moreover, because both of these multipliers are non-negative, we have $\tilde{\mu}^{\mathrm{st}} \geq 0$. The first-order condition for capital in equation (A.22) then implies that ϑ_t follows a geometric progression with $\vartheta_{t-1} = q^{\mathrm{st}} \cdot \vartheta_t$. Because $\mu_t = \tilde{\mu}^{\mathrm{st}} \cdot \vartheta_t$ and $\eta_t = \vartheta_t / \mathcal{W}_u$, these multipliers also follow geometric progressions with $\mu_{t-1} = q^{\mathrm{st}} \cdot \mu_t$ and $\eta_{t-1} = q^{\mathrm{st}} \cdot \eta_t$.

The steady state can be computed as the unique solution for $\tilde{\mu}^{\mathrm{st}}, q^{\mathrm{st}}, k^{\mathrm{st}}, c^{\mathrm{st}}, \ell^{\mathrm{st}}$ and V^{st} to the following system of equations:

$$\begin{aligned}
q^{\mathrm{st}} &= 1 + \tilde{\mu}^{\mathrm{st}} \cdot \mathcal{W}_u^{\mathrm{st}} \cdot \left(M_1^{\mathrm{st}} \cdot \frac{1}{\bar{M}^{\mathrm{st}} \cdot q^{\mathrm{st}}} + M_2^{\mathrm{st}} + M_3^{\mathrm{st}} \cdot \bar{M}^{\mathrm{st}} \cdot q^{\mathrm{st}} \right) \cdot k^{\mathrm{st}} \\
\bar{M}^{\mathrm{st}} \cdot q^{\mathrm{st}} &= f_k^{\mathrm{st}} + 1 - \delta + \tilde{\mu}^{\mathrm{st}} \cdot \left(f_k^{\mathrm{st}} + 1 - \delta - \bar{M}^{\mathrm{st}} \right) \\
0 &= f_\ell^{\mathrm{st}} - \nu'^{\mathrm{st}}) + \tilde{\mu}^{\mathrm{st}} \cdot \left(f_\ell - \frac{\nu'^{\mathrm{st}})}{1-\varrho} - \frac{\nu''(\ell^{\mathrm{st}}) \cdot \ell^{\mathrm{st}}}{1-\varrho} \right) \\
c^{\mathrm{st}} + g &= f(k^{\mathrm{st}}, \ell^{\mathrm{st}}, \theta^m(k^{\mathrm{st}}, \ell^{\mathrm{st}})) - \delta \cdot k^{\mathrm{st}} \\
g &= f(k^{\mathrm{st}}, \ell^{\mathrm{st}}, \theta^m(k^{\mathrm{st}}, \ell^{\mathrm{st}})) + (1-\delta) \cdot k^{\mathrm{st}} - \bar{M}^{\mathrm{st}} \cdot k^{\mathrm{st}} - \frac{\nu'(\ell^{\mathrm{st}}) \cdot \ell^{\mathrm{st}}}{1-\varrho} \\
V^{\mathrm{st}} &= \mathcal{W}(c^{\mathrm{st}} - \nu(\ell^{\mathrm{st}}), V^{\mathrm{st}}) \\
q^{\mathrm{st}} &\geq 1/\bar{M}^{\mathrm{st}}
\end{aligned}$$

These equations correspond to the limits of the first-order conditions in (A.20), (A.21) and (A.22); and the limits of the resource constraint in equation (A.12), the implementability condition in equation (A.17), and the recursive definition of utility in equation (A.8). Finally, the inequality $q^{\text{st}} \geq 1/\bar{M}^{\text{st}}$ is equivalent to the transversality condition.[41]

We now characterize the solution to this system of equations.

First, we show that, so long as $g > 0$, we must have $\tilde{\mu}^{\text{st}} > 0$. Suppose to obtain a contradiction that $\tilde{\mu}^{\text{st}} = 0$. The first equation of the block implies that $q^{\text{st}} = 1$ and the second equation implies that $\bar{M}^{\text{st}} = f_k^{\text{st}} + 1 - \delta$. The third equation of the block implies $\nu'(\ell^{\text{st}}) = f_\ell^{\text{st}}$. Thus, if $\tilde{\mu}^{\text{st}} = 0$, the steady-state allocation coincides with the first best. However, implementing the first-best allocation generates negative revenue for the government (as it has to subsidize labor and cannot tax capital), and so the IC cannot hold. To see this formally, multiply $\bar{M}^{\text{st}} = f_k^{\text{st}} + 1 - \delta$ by k^{st} and $\nu'(\ell^{\text{st}}) = f_\ell^{\text{st}}$ by ℓ^{st} and add these two equations to obtain

$$0 = f(k^{\text{st}}, \ell^{\text{st}}; \theta^m(k^{\text{st}}, \ell^{\text{st}})) + (1-\delta) \cdot k^{\text{st}} - \bar{M}^{\text{st}} \cdot k^{\text{st}} - \nu'(\ell^{\text{st}}) \cdot \ell^{\text{st}},$$

where we used the fact that $f(k^{\text{st}}, \ell^{\text{st}}; \theta^m(k^{\text{st}}, \ell^{\text{st}})) = f_k^{\text{st}} \cdot k^{\text{st}} + f_\ell^{\text{st}} \cdot \ell^{\text{st}}$. When $g > 0$, this equality implies that

$$\begin{aligned} g &> f(k^{\text{st}}, \ell^{\text{st}}, \theta^m(k^{\text{st}}, \ell^{\text{st}})) + (1-\delta) \cdot k^{\text{st}} - \bar{M}^{\text{st}} \cdot k^{\text{st}} - \nu'(\ell^{\text{st}}) \cdot \ell^{\text{st}} \\ &\geq f(k^{\text{st}}, \ell^{\text{st}}, \theta^m(k^{\text{st}}, \ell^{\text{st}})) + (1-\delta) \cdot k^{\text{st}} - \bar{M}^{\text{st}} \cdot k^{\text{st}} - \frac{\nu'(\ell^{\text{st}}) \cdot \ell^{\text{st}}}{1-\varrho}, \end{aligned}$$

which contradicts the IC constraint.

Second, we show that for any $\tilde{\mu}^{\text{st}} > 0$ and a given allocation, the first equation of the above block has a unique solution q^{st} such that $q^{\text{st}} \geq 1/\bar{M}^{\text{st}}$. Moreover, this solution satisfies that $q^{\text{st}} > 1$ if $\bar{M}'(V) > 0$ for all V, and $q^{\text{st}} = 1$ if $\bar{M}'(V^{\text{st}}) = 0$.

To show this, write the first equation of the block as $q^{\text{st}} = 1 + \mu^{\text{st}} \cdot \mathcal{W}_u^{\text{st}} \cdot h(q^{\text{st}}) \cdot k^{\text{st}}$, where

$$h(q) = M_1^{\text{st}} \frac{1}{\bar{M}^{\text{st}} \cdot q} + M_2^{\text{st}} + M_3^{\text{st}} \cdot \bar{M}^{\text{st}} \cdot q.$$

For $q^{\text{st}} \in [1/\bar{M}^{\text{st}}, 1]$, the function $h(q^{\text{st}})$ has an inverted U-shape with minima at the extremes, where $h(1/\bar{M}^{\text{st}}) = h(1) = \bar{M}'(V^{\text{st}}) > 0$. The fact that $h(1) = \bar{M}'(V^{\text{st}})$ follows from the

[41] The transversality condition of the Ramsey problem is $x_t = \beta_t \cdot \eta_t \cdot k^{\text{st}} \to 0$, which requires that

$$1 \leq \lim_{t\to\infty} \frac{x_{t-1}}{x_t} = \bar{M}^{\text{st}} \cdot q^{\text{st}},$$

and is thus equivalent to $q^{\text{st}} \geq 1/\bar{M}^{\text{st}}$.

observation that $M_3^{\rm st} \cdot \bar{M}^{\rm st} = M_1^{\rm st}$, and thus $M_1^{\rm st}/\bar{M}^{\rm st} + M_3^{\rm st} \cdot \bar{M}^{\rm st} = M_3^{\rm st} + M_1^{\rm st}$.[42] The fact that $h(q^{\rm st})$ has an inverted U-shape in this range follows from the fact that $h(q) = a/q + b + dq$, with $a, d \leq 0$.[43]

Suppose that $\bar{M}'(V) > 0$. If $q^{\rm st} \in [1/\bar{M}^{\rm st}, 1]$, we have $h(q^{\rm st}) \geq \bar{M}'(V^{\rm st}) > 0$, and the first equation of the block implies $q^{\rm st} = 1 + \tilde{\mu}^{\rm st} \cdot \mathcal{W}_u^{\rm st} \cdot h(q^{\rm st}) \cdot k^{\rm st} \geq 1 + \tilde{\mu}^{\rm st} \cdot \mathcal{W}_u^{\rm st} \cdot \bar{M}'(V^{\rm st}) \cdot k^{\rm st} > 1$, which contradicts the assumption that $q^{\rm st} \in [1/\bar{M}^{\rm st}, 1]$. Thus, any solution to the above system of equations must have $q^{\rm st} \geq 1$. We now show that there is a unique q that solves $q^{\rm st} = 1 + \tilde{\mu}^{\rm st} \cdot \mathcal{W}_u^{\rm st} \cdot h(q^{\rm st}) \cdot k^{\rm st}$, with $q^{\rm st} \geq 1$. At $q^{\rm st} = 1$, we have $q^{\rm st} < 1 + \tilde{\mu}^{\rm st} \cdot \mathcal{W}_u \cdot f(q^{\rm st}) \cdot k^{\rm st}$. However, as $q^{\rm st}$ increases, the left-hand side of this equation increases without bound and the right-hand side declines, which implies a unique solution $q^{\rm st}$ with $q^{\rm st} > 1$.

