NBER WORKING PAPER SERIES

DOES THE US TAX CODE FAVOR AUTOMATION?

Daron Acemoglu
Andrea Manera
Pascual Restrepo

Working Paper 27052
http://www.nber.org/papers/w27052

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2020

Prepared for the Brookings Institution Spring Conference of 2020. We are grateful to Jan Eberly, Larry Katz, James Poterba, Ivan Werning, Owen Zidar, and Eric Zwick for excellent comments and suggestions. We gratefully acknowledge financial support from Brookings Institution, Google, the National Science Foundation, Schmidt Sciences, the Smith Richardson Foundation and the Sloan Foundation. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Does the US Tax Code Favor Automation?
Daron Acemoglu, Andrea Manera, and Pascual Restrepo
NBER Working Paper No. 27052
April 2020
JEL No. J23,J24

ABSTRACT

We argue that the US tax system is biased against labor and in favor of capital and has become more so in recent years. As a consequence, it has promoted inefficiently high levels of automation. Moving from the US tax system in the 2010s to optimal taxation of capital and labor would raise employment by 4.02% and the labor share by 0.78 percentage points, and restore the optimal level of automation. If moving to optimal taxes is infeasible, more modest reforms can still increase employment by 1.14–1.96%, but in this case efficiency can be increased by imposing an additional automation tax to reduce the equilibrium level of automation. This is because marginal automated tasks do not bring much productivity gains but displace workers, reducing employment below its socially optimal level. We additionally show that reducing labor taxes or combining lower capital taxes with automation taxes can increase employment much more than the uniform reductions in capital taxes enacted between 2000 and 2018.

Daron Acemoglu
Department of Economics
MIT
50 Memorial Drive
Cambridge, MA 02142-1347
and NBER
daron@mit.edu

Andrea Manera
MIT
manera@mit.edu

Pascual Restrepo
Department of Economics
Boston University
270 Bay State Rd
Boston, MA 02215
and Cowles Foundation, Yale
pascual@bu.edu

An appendix is available at http://www.nber.org/data-appendix/w27052

1 Introduction

The last three decades have witnessed a declining share of labor in national income, stagnant median real wages and lower real wages for low-skill workers in the US economy (Elsby, Hobijn & Sahin, 2011; Acemoglu & Autor, 2011; Karabarbounis & Neiman, 2014). The labor share in non-farm private businesses declined from 63% in 1980 to 56% in 2017, while median real wages grew only by 16% (as compared to GDP per capita which doubled during the same period) and the real wages of male workers with a high school diploma fell by 6% between 1980 and 2017. In the meantime, production processes have become increasingly automated, as computerized numerical control machines, industrial robotics, specialized software and lately artificial intelligence technologies have spread rapidly throughout the economy. For instance, the US economy had a total of 2.5 industrial robots per thousand workers in manufacturing in 1993 and this number rose to 20 by 2019 (Acemoglu & Restrepo, 2020a). From a base of essentially zero in the mid-2000s, the share of vacancies posted for artificial intelligence-related activities increased to 0.75% by 2018 (Acemoglu et al., 2020).

A common perspective among economists is that even if automation is contributing to the decline in the labor share and the stagnation of wages, the adoption of these technologies is beneficial, and any adverse consequences should be dealt with redistributive policies and investments in education and training. But could it be that the extent of automation is excessive, meaning that businesses are adopting automation technologies beyond the socially optimal level? If this were the case, the policy responses to this trend would need to be rethought.

In this paper, we show that the US tax system is biased against labor and as a result generates excessive automation and suboptimally low levels of employment and labor share. We first introduce a task-based model of automation, building on Acemoglu & Restrepo (2018, 2019a,b) and Zeira (1998), to study the interplay between taxes and automation. Our first theoretical result establishes that optimal capital and labor taxes depend on the inverse supply elasticities of these factors and labor market frictions. Consistent with Diamond & Mirrlees (1971), once capital and labor taxes are set optimally, there is no reason to distort equilibrium automation decisions. Intuitively, optimal taxes undo any distortions and ensure that market prices reflect the social values of capital and labor. Automation decisions based on these prices are therefore optimal.[1]

[1] We assume that the labor market friction is common across tasks. When labor market frictions affect tasks differentially, there is an additional reason for excessive automation as shown in our companion paper, Acemoglu, Manera & Restrepo (2020), and in that case, distorting automation may be beneficial even when taxes are set optimally.

Yet this result does not imply that equilibrium automation decisions are optimal at arbitrary capital and labor taxes. Our second theoretical result shows that if a tax system is biased against labor and in favor of capital—i.e., taxes on labor are too high and taxes on capital are too low—then reducing automation at the margin improves welfare. We show that this reduction can be achieved with an *automation tax*, which is an additional tax on the use of capital in tasks where labor has a comparative advantage. An automation tax is beneficial because reducing automation below its equilibrium level has second-order costs and first-order benefits. The costs are second-order as the productivity gains from automating marginal tasks are small (or, equivalently, the automation of marginal tasks corresponds to "so-so automation" in the terminology of Acemoglu & Restrepo, 2019a,b). But, when the tax system is biased against labor and thus the level of employment is below the social optimum, limiting automation and avoiding the resulting displacement of labor has first-order benefits.

A common intuition is that if taxes are distorted, then the best policy remedy is to correct these distortions. Hence, if a tax system treats capital too favorably, we should directly tackle this distortion and increase capital taxes. We demonstrate that this intuition does not always apply in the presence of other constraints—for example, a lower bound on labor taxes. Our third theoretical result shows that a tax system distorted in favor of capital may call for reducing equilibrium automation even if raising capital taxes is possible. In fact, when moving to the unconstrained optimum is not feasible, constrained optimal policy may involve lower capital taxes in addition to a reduced level of automation because this combination avoids the displacement of workers from marginal tasks while ensuring that capital gets used intensively in tasks that are (and should be) automated. Both of these margins contribute to raising employment and welfare. This result underscores the importance of distinguishing between the choice of capital intensity in tasks where capital has a comparative advantage and automation, which involves the substitution of capital for labor in additional tasks. An automation tax is beneficial precisely because it does not reduce capital intensity uniformly but discourages the automation of marginal tasks

Armed with these theoretical results, we turn to measuring effective taxes on capital and labor in the US and comparing them to their optimal counterparts. We find that labor is much more heavily taxed than capital, and this difference has increased in recent years. Effective labor taxes in the US are in the range of 25.5–33.5%. Effective capital taxes on software and equipment, on the other hand, are much lower, 10% in the 2010s and 5% after the 2017 tax reforms, though they used to be about 20% in 2000.[2] About half of this decline

[2] Acemoglu & Restrepo (2019b) document that technological changes in the four decades after World War II involved less automation and more rapid advances in technologies that increased human productivity (such as the creation of new tasks for workers) than recently. Though there are other reasons for why the

is due to the greater generosity of depreciation allowances.

Using plausible ranges for the elasticities of the capital and labor supply and estimates of labor market distortions, we find that the US tax system is biased against labor. In fact, our baseline estimates suggest that optimal labor taxes are lower than capital taxes—a 18.22% labor tax compared to a 26.65% capital tax. Optimal taxes are lower for labor than for capital because empirically plausible ranges of supply elasticities for capital and labor are similar, but employment is further distorted by labor market imperfections. Moving from the current tax system to optimal taxes would reduce the range of automated tasks by 4.1% and increase employment by 4.02% and the labor share by 0.78 percentage points.[3]

Our quantitative results show that, as in our theory, reducing automation is socially beneficial. Specifically, with no changes in capital and labor taxes, an automation tax of 10.15%—which implies that only tasks where the substitution of labor for capital reduces unit costs by more than 10.15% are automated—maximizes welfare and raises employment by 1.14% and the labor share by 1.93 percentage points. If capital taxes can be reduced as well, then a 12.9% automation tax combined with a reduction in capital taxes from 10% to 8.39% would achieve even higher welfare gains and increase employment by 1.59% and the labor share by 2.44 percentage points. We further show that tax reforms that involve lower labor taxes or combine lower capital taxes with an automation tax would have increased welfare and expanded employment much more than the uniform capital tax reductions enacted between 2000 and 2018.

We conclude with two extensions. First, we show that if human capital is endogenous, the asymmetric treatment of labor becomes more costly as it distorts human capital investments as well, leading to even lower optimal taxes on labor and more excessive automation under the current system. Second, we consider endogenous development of automation technologies, which come at the expense of other types of innovations that are more beneficial for labor. In this case, there are reasons for not just preventing excessive adoption of automation technologies but also redirecting technological change away from further automation (and this is true even with optimal taxes on capital and labor).

Our paper is related to several classic and recent literatures, though, to the best of our knowledge, no other paper investigates whether the US tax system favors automation.