Suppose next that $\bar{M}'(V^{\rm st}) = 0$, then $q^{\rm st} = 1$ gives the unique solution with $q^{\rm st} > 1/\bar{M}^{\rm st}$ to the first equation of the above block.

Rearranging the first-order conditions for capital (the second equation of the block), and using the Euler equation in (A.10) to substitute for the marginal product of capital in terms of capital taxes, we obtain optimal capital taxes as

$$\frac{\tau^{k,r}}{1 - \tau^{k,r}} = \frac{1}{1 + \tilde{\mu}^{\rm st}} \frac{1}{1 - \mathcal{W}_V^{\rm st}} (q^{\rm st} - 1). \tag{A.23}$$

This equation implies that, if $\bar{M}'(V^{\rm st}) = 0$ so that $q^{\rm st} = 1$, we will have $\tau^{k,r} = 0$. However, if $\bar{M}'(V) > 0$ so that $q^{\rm st} > 1$, we have $\tau^{k,r} > 0$ as claimed.

Furthermore, we can approximate the optimal tax on capital as follows. A first-order

[42] Define the function $g(V_t, V_{t+1})$ implicitly by $V_t = \mathcal{W}(g(V_t, V_{t+1}))$. Denote the derivatives of g with respect to its arguments by $g_1 = \frac{\partial g}{\partial V_t}$ and $g_2 = \frac{\partial g}{\partial V_{t+1}}$. The definition of g implies that $1 = \mathcal{W}_u \cdot g_1$ and $-\mathcal{W}_V = \mathcal{W}_u \cdot g_2$. Turning to the definition of M, we have

$$M(V_{t-1}, V_t, V_{t+1}) = \frac{1}{\mathcal{W}_V(g(V_{t-1}, V_t), V_t)} \frac{\mathcal{W}_u(g(V_{t-1}, V_t), V_t)}{\mathcal{W}_u(g(V_t, V_{t+1}), V_{t+1})}.$$

It follows that

$$M_1^{\rm st} = \lim_{t\to\infty} \frac{\partial M}{\partial V_{t-1}} = -\frac{\mathcal{W}_{Vu}^{\rm st} \cdot g_1^{\rm st}}{\mathcal{W}_V^{\rm st2}} + \frac{\mathcal{W}_{uu}^{\rm st} \cdot g_1^{\rm st}}{\mathcal{W}_V^{\rm st} \cdot \mathcal{W}_u^{\rm st}} = \frac{\mathcal{W}_{uu}^{\rm st} \cdot \mathcal{W}_V^{\rm st} - \mathcal{W}_{Vu}^{\rm st} \cdot \mathcal{W}_u^{\rm st}}{\mathcal{W}_V^{\rm st2} \cdot \mathcal{W}_u^{\rm st2}},$$

and

$$M_3^{\rm st} = \lim_{t\to\infty} \frac{\partial M}{\partial V_{t+1}} = -\frac{\mathcal{W}_{uu}^{\rm st} \cdot \mathcal{W}_u^{\rm st} \cdot g_2^{\rm st} + \mathcal{W}_{uV}^{\rm st} \cdot \mathcal{W}_u^{\rm st}}{\mathcal{W}_V^{\rm st} \cdot \mathcal{W}_u^{\rm st2}} = \frac{\mathcal{W}_{uu}^{\rm st} \cdot \mathcal{W}_V^{\rm st} - \mathcal{W}_{uV}^{\rm st} \cdot \mathcal{W}_u^{\rm st}}{\mathcal{W}_V^{\rm st} \cdot \mathcal{W}_u^{\rm st2}}.$$

Dividing these formulae, we obtain $M_3^{\rm st}/M_1^{\rm st} = \mathcal{W}_V^{\rm st}$, which implies $M_3^{\rm st} \cdot \bar{M}^{\rm st} = M_1^{\rm st}$.

[43] In our system, $a = M_1^{\rm st}/\bar{M}^{\rm st}$ and $b = M_3^{\rm st} \cdot \bar{M}^{\rm st}$. The claim that $a, d \leq 0$ is thus equivalent to $M_1^{\rm st}, M_3^{\rm st} \leq 0$ which holds if and only if M_t is decreasing in c_{t-1}, as we assumed was the case. To show this formally, note that an increase in c_{t-1} holding constant consumption at all other dates is equivalent to an increase in V_{t-1} holding V_t and V_{t+1} constant. Thus, M_t is decreasing in c_{t-1} if and only if $M_{1t} \leq 0$. Because $M_3^{\rm st} \cdot \bar{M}^{\rm st} = M_1^{\rm st}$, we have $M_3^{\rm st}, M_1^{\rm st} \leq 0$.

Taylor expansion of the first equation of the block gives

$$q^{\text{st}} - 1 = \tilde{\mu}^{\text{st}} \cdot \mathcal{W}_u^{\text{st}} \cdot \bar{M}'(V^{\text{st}}) \cdot k^{\text{st}} + (q^{\text{st}} - 1) \cdot \tilde{\mu}^{\text{st}} \cdot \mathcal{W}_u^{\text{st}} \cdot h'(1) + \mathcal{O}((q^{\text{st}} - 1)^2).$$

This equation implies

$$\tilde{\mu}^{\text{st}} = D \cdot (q^{\text{st}} - 1) + \mathcal{O}((q^{\text{st}} - 1)^2),$$

for some constant D, and we can therefore write

$$q^{\text{st}} - 1 = \tilde{\mu}^{\text{st}} \cdot \mathcal{W}_u^{\text{st}} \cdot \bar{M}'(V^{\text{st}}) \cdot k^{\text{st}} + \mathcal{O}((q^{\text{st}} - 1)^2).$$

Substituting this expression into equation (A.23) yields

$$\frac{\tau^{k,r}}{1 - \tau^{k,r}} = \frac{\tilde{\mu}^{\text{st}}}{1 + \tilde{\mu}^{\text{st}}} \frac{\mathcal{W}_u^{\text{st}} \cdot \bar{M}'(V^{\text{st}}) \cdot k^{\text{st}}}{1 - \mathcal{W}_V^{\text{st}}} + \mathcal{O}((q^{\text{st}} - 1)^2).$$

The formula in the proposition follows from the fact that $\mathcal{O}((q^{\text{st}} - 1)^2) = \mathcal{O}({\tau^{k,r}}^2)$, which is a direct implication of equation (A.23).

Finally, the optimal tax on labor, we can combine the first-order condition for labor (the third equation of the block) and the labor market-clearing condition in equation (A.11) to write optimal taxes on labor as

$$\frac{\tau^{\ell}}{1 - \tau^{\ell}} = \frac{\tilde{\mu}^{\text{st}}}{1 + \tilde{\mu}^{\text{st}}} \cdot \frac{\nu''(\ell^{\text{st}}) \cdot \ell^{\text{st}}}{\nu'(\ell^{\text{st}})} - \frac{1}{1 + \tilde{\mu}^{\text{st}}} \varrho,$$

which completes the proof. ■

The proposition above deals with long-run taxes and explores what happens when the allocation converges. We next establish that capital and labor taxes away from the long-run limit are given by similar expressions.