First, there is an emerging literature on redistribution and taxation of automation technologies (Guerreiro, Rebelo & Teles, 2017; Thuemmel, 2018; Costinot & Werning, 2018).

---

direction of technology altered, the lower taxation of equipment and software capital may have also played a role.

[3]Despite these large changes in employment, the increase in welfare is given by a "Harberger triangle" and is thus smaller—0.38% in consumption-equivalent terms.

This literature studies whether adverse distributional effects of automation call for taxes on automation technologies. Our paper is complementary to this literature, as it focuses on situations in which the tax system is biased against labor and the key policy objective is to raise employment (not to redistribute income).

Second, our paper is related to the literature on optimal capital taxation (e.g., Atkinson & Stiglitz, 1972, Judd, 1985, Chamley 1986, and Straub & Werning, 2020). Our contribution is to show that in both two-period and infinite-horizon settings, provided that the government must run a balanced budget at each date, optimal taxes are given by the same inverse-elasticity formulae (with an additional term adjusting for labor market frictions). In contrast, this literature typically assumes that the government can freely accumulate assets and concludes that zero capital taxation is optimal in the long run. Straub & Werning (2020) show that if the supply of capital is not perfectly elastic (which means utility is not time-additive), then the government accumulates sufficient assets so that both capital and labor face zero taxes in the long run. We demonstrate in the Appendix that in the presence of labor market frictions, the same reasoning leads to a subsidy to labor. Thus, in the empirically relevant case of a finite supply elasticity of capital, even without the balanced budget assumption, the US tax system with low capital taxes and high labor taxes is far from optimal.

Third, our paper relates to the literature on the effects of tax reforms on investment and labor market outcomes. A branch of this literature estimates the differential responses of investment across firms facing different taxes (Goolsbee, 1998; Hasset & Hubbard, 2002; Edgerton, 2010; Yagan, 2015).[4] However, these estimates are informative about firms' demand for capital, not about the (long-run) elasticity of the supply of capital, which is the relevant object for optimal taxes. We discuss below estimates of this elasticity based on the response of the supply of capital to wealth and capital income taxes (see Kleven & Schultz, 2014; Zoutman, 2018; Brülhart et al., 2019; Jakobsen et al., 2019; Durán-Cabré et al., 2019). More closely connected to our work is a branch of this literature on the labor market implications of tax reforms. Suárez Serrato & Zidar (2016) exploit the incidence of tax changes across US counties and estimate that a 1% increase in the keep rate of corporate taxes raises employment by 3.5% and wages by 0.8%, and that workers bear 35% of the incidence. Garret, Ohrn & Suárez Serrato (2020) compare counties at the 75th percentile of exposure to bonus depreciation allowances to those at the 25th percentile and find a 2%

[4] Modal results in this literature find investment elasticities with respect to the keep rate (one minus tax rate) between 0.5 and 1. More recent work by House & Shapiro (2008) documents a larger investment response and argue that this was due to the temporary nature of the bonus, while Zwick & Mahon (2017) estimate investment elasticities with respect to the keep rate that are around 1.5 for most firms.

increase in employment, no changes in wages, and a 3.3% increase in investment in response to the reform. These estimates point to a fairly elastic response of employment and a less than perfectly elastic response of capital in local labor markets (a perfectly elastic response of capital would cause workers to bear the full incidence).

Finally, our modeling of automation builds on Zeira (1998), Autor, Levy & Murnane (2003), Acemoglu & Autor (2011) and most closely, Acemoglu & Restrepo (2018, 2019a,b). The task-based framework is useful in our setting because it shows how automation (substituting capital for labor in tasks previously performed by humans) creates a displacement effect while automating marginal tasks generates limited productivity gains (because firms are approximately indifferent between automating these tasks or producing with labor). This combination of displacement effects and small productivity gains is at the root of our result that the planner would like to reduce automation at the margin when the tax system is biased against labor. Our framework also clarifies how policy can affect the level of automation and why taxing automation is not the same as taxing capital.

The rest of the paper is organized as follows. Section 2 introduces our conceptual framework and derives our theoretical results. Section 3 provides a detailed discussion of the US tax system and maps the complex US tax code into effective capital and labor income taxes. Section 4 then explores whether these taxes are biased and how they compare against optimal taxes. Section 5 discusses two extensions of our framework, while Section 6 concludes. The Appendix contains proofs of the results stated in the text, various theoretical generalizations, and further details for and robustness checks on our empirical work.

2 Conceptual Framework

This section presents our conceptual framework for evaluating the optimality of capital and labor taxes and the extent of automation. To facilitate the exposition, we focus on a two-period model and generalize our main results to an infinite-horizon setting in the Appendix.

2.1 Environment

There is a unique final good, produced at time $t = 1$ by combining a unit measure of tasks,

$$y = \left(\int_0^1 y(x)^{\frac{\lambda-1}{\lambda}} dx \right)^{\frac{\lambda}{\lambda-1}}.$$

Tasks are allocated between capital and labor, and performed with the following task-level production function:

$$y(x) = \psi^\ell(x) \cdot \ell(x) + \psi^k(x) \cdot k(x), \tag{1}$$

where $\ell(x)$ is labor employed in task x, $k(x)$ is capital used in the production of task x, and $\psi^\ell(x)$ and $\psi^k(x)$ denote, respectively, the productivities of labor and capital in task x. We order tasks such that $\psi^\ell(x)/\psi^k(x)$ is nondecreasing and simplify the exposition by assuming that it is strictly increasing. We also suppose that when indifferent between producing a task with capital or labor, firms produce with capital. Therefore, there exists a threshold task θ such that tasks in $[0, \theta]$ are produced with capital and tasks in $(\theta, 1]$ are produced with labor. For now, there is no distinction between the adoption and the development of such technologies. We explore the implications of this distinction in Section 5.2.

The household side is inhabited by a representative household who lives for two periods, $t = 0$ and $t = 1$. There is no production in period 0, but the representative household is endowed with $\bar{y}$ units of output. Out of this, it consumes c_0 and saves the remaining $k = \bar{y} - c_0$ units, which are allocated to producing capital. Capital is used during period 1, is subject to depreciation at the rate δ, and is rented to firms at the rental rate R, so that households earn an after tax return of $(R - \delta) \cdot (1 - \tau^k)$. The period 1 budget constraint facing the household is

$$c \leq (1 + (R - \delta) \cdot (1 - \tau^k)) \cdot k + w \cdot (1 - \tau^\ell) \cdot \ell,$$

where R is the rental rate on capital paid by firms and w is the wage rate. Tax revenues are used for financing a fixed level of government expenditure, denoted by g.

The household chooses consumption, and the supply of capital and hours to maximize

$$u(\bar{y} - k) + c - \nu(\ell).$$

Here, $u(\bar{y} - k)$ is a concave function representing the utility from consuming $\bar{y} - k$ units of output in period 0; c denotes the utility from consumption in period 1; and $\nu(\ell)$ is a convex function representing the disutility from working. Quasi-linearity in period 1 is imposed for simplicity (see the Appendix for more general preferences).

We allow for various types of frictions in the labor market, modeled as introducing a wedge between the market wage and the representative household's marginal cost of supplying labor.

We denote this wedge by $\varrho \geq 0$.[5]

Market clearing for capital and labor requires $k = \int_0^1 k(x)dx$ and $\ell = \int_0^1 \ell(x)dx$.

To ensure uniqueness of optimal taxes below, we suppose that $u'(\bar{y} - k) \cdot k$ and $\nu'(\ell) \cdot \ell$ are convex. In addition, we assume that the equilibrium involves a positive net rate of return on investment. Finally, we denote by $\varepsilon^k(k)$ and $\varepsilon^\ell(\ell)$ the Hicksian elasticities of capital and labor. These are given by the response of capital and labor supply to a permanent percent change in the relevant keep rates (one minus the tax rates):

$$\varepsilon^k(k) = \frac{d\ln k}{d\ln(1-\tau^k)} = -\frac{u'(\bar{y}-k)-1}{u''(\bar{y}-k)\cdot k} \geq 0 \qquad \varepsilon^\ell(\ell) = \frac{d\ln \ell}{d\ln(1-\tau^\ell)} = \frac{\nu'(\ell)}{\nu''(\ell)\cdot \ell} \geq 0.$$

As the equation for $\varepsilon^k(k)$ makes clear, the concavity of period -1 utility, $u(\bar{y}-k)$, ensures that the marginal rate of substitution between consumption today and tomorrow is increasing in k and thus the supply of capital is not perfectly elastic (otherwise $\varepsilon^k(k)$ would be infinite).[6]

Note that our formulation assumes that τ^k is a tax on net—after depreciation—returns, not gross returns, and our formula for $\varepsilon^k(k)$ computes it as the elasticity of capital to a percent change in one minus the net tax on capital.