Proposition A.6 *Consider the Ramsey problem of maximizing V_0 subject to the recursion in (A.8), the resource constraint (A.12), and the sequence of Implementability Constraints in (A.17). Optimal policy involves taking $\theta_t^r = \theta^m(k_t, \ell_t)$, and for $t \geq 1$, setting taxes on capital and labor of*

$$\text{(A.24)} \quad \frac{\tau_t^k}{1 - \tau_t^k} = \frac{\tilde{\mu}_t}{1 + \tilde{\mu}_t} \frac{M_t}{M_t - 1} \mathcal{W}_{ut} \cdot \left(M_{1t+1} \cdot \mathcal{W}_{Vt} \frac{\mu_{t+1}}{\mu_t} k_{t+1} + M_{2t} \cdot k_t + M_{3t-1} \frac{1}{\mathcal{W}_{Vt-1}} \frac{\mu_{t-1}}{\mu_t} k_{t-1} \right)$$

$$\text{(A.25)} \quad \frac{\tau_t^{\ell}}{1 - \tau_t^{\ell}} = \frac{\tilde{\mu}_t}{1 + \tilde{\mu}_t} \frac{1}{\varepsilon^{\ell}(\ell_t)} - \frac{1}{1 + \tilde{\mu}_t} \varrho,$$

where $\tilde{\mu}_t$ gives the value of government funds in terms of units of the consumption good.

Proof. Combining equations (A.20) and (A.22), we obtain

$$(\vartheta_t + \mu_t) \cdot (f_{kt} + 1 - \delta - M_t) =$$
$$\mu_t \cdot M_t \cdot \mathcal{W}_{ut} \cdot \left(M_{1t+1} \cdot \mathcal{W}_{Vt} \frac{\mu_{t+1}}{\mu_t} k_{t+1} + M_{2t} \cdot k_t + M_{3t-1} \frac{1}{\mathcal{W}_{Vt-1}} \frac{\mu_{t-1}}{\mu_t} k_{t-1} \right).$$

Dividing both sides by $M_t - 1$ and using the Euler equation (A.10) to substitute for the marginal product of capital in terms of taxes, we obtain the formula in the Proposition.

Turning to labor, we can rewrite (A.21) as

$$\vartheta_t \cdot (f_{\ell t} - \nu'(\ell_t)) + \mu_t \cdot \left(f_{\ell t} - \frac{\nu'(\ell_t)}{1-\varrho} \right) = \frac{\nu''(\ell_t) \cdot \ell_t}{1-\varrho}.$$

Dividing both sides by $\nu'(\ell_t)/(1-\varrho)$ and using the labor market-clearing condition in equation (A.11) to substitute for the marginal product of labor in terms of taxes, we obtain the formula in the Proposition. ■

Proposition A.5 provides an approximation for optimal capital taxes in terms of Hicksian elasticities, and Proposition A.6 shows that a similar formula applies along the transition. We now show that for some commonly used preferences, the approximation in Proposition A.5 is exact and holds along the transition as well, thus providing an exact analog to the results in Proposition 1.

Corollary A.1 *If preferences are generated by an Epstein–Hynes aggregator of the form*

$$\mathcal{W}(c - \nu(\ell), V) = (-1 + V) \cdot \exp(-c + \nu(\ell)),$$

or by a Koopmans–Diamond–Williamson aggregator of the form

$$\mathcal{W}(c - \nu(\ell), V) = \frac{1}{\theta} \ln \left(1 + \beta(c - \nu(\ell)) + \delta V\right),$$

the optimal policy sets $\theta_t^r = \theta^m(k_t, \ell_t)$, and for $t \geq 1$, capital and labor taxes are given by

$$\frac{\tau_t^{k,r}}{1-\tau_t^{k,r}} = \frac{\tilde{\mu}_t}{1+\tilde{\mu}_t} \frac{1}{\varepsilon_t^k} \qquad \frac{\tau_t^{\ell,r}}{1-\tau_t^{\ell,r}} = \frac{\tilde{\mu}_t}{1+\tilde{\mu}_t} \frac{1}{\varepsilon_t^\ell} - \frac{1}{1+\tilde{\mu}_t} \varrho.$$

Proof. For the Epstein–Hynes preferences, we have

$$M_t = 1 - \frac{1}{V_t},$$

which implies $M_{1t+1} = M_{3t-1} = 0$.

For the Koopmans–Diamond–Williamson preferences, we have

$$M_t = \frac{\theta}{\delta} \exp(\theta \cdot V_t),$$

which implies that $M_{1t+1} = M_{3t-1} = 0$.

The result follows from the formulae in Proposition A.6 by setting $M_{1t+1} = M_{3t-1} = 0$. ■

A.4 Comparison to Atkinson–Stiglitz

In their seminal contribution, Atkinson & Stiglitz (1972) established several principles of efficient commodity taxation. One implication of these principles is that, if consumption and labor supply decisions are separable, then all commodities should face an homogeneous tax, which in a context with multiple periods would imply no taxes on capital.

This section explains why Propositions 1 and A.5 deviate from this paradigm, and provide formulae where optimal taxes on capital are linked to the elasticity of capital supply—how responsive savings are to changes in returns.

The key difference is that in Atkinson–Stiglitz the government is free to transfer resources across periods, whereas the key assumption behind our optimal tax formulae is that the government must run a balanced budget.

Proposition A.4 already showed that, if the government is allowed to accumulate assets, optimal policy dictates that the government accumulates enough assets to finance all of its expenditures out of interest income, reaching the first-best allocation in the long run. Here, optimal taxes on capital converge to zero independently of how elastic its supply is. The contrast between Propositions A.4 and A.5 thus underscores the importance of restricting the government to run a balanced budget.

We now return to a general version of our two-period model to elaborate on this point and explain the connection of our results to those of Atkinson & Stiglitz (1972).

A.4.1 A General Two-Period Model

The economy operates for two periods, $t = 0$ and $t = 1$. As in the main text, we use the subscript 0 for variables in period 0 and no subscripts for variables in period 1. Households are endowed with k_0 units of capital and face labor and capital taxes in each period. The decide how much labor to supply and how many resources to save in order to maximize their utility:

$$\begin{aligned} & \max u(c_0, c, \ell_0, \ell) \\ \text{subject to: } & c_0 \leq (1-\tau_0^\ell) \cdot w_0 \cdot \ell_0 + (1 + (1-\tau_0^k) \cdot (R_0 - \delta))k_0 - a^h \\ & c \leq \cdot (1-\tau^\ell) \cdot w \cdot \ell + (1 + (1-\tau^k) \cdot (R-\delta)) \cdot a^h, \end{aligned}$$

where a^h are assets saved by households in period 0.

The government faces the following budget constraints:

$$
\begin{aligned}
g_0 \leq &\tau_0^\ell \cdot w_0 \cdot \ell_0 + \tau_0^k \cdot (R_0 - \delta) \cdot k_0 - a^g \\
g \leq &\tau^\ell \cdot w \cdot \ell + \tau^k \cdot (R - \delta) \cdot k + (1 + (1 - \tau^k) \cdot (R - \delta)) \cdot a^g,
\end{aligned}
$$

where, analogously to the household side, a^g are assets saved by the government. When there are no restrictions on a^g, the two budget constraints can be combined into a single intertemporal constraint.

For simplicity, we work with a generic production function given by $y_0 = f_0(k_0, \ell_0)$ in period 0 and $y = f(k, \ell)$ in period 1. We also simplify the notation by setting $\varrho = 0$, so that labor market frictions, which are not important in the following analysis, are removed (this has no effect on any of the analysis except for simplifying some of our expressions).

A competitive equilibrium is given by an allocation $\{c_0, c, \ell_0, \ell, k, a^h, a^g\}$ and factor prices $\{w_0, w, R_0, R\}$ such that:

- the Euler equation of households holds:

$$
\frac{u_{c0}}{u_c} = (1 + (R - \delta) \cdot (1 - \tau^k));
$$

- the supply of labor satisfies:

$$
-\frac{u_{\ell 0}}{u_{c0}} = w_0 \cdot (1 - \tau_0^\ell) \qquad\qquad -\frac{u_{\ell 1}}{u_c} = w \cdot (1 - \tau^\ell);
$$

- factor prices are given by

$$
w_0 = f_{\ell 0} \qquad\qquad w = f_\ell \qquad\qquad R = f_k;
$$

- the market for capital clears:

$$
k = a^g + a^h.
$$

- the resource constraint at time 0 and 1 holds :

$$
c_0 + g_0 + k \leq f_0(k_0, \ell_0) + (1 - \delta) \cdot k_0 \qquad\qquad c + g \leq f(k, \ell) + (1 - \delta) \cdot k.
$$

The model used in the main text is a particular case of this one where we imposed the following simplifications:

- government must run a balanced budget, and so $a^g = 0$;
- in period 0, $f_0(k_0, \ell_0) = k_0$ and $\bar{y} = (1-\delta)\cdot k_0$; and in period 1 $y = f(k, \ell, \theta^m(k, \ell))$;
- quasi-linear preferences in c.

A.4.2 Implementability Conditions

As in Section A.2, we transform the government budget constraints into a series of implementability conditions. The implementability condition at time 0 becomes

$$g_0 \leq f_0(k_0, \ell_0) + (1-\delta)\cdot k_0 + \frac{u_{\ell 0}}{u_{c0}}\ell_0 - (1 + (1-\tau_0^k)\cdot(f_{k0} - \delta))\cdot k_0 - a_g,$$

which can be combined with the resource constraint at time 0 into

$$u_{c0}\cdot(1 + (1-\tau_0^k)\cdot(f_{k0}-\delta))\cdot k_0 \leq u_{c0}\cdot c_0 + u_{\ell 0}\cdot \ell_0 + u_{c0}\cdot(k - a_g).$$

The implementability condition at time 1 becomes

$$g \leq f(k, \ell) + (1-\delta)\cdot k + \frac{u_\ell}{u_c}\ell - \frac{u_{c0}}{u_c}\cdot(k - a^g),$$

which can be combined with the resource constraint at time 0 into

$$0 \leq u_c \cdot c + u_\ell \cdot \ell - u_{c0}\cdot(k - a^g).$$

Because k_0 is given, taxes on capital income at time 1 are lump sum.

Below we will consider two different scenarios. In the first scenario, a^g is unconstrained and we can combine both implementability conditions into a single one:

$$u_{c0}\cdot(1 + (1-\tau_0^k)\cdot(f_{k0}-\delta))\cdot k_0 \leq u_{c0}\cdot c_0 + u_{\ell 0}\cdot \ell_0 + u_c\cdot c + u_\ell \cdot \ell. \tag{A.26}$$

Alternatively, we could have a scenario where we restrict $a^g \leq 0$ and this restriction binds. This implies that only the first period IC constraint binds and can be written as

$$g \leq f(k, \ell) + (1-\delta)\cdot k + \frac{u_\ell}{u_c}\ell - \frac{u_{c0}}{u_c}\cdot k, \tag{A.27}$$

which coincides with the IC constraint used in the main text.

A.4.3 The Atkinson–Stiglitz Theorem

The following is the version of the Atkinson–Stiglitz theorem that applies in our economy.

PROPOSITION A.7 (ATKINSON–STIGLITZ) *Suppose that utility is separable in consumption and leisure and homothetic in c_0 and c. If the government can tax all capital income at time 0 and accumulate assets in an unrestricted way, the optimal tax on capital income in period 1 is zero.*

PROOF. The government will expropriate all capital income at period zero and ensure that $(1 + (1 - \bar{\tau}) \cdot (f_{k0} - \delta)) \cdot k_0 = 0$.

The assumptions on the utility function imply that we can write utility as

$$u(G(c_0, c), \ell_0, \ell)$$

for some homogeneous of degree 1 function G.

The Ramsey problem becomes

$$\begin{aligned} \max \quad & u(G(c_0, c), \ell_0, \ell) \\ \text{subject to: } & c_0 + g_0 + k \leq f_0(k_0, \ell_0) + (1 - \delta) \cdot k_0 \\ & c + g \leq f(k, \ell; \theta) + (1 - \delta) \cdot k \\ & 0 \leq u_G \cdot G_{c0} \cdot c_0 + u_G \cdot G_c \cdot c + u_{\ell 0} \cdot \ell_0 + u_\ell \cdot \ell. \end{aligned}$$

Denote by μ the multiplier on the IC constraint, by ϑ_0 the multiplier on the resource constraint at time 0, and by ϑ the multiplier on the IC constraint at time 1.

The first-order condition for capital is

$$\frac{\vartheta_0}{\vartheta} = 1 + f_k - \delta.$$

The first-order conditions for c_0 and c are given by:

$$\begin{aligned} \vartheta_0 =& (1 + \mu) \cdot u_G \cdot G_{c0} + \mu \cdot u_G \cdot (G_{c0c0} \cdot c_0 + G_{cc0} \cdot c) + \mu \frac{\partial (u_G \cdot G_{c0} \cdot c_0 + u_G \cdot G_c \cdot c + u_{\ell 0} \ell_0 + u_\ell \ell)}{\partial G} G_{c0}, \\ \vartheta =& (1 + \mu) \cdot u_G \cdot G_c + \mu \cdot u_G \cdot (G_{c0c} \cdot c_0 + G_{cc} \cdot c) + \mu \frac{\partial (u_G \cdot G_{c0} \cdot c_0 + u_G \cdot G_c \cdot c + u_{\ell 0} \ell_0 + u_\ell \ell)}{\partial G} G_c. \end{aligned}$$

Using the fact that G is homogeneous of degree 1, Euler's theorem implies $0 = G_{c0c0} \cdot c_0 + G_{cc0} \cdot c$ and $0 = G_{c0c} \cdot c_0 + G_{cc} \cdot c$. Dividing the first-order conditions for c_0 and c and using these identities, we obtain

$$\frac{\vartheta_0}{\vartheta} = \frac{G_{c0}}{G_c}.$$

Using this expression, the first-order condition for capital becomes

$$\frac{G_{c0}}{G_c} = 1 + f_k - \delta,$$

and the Euler equation for households then requires zero taxes on capital as claimed. ■

Note that because we are using linear taxes, we need to impose the stronger assumption that utility is homogeneous in c_0 and c. The original Atkinson–Stiglitz result only requires separability between consumption and leisure because it allows for non-linear taxes.

As the above derivation shows, the Atkinson-Stiglitz result hinges on three crucial assumptions: separable (and homothetic) preferences, no other restrictions on taxes, and the ability of the government to accumulate assets with no restriction.

In what follows, we explore the consequences of requiring the government to run a balanced budget. This is sufficient to break the result of zero taxes on capital and implies that optimal capital taxes are linked to its supply elasticity. This is also the key assumption used to derive Proposition (A.5) in the infinite horizon model.

A.4.4 Implications of Imposing a Balanced Budget

Before characterizing optimal policy in this case, we introduce some definitions.

First, define the Hicksian elasticity of capital supply, ε^k, as the percent change in k following a compensated change in the keep rate $1 - \tau^k$ announced after households have already derived all of their income in period 0. Thus when this tax change takes place, households can re-optimize their saving decisions and labor supply decisions in period 1, but cannot adjust their hours in of work in period 0. Moreover, define by $\sigma_{k\ell}$ the percent change in employment induced by the percent change in savings.

We now provide formulae for ε^k and $\sigma_{k\ell}$. Let $M^\ell = -\frac{u_\ell}{u_c} > 0$ denote the marginal rate of substitution between leisure and consumption, and analogously to the previous section, let $M^k = \frac{u_{c0}}{u_c} > 0$ be the intertemporal marginal rate of substitution between consumption at time 0 and time 1. Because the tax change is compensated, we have

$$dc_0 = -k \qquad\qquad dc = M^k \cdot dk + M^\ell \cdot d\ell.$$

That is, consumption only changes due to the behavioral response of savings and labor supply decisions, but not because of the changes in after-tax prices obtained by households.

Optimal saving decisions satisfy

$$(R-\delta)\cdot(1-\tau^k) = M^k - 1.$$

Totally differentiating this expression and using the formulae for dc_0 and dc, we obtain

$$(M^k-1)\cdot d\ln(1-\tau^k) = (M_c^k\cdot M^k - M_{c0}^k)\cdot k\cdot d\ln k + (M_c^k\cdot M^\ell + M_\ell^k)\cdot \ell\cdot d\ln \ell. \tag{A.28}$$

Optimal labor supply decisions satisfy

$$w\cdot(1-\tau^\ell) = M^\ell.$$

Totally differentiating this expression and using the formulae for dc_0 and dc, we obtain

$$0 = (M_c^\ell\cdot M^k - M_{c0}^\ell)\cdot k\cdot d\ln k + (M_c^\ell\cdot M^\ell + M_\ell^\ell)\cdot \ell\cdot d\ln \ell. \tag{A.29}$$

Equations (A.28) and (A.29) form a system of two equations on two unknowns and imply that

$$\frac{1}{\varepsilon^k} = \frac{(M_c^k\cdot M^k - M_{c0}^k)\cdot k}{M^k-1} + \frac{(M_c^k\cdot M^\ell + M_\ell^k)\cdot \ell}{M^k-1}\sigma_{k\ell}, \qquad \sigma_{k\ell} = -\frac{(M_c^\ell\cdot M^k - M_{c0}^\ell)\cdot k}{(M_c^\ell\cdot M^\ell + M_\ell^\ell)\cdot \ell}$$

Following analogous steps, the Hicksian elasticity of labor supply, ε^ℓ, and $\sigma_{\ell k}$ can be computed as

$$\frac{1}{\varepsilon^\ell} = \frac{(M_c^\ell\cdot M^\ell + M_\ell^\ell)\cdot \ell}{M^\ell} + \frac{(M_c^\ell\cdot M^k - M_{c0}^\ell)\cdot k}{M^\ell}\sigma_{\ell k}, \qquad \sigma_{\ell k} = -\frac{(M_c^k\cdot M^\ell + M_\ell^k)\cdot \ell}{(M_c^k\cdot M^k - M_{c0}^k)\cdot k}.$$

Finally, denote by $\alpha^k = (M^k-1)\cdot k/((M^k-1)\cdot k + M^\ell\cdot\ell)$ and $\alpha^\ell = M^\ell\cdot\ell/((M^k-1)\cdot k + M^\ell\cdot\ell)$ the share of capital and labor income in household income.