2.2 Equilibrium

Given taxes $\{\tau^k, \tau^\ell\}$ and the labor wedge ϱ, a market equilibrium is defined by factor prices $\{w, R\}$, a tuple of current output, consumption, capital and labor, $\{y, c, k, \ell\}$, and an allocation of tasks to factors, such that this allocation minimizes the after-tax cost of producing each task and the markets for capital, labor and the final good clear. The Appendix shows that the equilibrium level of output can be represented as:

$$(2) \qquad y = f(k, \ell; \theta) = \left(\left(\int_0^\theta \psi^k(x)^{\lambda-1} dx \right)^{\frac{1}{\lambda}} \cdot k^{\frac{\lambda-1}{\lambda}} + \left(\int_\theta^1 \psi^\ell(x)^{\lambda-1} dx \right)^{\frac{1}{\lambda}} \cdot \ell^{\frac{\lambda-1}{\lambda}} \right)^{\frac{\lambda}{\lambda-1}},$$

where the threshold task θ satisfies

$$(3) \qquad \theta = \theta^m(k, \ell) \equiv \underset{\theta \in [0,1]}{\arg\max}\, f(k, \ell; \theta).$$

[5]As shown in Acemoglu, Manera & Restrepo (2020), this wedge can be derived from bargaining between workers and firms or from efficiency wage considerations.

[6]A complementary reason for finite $\varepsilon^k(k)$ is that the technology for investment is convex (for example, the production of k units of capital requires $\phi(k)$ units of period 0 resources, where ϕ is strictly convex). If the profits from producing capital cannot be directly taxed, our optimal tax formulae apply regardless of whether $\varepsilon^k(k)$ reflects changes in the marginal rate of substitution between consumption today and tomorrow as a function of k or a convex investment technology.

Moreover, factor prices are given by the usual marginal conditions $f_k = R$ and $f_\ell = w$. Consequently, the market-clearing condition for capital is

$$u'(\bar{y} - k) = 1 + (f_k - \delta) \cdot (1 - \tau^k), \tag{4}$$

while the market-clearing condition for labor is

$$\nu'(\ell) = f_\ell \cdot (1 - \varrho) \cdot (1 - \tau^\ell), \tag{5}$$

so that the wedge ϱ and the labor tax τ^ℓ distort the labor market in similar ways.

Finally, the government budget constraint takes the form

$$g \le \tau^k \cdot (f_k - \delta) \cdot k + \tau^\ell \cdot f_\ell \cdot \ell. \tag{6}$$

A couple of points about this equilibrium are worth noting. As emphasized in Acemoglu & Restrepo (2018, 2019b), though the output level in the economy can be represented by a constant elasticity of substitution (CES) aggregate of capital and labor, the implications of this setup are very different from models that assume a CES production function with factor-augmenting technologies. First, there is a crucial distinction between capital intensity of production given a fixed allocation of tasks to factors and automation, represented by an increase in θ—which involves the substitution of capital for tasks previously performed by labor. This can be seen from the fact that holding the task allocation constant, the elasticity of substitution between capital and labor is λ, but when θ adjusts, the elasticity is greater. Second, further automation increases productivity but can easily reduce labor demand and the equilibrium wage because of the displacement it creates (mathematically, this works by changing the share parameters of the CES). In contrast, with a standard CES production function labor demand necessarily increases when capital becomes more productive. Third, and for the same reason, automation always reduces the labor share. Finally, our framework also clarifies that marginal increases in automation have second-order effects on aggregate output (because of (3)).

2.3 Optimal Policy

We now characterize optimal policy by considering the choices of a benevolent social planner that sets capital and labor taxes τ^k and τ^ℓ, and directly controls the extent of automation, represented by θ. We refer to the maximization problem of this planner as the *Ramsey problem*. As usual, this problem can be transformed so that the planner chooses directly an allocation $\{c, \ell, k, \theta\}$ that maximizes household utility subject to the resource constraint of the economy and a single *Implementability Condition*, which combines the government budget constraint in (6) and input market equilibrium conditions (4) and (5):

$$(7) \qquad \max_{c,\ell,k,\theta} \quad u(\bar{y} - k) + c - \nu(\ell)$$

$$\text{subject to:} \quad c + g = f(k, \ell; \theta) + (1 - \delta) \cdot k \qquad \text{(Resource constraint)}$$

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

Because the planner is assumed to choose the level of automation θ, we do not impose $\theta = \theta^m(k, \ell)$ as an additional constraint. We discuss issues of how the planner's choice of automation can be implemented below. Throughout, we use $\mu > 0$ to denote the multiplier on the Implementability Condition, which also gives the social value of public funds.