PROPOSITION A.8 *Suppose that the government is restricted and must set $a^g \le 0$, and that this constraint binds. Optimal taxes are given by*

$$\frac{\tau^k}{1-\tau^k} = \frac{\mu}{\vartheta+\mu}\frac{1}{\varepsilon^k} - \frac{\alpha^\ell}{\alpha^k}\sigma_{k\ell}\frac{\tau^\ell}{1-\tau^\ell} \qquad \frac{\tau^\ell}{1-\tau^\ell} = \frac{\mu}{\vartheta+\mu}\frac{1}{\varepsilon^\ell} - \frac{\alpha^k}{\alpha^\ell}\sigma_{\ell k}\frac{\tau^k}{1-\tau^k}, \tag{A.30}$$

where $\mu > 0$ denote the multiplier on the first-period IC constraint and ϑ the multiplier on the first period resource constraint.

PROOF. Because the constraint $a^g \le 0$ binds, we have that the IC constraint in period 0 is slack and the IC constraint in period 1 binds. Thus we can rewrite the Ramsey problem as

$$\max_{\{k_0,k,c_0,c,\ell_0,\ell\}} u(c_0,c,\ell_0,\ell)$$

$$\begin{aligned}\text{subject to: } c_0 + g_0 + k &\leq f_0(k_0,\ell_0) + (1-\delta)\cdot k_0 \\ c + g &\leq f(k,\ell;\theta) + (1-\delta)\cdot k \\ g &\leq f(k,\ell) + (1-\delta)\cdot k - M^k \cdot k - M^\ell \cdot \ell,\end{aligned}$$

where $M^\ell = -\frac{u_\ell}{u_c} > 0$ denotes the marginal rate of substitution between leisure and consumption and $M^k = \frac{u_{c0}}{u_c} > 0$ denotes the intertemporal marginal rate of substitution between consumption in periods 0 and 1. Both marginal rates of substitution are functions of c_0, c, ℓ_0, ℓ.

Denote by $u_{c0} \cdot \vartheta_0$ the multiplier on the resource constraint at time 0, $u_c \cdot \vartheta$ the multiplier on the resource constraint in period 1, and $u_c \cdot \mu$ the multiplier on the IC constraint.

The first-order condition for c_0 is given

$$M^k = M^k \cdot \vartheta_0 + \mu \cdot \left[M^k_{c0} \cdot k + M^\ell_{c0} \cdot \ell\right].$$

The first-order condition for c is given

$$1 = \vartheta + \mu \cdot \left[M^k_c \cdot k + M^\ell_c \cdot \ell\right].$$

Turning to labor, we obtain the first-order condition:

$$\vartheta \cdot f_\ell - M^\ell + \mu \cdot (f_\ell - M^\ell) = \mu \cdot \left[M^k_\ell \cdot k + M^\ell_\ell \cdot \ell\right].$$

Plugging in the first-order condition for c, we obtain

$$\vartheta \cdot f_\ell - M^\ell \cdot \left(\vartheta + \mu \cdot \left[M^k_c \cdot k + M^\ell_c \cdot \ell\right]\right) + \mu \cdot (f_\ell - M^\ell) = \mu \cdot \left[M^k_\ell \cdot k + M^\ell_\ell \cdot \ell\right],$$

which can be rearranged to

$$\frac{\tau^\ell}{1-\tau^\ell} = \frac{\mu}{\vartheta + \mu} \frac{\left[M^k_\ell \cdot k + M^\ell_\ell \cdot \ell\right] + M^\ell \cdot \left[M^k_c \cdot k + M^\ell_c \cdot \ell\right]}{M^\ell}.$$

Turning to capital, the first-order condition is

$$\vartheta \cdot (f_k + 1 - \delta) + \mu \cdot (f_k + 1 - \delta - M^k) = M^k \vartheta_0,$$

which, using the first-order conditions for c_0 and c, can be rewritten as

$$\begin{aligned}\vartheta \cdot (f_k + 1 - \delta) + \mu \cdot (f_k + 1 - \delta - M^k) =& M^k \left(\vartheta + \mu \cdot \left[M_c^k \cdot k + M_c^\ell \cdot \ell\right]\right) \\ & - \mu \cdot \left[M_{c0}^k \cdot k + M_{c0}^\ell \cdot \ell\right].\end{aligned}$$

This can be rearranged to

$$\frac{\tau^k}{1-\tau^k} = \frac{\mu}{\vartheta + \mu} \frac{M^k \cdot [M_c^k \cdot k + M_c^\ell \cdot \ell] - [M_{c0}^k \cdot k + M_{c0}^\ell \cdot \ell]}{M^k - 1}$$

Using the definition of the Hicksian elasticities introduced above, we can rewrite optimal taxes as

$$\begin{aligned}\frac{\tau^k}{1-\tau^k} =& \frac{\mu}{\vartheta+\mu}\frac{1}{\varepsilon^k} - \frac{M^\ell \cdot \ell}{(M^k - 1) \cdot k} \sigma_{k\ell} \frac{\tau^\ell}{1-\tau^\ell} \\ \frac{\tau^\ell}{1-\tau^\ell} =& \frac{\mu}{\vartheta+\mu}\frac{1}{\varepsilon^\ell} - \frac{(M^k - 1) \cdot k}{M^\ell \cdot \ell} \sigma_{\ell k} \frac{\tau^k}{1-\tau^k},\end{aligned}$$

which coincide with the formulae in the proposition. ■

The optimal tax formulae in equation (A.30) are a generalization of those provided in Proposition 1 (except that we abstracted from labor market frictions for the purposes of this part of the Appendix). Here, optimal taxes are inversely linked to their supply elasticities. But now, optimal capital tax also depends on its effects on employment via the cross-elasticity $\sigma_{k\ell}$ (put differently, the optimal capital tax now depends on the fiscal externalities it creates by raising or lowering employment). Similarly, the optimal labor tax depends on its impact on savings via the cross-elasticity $\sigma_{\ell k}$.

In practice, income effects on labor supply are weak. This implies that the cross-elasticities $\sigma_{k\ell}$ and $\sigma_{\ell k}$ are small and these terms have a small effect on optimal taxes. In particular, in the limit case with no income effects on labor supply (as in the quasi-linear preferences used in the main text and in the infinite horizon version), the above formulae boil down to those provided in Proposition 1 (except that we have simplified the expressions by setting $\varrho = 0$).

Corollary A.2 *If there are no income effects on labor supply so that utility is given by*

$$u(c_0 - \nu(\ell_0), c - \nu(\ell)),$$

optimal taxes are given by

$$\frac{\tau^k}{1-\tau^k} = \frac{\mu}{\vartheta+\mu}\frac{1}{\varepsilon^k} \qquad \frac{\tau^\ell}{1-\tau^\ell} = \frac{\mu}{\vartheta+\mu}\frac{1}{\varepsilon^\ell},$$

where $\mu > 0$ *denote the multiplier on the first-period IC constraint and* ϑ *the multiplier on the first period resource constraint.*

Proof. With these preferences we have $\sigma_{k\ell} = \sigma_{\ell k} = 0$, which follow from the definitions of the cross elasticities. ■

This corollary shows that the formulae in the main text only require quasi-linearity within periods. The stronger assumption of quasi-linearity on c used in the main text is imposed to simplify the exposition by removing the cross-elasticity effects.

The contrast between Propositions A.7 and A.8 underscores the role of allowing the government to have a single intertemporal budget constraint. Once we depart from this by constraining the government's capacity to regulate assets, the commodity taxation principles in Atkinson–Stiglitz (1972) are no longer valid. In particular, Proposition A.8 has established that when the government has to run a balanced budget, the formulae in equation (A.30) apply exactly and reinstate the intuitive notion that optimal taxes should depend on the supply elasticities of the relevant factors as well as on cross-elasticity effects. In particular, it is optimal to set a lower tax on capital than on labor only when capital taxes reduce capital and labor supplies more than labor taxes.

A.5 Computation of Effective Tax Rates

Summary Table A.9 presents the sources and main computations required to obtain our measure of net operating surplus. The following sections describe the procedure we followed to compute the average taxes of interest, and the sources for our depreciation, investment price, and interest rate series.

Average Business Tax Rate on C-Corporations, $\tau^{\mathbf{c}}$

We first determine the average tax rates imposed on firms' profits *net of depreciation allowances*. The BEA produces series for capital consumption allowances for corporate and non-corporate taxpayers. We recover these series from FRED.[44] The national accounts classify as "corporate" all taxpayers that are subject to filing a form of the IRS series 1120, as reported in the NIPA Handbook. In particular, both C- and S-corporations are considered corporate. Notably, S-corporations are exempt from federal corporate income taxation, but this favorable treatment does not extend to all state and local business taxes.[45] In keeping with the foregoing discussion, we compute the tax base for state and local corporate taxes $\tau^{c,\mathrm{SL}}$ as the *net* operating surplus of C-corporations, $\mathrm{NOSCORP}^{\mathrm{IRS}}$, defined as the difference between the gross operating surplus of corporations. We calculate this measure as the sum of net operating surplus of corporations (line 8 of the BEA NIPA Table 1.14) and the consumption of fixed capital of corporations (line 12 of the BEA NIPA Table 1.14) minus the capital consumption allowances for corporations. We cannot directly use the consumption of fixed capital reported in the NIPA tables because they estimate the economic depreciation of the capital stock, while we are interested in recovering a measure of the *fiscal* depreciation of capital stocks. For this reason, we need to add back the NIPA consumption of fixed capital and then subtract the relevant allowances. The state and local tax revenues corresponding to the corporate tax base are given by the tax revenues from corporations at the state and local level (line 5 of BEA NIPA Table 3.3), $\mathrm{CT}^{\mathrm{SL}}$. We can thus estimate the capital tax faced by the corporate sector as

$$\tau_t^{c,\mathrm{SL}} = \frac{\mathrm{CT}_t^{\mathrm{SL}}}{\mathrm{NOSCORP}_t^{\mathrm{IRS}}}.$$

[44] The corresponding codes are A677RC1A027NBEA and A1700C0A144NBEA.

[45] For example, New York City, New Hampshire, California, Texas, and Tennessee do not recognize S-corporations for tax purposes. Other states have special rules on S-corporation election which do not necessarily match the federal criteria.

As mentioned above, the BEA also considers non-C-corporations as part of the corporate sector. However, only C-corporations are subject to federal taxes at the entity level, and the relevant tax base for federal corporate income taxes, $\tau^{c,\mathrm{SL}}$ is given by the net operating surplus of corporation that can be attributed to C-corporations, $\mathrm{NOSCORP}_t^{\mathrm{C,IRS}}$. Since NIPA tables do not provide a breakdown of corporate income by legal form of organization, we obtain the net income of corporations from the IRS SOI Tax Stats-Integrated Business data (IRS IBD), and compute share of C-corporations' net income in total corporate net income reported in IRS IBD Table 1,[46] which provides our estimate of the net operating surplus of C-corporations, $\mathrm{NOSCORP}_t^{\mathrm{C,IRS}}$. The federal revenues corresponding to this corporate tax base are given by the tax revenues from corporations at the federal level (line 5 of BEA NIPA Table 3.2), $\mathrm{CT}^{\mathrm{Fed}}$. Accordingly, the federal tax rate on capital income from C-corporations can be estimated as

$$\tau_t^{c,\mathrm{Fed}} = \frac{\mathrm{CT}_t^{\mathrm{Fed}}}{\mathrm{NOSCORP}_t^{\mathrm{C,IRS}}}.$$

Combining the federal, stat and local taxes, the overall entity-level tax rate on C-corporations is

$$\tau_t^c = \tau_t^{c,\mathrm{SL}} + \tau_t^{c,\mathrm{Fed}}.$$

Average Personal Tax Rates on Income from C-Corporations, $\tau^{e,c}$ and $\tau^{b,c}$

In addition to entity-level taxes, incomes distributed from C-corporations are subject to personal taxation. As described in the main text, we compute the corresponding tax rate as

$$\tau_t^{e,c} = \begin{matrix}\text{share directly}\\ \text{owned}_t\end{matrix} \cdot \left(\begin{matrix}\text{share short-}\\ \text{term ordinary}_t\end{matrix} \cdot \tau_t^o + \begin{matrix}\text{share long-}\\ \text{term qualified}_t\end{matrix} \cdot \tau_t^q + \begin{matrix}\text{share held}\\ \text{until death}_t\end{matrix} \cdot 0\% \right),$$

where τ_t^o is the average tax rate on short-term ordinary capital gains and dividends, and τ_t^q is the average tax rate on long-term qualified capital gains and dividends. For each year, we compute the share of corporate stocks directly owned by households as the ratio of share of equity held by households and non-profit organizations serving households over total corporate equity using data from FRED.[47] We build the share of profits realized through ordinary dividends and short-term capital gains on stocks directly owned by households using data from the IRS Individual Complete Report (Publication 1304, Table

[46] This series only span the period 1980-2013, with a missing data point in 1990, which we fill by linear interpolation. We assume that the share of net income of C-corporations in the total corporate sector has remained constant after 2013.

[47] The corresponding FRED series are HNOCEAQ027S and BOGZ1LM893064105Q, respectively

A) for the period 1990-2017 and the IRS SOI Tax Stats (Sales of Capital Assets Reported on Individual Tax Returns) for the period 1990-2012. Publication 1304 reports households' ordinary dividend income from corporate stocks, while the SOI Tax Stats reports the short-term capital gains on corporate stocks. The share of profits realized by households in the form of short-term gains or ordinary dividends can then be obtained by dividing the overall income from theses two sources by the net operating surplus of C-corporations.[48,49] We set "share short-term ordinary$_t$" to the average of the same variable over the period 1990-2012 for all years in our sample. The shares of profit realized long-term or until death are then computed assuming that half of the profits not realized in the short-term are never realized. This is in keeping with the findings reported in Table 14 of CBO (2006). Accordingly, the share of profits taxed at rate τ_t^q can be obtained as a residual, equal to half of the share of profits not taxed at the rate τ_t^o. The remaining share of profits is assumed to be unrealized until death, and subject to zero income taxation.

The average tax rates, τ_t^q and τ_t^o, are computed using data from the Office for Tax Analysis (OTA) for 1980–2014 (2019). Since both series exhibit trends over time, we extrapolate the data point for 2014 for the years 2015–2018. Ideally, τ_t^q should be the average marginal maximum tax rate for individuals realizing long-term capital gains and qualified dividends, and τ_t^o should be the average marginal ordinary income tax rate. However, these rates cannot be recovered from OTA data without detailed information on individual tax returns. We therefore proxy this quantity using the average tax rate on realized long-term capital gains provided by the OTA, which provides us with a measure for τ_t^q.[50] In addition to this average rate, the same source also reports the average long-term capital gains realized and the corresponding tax receipts. We combine this information with OTA data on total net capital gains and total taxes paid on net capital gains to obtain our measure of average taxes on short-term gains and ordinary dividends, τ_t^o.[51] In particular, we compute the tax

[48] In the notation above, this corresponds to the product

$$\begin{matrix}\text{share directly} \\ \text{owned}_t\end{matrix} \cdot \begin{matrix}\text{share short-} \\ \text{term ordinary}_t\end{matrix} .$$

[49] In practice, the share of profits taxed at ordinary rates is not limited to the short-term and ordinary dividends that accrue to the household from directly-owned corporate stocks. Capital gain distributions and IRA distributions—which originate from indirectly owned stocks—are also taxed at the ordinary rate, and constitute about 23% of realized profits over the period we considered. As a result, in our computations the share of profits taxed at ordinary income rates is 48%. Of this number, 25% comes form short-term gains and ordinary dividends from directly owned stocks (37% of stocks owned directly by households times 60% of profits realized in the form of short-term gains or ordinary dividends).

[50] Recovered at `https://www.treasury.gov/resource-center/tax-policy/tax-analysis/Documents/Taxes-Paid-on-Long-Term-Capital-Gains.pdf`.

[51] We recover this data at `https://www.treasury.gov/resource-center/tax-policy/tax-analysis/`

base for τ_t^o by subtracting realized long-term capital gains from total net capital gains. The relevant tax revenue is computed analogously, by subtracting total taxes paid on long-term capital gains from total taxes paid on total realized net capital gains. The ratio of these two quantities provides us with our estimate for τ_t^o.

The tax on interest income from C-corporations, $\tau^{b,c}$, is computed as explained in the main text. We obtained the share of fully taxable and temporarily deferred interest income and the average marginal tax rate on interest income for 2014 from Tables A-3 and A-4 of CBO (2014).

Average Tax Rates on Profits from S-Corporations, $\tau^{o,s}$ and $\tau^{b,s}$

Although S-corporations do not pay corporate income tax, the capital income from these corporations is taxed on the household side and the tax rate depends on how this income is realized. Long-term gains are taxed at the maximum marginal tax rate, while profits realized as short-term gains or net business income are taxed at the ordinary marginal tax rates. We obtained short-term capital gains from the sales of partnerships and S-corporations from the SOI Tax statistics Complete Year Data, Table 1 for years 1995-2011, and short-term gains to S-corporations in proportion to their share on the net income of partnerships plus S-corporations. We obtained profits realized through net business income from IRS Publication 4801, which provides yearly estimates for each item in IRS form 1040. In particular, taxpayers use columns (f)–(j) of Schedule E to register passive and non-passive income and losses from S-corporations and section 179 deductions.[52,53] These data are available for 2003-2017 and are reported in the yearly files of line-item estimates that can be downloaded from the IRS website.[54] This allows us to compute S-corporation profits realized in the form net business income as the sum net passive and active income minus the Section 179 deductions. We add this term to the realized short-term gains attributed to S-corporations as explained above, and divide this quantity by the net operating surplus attributable to S-corporations to obtain the short-term gain and business income share of S-coporation profits. Once again, we use IRS IBD Table 1 to attribute a fraction of $\text{NOSCORP}^{\text{IRS}}$ to S-corporations in proportion to their share of the net income of corporations. The long-term gain share of S-corporation income is then simply obtained as the complement of the short-term and business income

Documents/Taxes-Paid-on-Capital-Gains-for-Returns-with-Positive-Net-Capital-Gains.pdf.

[52]Passive income and losses are reported for taxpayers who own S-corporations but do not participate actively to their administration. Active income and losses are for owners of S-corporation who actively administer the business or provide labor services to it.

[53]Section 179 of the tax code allows business owners to deduct investment expenses below a certain amount. The TCJA of 2017 set the maximum section 179 deduction at $1 million.

[54]https://www.irs.gov/statistics/soi-tax-stats-individual-income-tax-returns-line-item-estimates

share. We set this share equal to its average over the period 2003–2011 for all our sample. As mentioned in the previous section, the tax status of S-corporations is not recognized by all state and local government authorities. To account for these additional taxes, we computed the ordinary income tax rate for S-corporation owners as the sum of their personal income tax and the state and local business tax rate, $\tau^{c,\text{SL}}$, described above. We estimate the average marginal income tax rate of S-corporation proprietors as the average income tax rate applied to short-term capital gains realized by owners of C-corporation stocks. Doing so amounts to assuming that the distribution of income of S-corporation proprietors coincides with that of C-corporation investors. This assumption is supported by Table A.4 in CBO (2014), which shows almost no difference between the average marginal tax rate on short-term capital gains of corporations and the average marginal tax rate on passthrough business profits. Finally, we calculate the tax rate on debt-financed investment analogously to C-corporations using the data provided in Table A.3 and A.4 in CBO (2014).

Assigning Depreciation Schedules

Table A.10 presents the sources we used to assign depreciation schedules to specific (fixed) asset types from BEA Table 2.7 of BEA FAT together with the resulting class lives and depreciation systems. Tables B.1-2 of IRS Publication 946 detail the class lives and the depreciation method according to the MACRS system, which applies to assets installed starting from the 1986 fiscal year. This allows us to match each of the fixed assets categories in BEA Table 2.7 to a class life. We then use the same class life to obtain the depreciation schedules according to the ACRS system from IRS Publication 534, which applies to property put in service in fiscal years 1981-1985.

Tables B.1-B.2 of IRS Publication 946 divide all types of capital into asset classes with corresponding class lives and depreciation methods. Tables B.1 collects asset classes for general-purpose capital (e.g., autos, trucks, office equipment). Table B.2 instead attributes class lives to the remaining asset classes according to the specific sector and application in which capital is employed, with considerable degree of detail. For example, all equipment used in the manufacturing of tobacco products (asset class 21.0) has a class life of 7 years, while the equipment used for knitting goods (asset class 22.1) has a class life of 5 years. Since BEA Table 2.7 does not allow us to distinguish the sector of application of many asset classes, we use the following strategy to build the crosswalk in Table A.10: Tables 6.A-B in the BEA NIPA Handbook contain the deflators (PPI's) that the BEA uses to build quantity indexes for each of the asset classes in BEA Table 2.7. This allows us to recover information on the underlying sectors of application for each type of capital, which we then match to the types of assets mentioned in the description of asset classes contained in Table B.2 of IRS Publication 946. For example, we match asset class 22.1 (equipment used for the production of knitting goods) and 22.4 (nonwoven fabrics) in Publication 946 to "Special Industry Machinery" in BEA Table 2.7, since the latter cites the PPI for textile machinery among the PPI's used to build the quantity index for "Special Industry Machinery". As this example illustrates, items in BEA Table 2.7 often correspond to multiple asset classes in Publication 946, each with potentially different class lives. We set the class life of each item in BEA Table 2.7 to the the mode of Publication 946 matching the class lives of asset classes, obtained as in the example above.[55]

The column "Sources P. 946" reports the items of Tables B.1-2 used to assign class lives. In some instances, Publication 946 refers to other sections of the tax code, or provides specific exceptions to the depreciation method that would apply following Tables B.1-2. When this is the case, the column "Sources P. 946" cites either the passage of Publication 946 listing the property under consideration, or the section of the tax code. We report the modal class life according to the Asset Depreciation Range system (ADR), the Accelerated Cost Recovery System (ACRS), and Modified Accelerated Cost Recovery System (MACRS). The ADR applies to assets installed in 1970-1980, ACRS to assets installed in 1981-1986, and MACRS applies to capital installed from 1986 onwards. MACRS consists of two depreciation systems, the General Depreciation System (GDS) and the Alternative Depreciation System (ADS), each prescribing different depreciation schedules. The ADS only applies to specific class lives and uses of property. However, the level of detail in BEA Fixed Asset (FA) Tables is not sufficient to attribute assets to this system precisely. We therefore follow the relevant GDS schedules when computing allowances and apply the MACRS system listed in the "GDS". "SL" denotes the straight-line method, while 200 and 150 denote the declining-balance (DB) methods with 200% and 150% accelerated depreciation, respectively. Finally, the column "HS" reports the classification in House and Shapiro (2008) when this is available.

We use GDS depreciation schedules with half-year convention from Appendix A of IRS Publication 946 (MACRS), and IRS Publication 534 (ACRS), and we apply the straight-line method for ADR, with the class lives listed in Tables B.1-2 of IRS Publication 946.[56] The MACRS provides schedules for assets installed in specific quarters or months of the year. We choose the half-year convention since we rely on annual data which does not allow us to

[55] In only one case—fabricated metal products—we chose the generic equipment class life of 7 years following House and Shapiro (2008), instead of the modal life of 20 that follows from our method.

[56] Using the class lives in Table B.1-2 and the straight-line method is likely to lead to some imprecision since the ADR system allowed substantial discretion in the choice of class lives, as much as ±20% from the baseline IRS class life. The choice of the depreciation method was also left to taxpayer discretion.

establish when capital was installed during the year.

Computing Total Discounts from Allowances

As discussed in the main text, depreciation allowances give rise to a discount on the purchase price of capital goods. This discount is given by the present discounted value of current and future tax payments that the business can deduct expensing the statutory allowance in each year. Assuming that the business in question correctly anticipates future changes in taxes and interest rate, and given a sequence of business tax rates $\{\tau_t\}$ and depreciation schedules $\{d_t^j\}$, tax discounts are given as:

$$\text{total discount from allowances}_t^j = d_t^j \cdot \tau_t + \sum_{s=0}^{\infty} d_{t+s+1}^j \cdot \tau_{t+s+1} \cdot \prod_{k=0}^{s} \frac{1 - d_{t+k}^j}{1 + r_{t+k+1}}.$$

Under our baseline assumption that depreciation rates and taxes are not changing, this expression simplifies to:

$$\text{total discount from allowances}_t^j = d_0^j \cdot \tau_t + \sum_{s=0}^{\infty} d_{s+1}^j \cdot \tau_t \cdot \prod_{k=0}^{s} \frac{1 - d_k^j}{1 + r_{t+k+1}}.$$

This term equals $\alpha_t \tau_t$ in the notation of the main text.

Computing Effective Taxes on Different Types of Capital

The final step to compute the average effective capital taxes reported in the main text consists involves averaging the effective taxes for the various (legal) form of organization and type of financing obtained above. To do, we first compute the share of debt and equity financing for each legal form of organization. We obtain the series for total equity and debt of the corporate and non-corporate sector from FRED.[57] This allows us to directly compute the share of capital financed through debt and equity in the non-corporate sector. We follow the CBO (2006, 2014) and attribute debt and equity to C-corporations and S-corporations. The IRS SOI provides income tax returns for all corporations for 1994-2013. We compute total equity as the sum of the capital stock, paid-in capital, retained earnings and adjustment to shareholders' equity, minus the treasury stock cost. We then compute the share of total corporate equity in the tax returns that relates to S-corporations, and attribute to them the relevant part of the aggregate stock of corporate equity (about 4% of the total). The remaining fraction is attributed to C-corporations. We assigned debt to

[57]Series BCNSDODNS, NCBEILQ027S, NNBCMIA, TNWBSNNB.

the two forms of organization in proportion to their share in total interest deductions of corporations, as reported by the IRS SOI. The share of debt financing for each legal form of organization is therefore given by its stock of debt over the sum of debt and equity, while its complement measures equity financing. Since the series exhibit trends, we use closest-neighbor extrapolation to fill in the missing data for the years before 1994 and after 2013. Armed with these shares, we can compute effective taxes on capital for each legal form of organization. Finally, we construct the economy-wide average effective capital tax by weighing the tax rate of each legal form of organization by its share of net business income in each year. The source for net business income by form of organization is once again the IRS IBD.

Sources for the Computation of the Effective Labor Tax Rate, τ^ℓ

We calculate the effective labor tax rate, τ^ℓ as the weighted average of labor income and payroll taxes and the wedge introduced by imperfect valuation of employer-provided pension and health insurance contributions:

$$\tau^\ell = \frac{\text{salaries} \cdot (\tau^h + \tau^p) + \text{benefits} \cdot (1 - \varphi)}{\text{compensation}}.$$

Line 2 in NIPA Tables 6.11B-D contains the value of employers' contributions for employee pension and health insurance funds, while line 2 of BEA NIPA Table 1.10 provides the total compensation of employees in the economy. Subtracting employers' contributions from total compensations gives us total salaries. We use the average personal income tax rate of the bottom 95% of the income distribution from IRS SOI Tax Stats for 1986–2017 as our measure of the personal income tax rate, τ^h.[58] The payroll tax rate, τ^p, is computed as the sum of the Old-Age, Survivors, and Disability Insurance (OASDI) and Medicare's Hospital Insurance (HI) rates for each year, that we retrieve from the Social Security Administration Website.[59]

Other Sources

We obtained investment in private fixed assets by type from BEA FAT 2.7. We computed the depreciation rate of each type of fixed assets in each year, dividing current-cost depreciation from BEA FAT Table 2.4 by the current-cost stock of each type from Table 2.4. The source for fixed asset price changes is BEA FAT Table 2.8. When computing effective capital taxes by category for equipment, software and nonresidential structures, we weigh the effective

[58] "Individual Statistical Tables by Tax Rate and Income Percentile", Table 2.
[59] `https://www.ssa.gov/OACT/ProgData/taxRates.html`.

capital tax constructed for each type of asset by the share of investment in each category as listed in BEA FAT Table 2.7. As mentioned in the text, we use Moody's Seasoned AAA Corporate Bond Yield from FRED (series AAA) deflated by the CPI for all urban consumers (CPIAUCSL). For robustness, we also used allowances and effective tax series using the lending interest rate from the World Bank adjusted for inflation using the GDP deflator (World Bank indicator FR.INR.RINR). This has a minimal impact on our results, slightly raising the present discounted value of depreciation allowances. The average real return on S&P 500 stocks over the period 1957–2008 is computed deflating the FRED series SP500 by the CPI for all urban consumers.

Table A.9: Components for the computation of capital taxation

Variable Name	Full Name	Components/Formula	Elaboration on:	Notes
NOSPCU	Net operating surplus of private enterprises	Includes: net interest payments of domestic businesses; net transfer payments; proprietors' income; rental income of persons; corporate profits gross of corporate taxes; all variables are adjusted for inventory valuation and capital consumption.	BEA Table 1.10	
NOSCORP	Net operating surplus of corporations	Component of the above	BEA Table 1.14	
CFCPCU, CFC-CORP	Consumption of fixed capital of private enterprises, corporations	—	BEA Tables 1.1.10, 1.1.14	Represents *economic* depreciation of the capital stock
OS*	*Gross* operating surplus of private enterprises, corporations	NOS* + CFC*	BEA Tables 1.1.10, 1.1.14	Represents the *economic* tax base before allowing for depreciation
OSPUE	*Gross* operating surplus of private *unincorporated* enterprises	OSPCU - OSCORP	—	—
CCAll*	Capital consumption allowances for PUE, corporations	—	BEA data from FRED.	—

Table A.10: Class lives and depreciation schedules for equipment, structures, and intellectual property products

Type	Sources P. 946	ADR	ACRS	MACRS	HS	GDS
Computers and peripheral equipment	0.12	6	5	5	5	200
Communication equipment	36, 48.2, 48.37, 48.42–.45, 48.13, 00.11, 48.35–.36,48.38–42, 48.31, 48.34	10	5	7	5	200
Medical equipment and instruments	sec. 168(B)iv,		5	5	7	200
Nonmedical instruments	36,48.37, 48.39, 48.44, 26.1, 37.2	6	5	7	7	200
Photocopy and related equipment	0.13,	6	5	5	5	200
Office and accounting equipment	0.13,	6	5	5	5	200
Fabricated metal products	48.42, section 168(C), 49.12, 40.52, 49.11, 49.13, 49.21, 49.221, 49.3, 49.4, 51	6	10	7	7	150
Engines and turbines	6	5	7	15	200	
Computers and peripheral equipment	0.12	6	5	5	5	200
Metalworking machinery	34.01, 37.12, 33.21, 37.33, 33.2, 33.4, 34.0, 35.0, 37.11, 37.2, 37.31, 37.41, 37.42	12	5	7	7	200

Type	Sources P. 946	ADR	ACRS	MACRS	HS	GDS
Special industry machinery, n.e.c.	20.5, 30.11, 30.21, 32.11, 22.1, 22.3,22.4, 23.0, 24.1, 24.3, 28.0, 36, 36.1, 57.0 , 20.4, 22.2, 22.5, 24.2, 24.4, 26.1, 26.2, 27.0, 30.1, 30.2, 31.0, 32.1, 32.3, 79.0 80.0, 13.3, 20.1–.3, 32.2	10	5	7	7	200
General industrial, including materials handling, equipment	00.241,00.242[60]	6	5	5	7	200
Electrical transmission, distribution, and industrial apparatus	48.38, 48.31, , 0.4, 49.11, 49.13, 49.14	10	5	20	7	150
Trucks, buses, and truck trailers	00.23–00.242	4	5	5	5	200
Light trucks (including utility vehicles)	0.241,	4	3	5	5	200
Other trucks, buses, and truck trailers	00.23,00.242,	6	5	5	5	200
Autos	0.22,	3	3	5	5	200
Aircraft	0.21,	6	5	5	7	200
Ships and boats	0.28,	10	5	10	10	150
Railroad equipment	40.1,	14	5	7	7	200
Furniture and fixtures	0.11,	10	5	7		200
Agricultural machinery	1.1,	10	5	7	7	150
Construction machinery	15,	6	5	5	5	200
Mining and oilfield machinery	13, 13.1, 10,13.2,	6	5	5	7	200

[60] Exclusion of general purpose from most sectoral class lives of conveyor belts and general-purpose tools.

Type	Sources P. 946	ADR	ACRS	MACRS	HS	GDS
Service industry machinery	57, 79–80,	9	5	7	7	200
Electrical equipment, n.e.c.	Ch.4, p.28	6	5	7	7	200
Other nonresidential equipment	Ch.4, p.28	6	5	7	7	200
Residential equipment	Ch.4, p.28					
Structures			15–18–19			SL
Nonresidential structures[61]	Ch. 4 p. 31			39[62]		
Commercial and health care	Ch. 4 p. 31			39	39	
Manufacturing structures	Ch. 4 p. 31			39	39	
Electric structures	49.12,49.15, 49.11, 49.13, 49.14	20		20	20	150
Other power structures	49.23, 49.24, 49.25	14		15	15	150
Communication	48.14,	15		15	15	150
Mining exploration, shafts, and wells						
Petroleum and natural gas	13.0, 13.1, 13.2	6	10	5	5	200
Mining structures	10,	10	10	7	5	200
Farm structures				20	20	150
Residential structures				27.5		SL
Nonresidential intellectual property products[63]	section 197			15		SL

[61] Applies to religious, education, lodging, amusement and other nonresidential structures that are not explicitly mentioned below

[62] As per publication 946, structures put in service before 1994 should have a useful life of 31.5 years. For simplicity, we use 39 for all years.

[63] Applies to all intellectual property products not explicitly mentioned below.

Type	Sources P. 946	ADR	ACRS	MACRS	HS	GDS
Software					5	
Prepackaged	Ch. 1, p. 10, not sec.197			3		SL
Custom	Ch. 1, p. 10, not sec.197			3		SL
Own account	sec. 167(f)1			15 years		SL
Research and development	item 5, sec 197			15		SL