Proposition 1 ***(Optimal capital and labor taxes and automation)*** *The unique solution to the Ramsey problem in (7) satisfies*

$$(8) \qquad \frac{\tau^{k,r}}{1 - \tau^{k,r}} = \frac{\mu}{1 + \mu} \frac{1}{\varepsilon^k(k)} \qquad \qquad \frac{\tau^{\ell,r}}{1 - \tau^{\ell,r}} = \frac{\mu}{1 + \mu} \frac{1}{\varepsilon^\ell(\ell)} - \frac{\varrho}{1 + \mu},$$

and $\theta^r = \theta^m(k, \ell)$.

The proof of this proposition, like those of all other results in this paper, is provided in the Appendix. The optimal tax formulae in equation (8) follow from the first-order conditions for the maximization problem in (7). Uniqueness follows from the fact that the Ramsey program is convex (the objective function is quasi-concave and the constraint set is convex).

This proposition provides simple and intuitive formulae for the optimal taxes on capital and labor related to the social value of public funds and the inverse of the elasticity of supply of these factors. The formulae show that taxes should be lower for more elastic factors, and in addition, the optimal labor tax is further lowered by the presence of labor market frictions. This latter feature is intuitive: labor market frictions reduce employment beyond the socially optimal level, and the planner corrects for this by reducing labor taxation.

An immediate corollary of this proposition provides one set of sufficient conditions for uniform (symmetric) taxation of capital and labor—$\varepsilon^k(k) \simeq \varepsilon^\ell(\ell)$ and $\varrho \simeq 0$.

Corollary 1 *If* $\varepsilon^k(k) = \varepsilon^\ell(\ell)$ *and* $\varrho = 0$*, uniform taxation of capital and labor is optimal.*

In Section 4 we will see that realistic values of these parameters are not too far from $\varepsilon^k(k) \simeq \varepsilon^\ell(\ell) > 0$, but labor market imperfections imply $\varrho > 0$, so that our framework yields lower labor taxes than capital taxes in the optimum.

Although the formulae in equation (8) apply in a two-period model, the Appendix shows that, under the key assumption that the government must run a balanced budget, these formulae extend to an infinite-horizon setting.[7] The Appendix also derives similar formulae for general preferences over consumption and leisure, and clarifies the relationship between our result and Atkinson & Stiglitz's (1972) principles of optimal commodity taxation.

In line with Diamond & Mirrlees (1971), Proposition 1 also shows that, once optimal taxes are imposed on capital and labor, the planner has no reason to deviate from equilibrium automation decisions, $\theta^r = \theta^m(k, \ell)$. This is because any distortions in the labor market are corrected by optimal taxes, and thus, factor prices accurately reflect the social values of capital and labor. Consequently, profit-maximizing automation decisions are optimal as well. We will see that this is no longer true when taxes are not optimal or are subject to additional constraints.

2.4 Excessive Automation with Tax Distortions

Naturally, taxes in practice need not coincide with those characterized in Proposition 1 both because of additional constraints and for political economy reasons (policy-makers have other objectives and face political or other, unmodeled economic constraints). When that is the case, either capital or labor taxes can be (relatively) too low. The interesting case for us, both for conceptual and empirical reasons, is the one where capital taxes are too low and labor taxes are too high, and the necessary and sufficient condition for this follows from equation (8) in Proposition 1 and is presented in the next corollary.

Corollary 2 *If the tax system $\{\tau^k, \tau^\ell\}$ is below the peak of the Laffer curve and satisfies*

$$\frac{\frac{\tau^\ell}{1-\tau^\ell} + \varrho}{\frac{1}{\varepsilon^\ell(\ell)} - \frac{\tau^\ell}{1-\tau^\ell}} > \frac{\frac{\tau^k}{1-\tau^k}}{\frac{1}{\varepsilon^k(k)} - \frac{\tau^k}{1-\tau^k}}, \tag{9}$$

then $\tau^\ell > \tau^{\ell,r}$ and $\tau^{k,r} > \tau^k$—that is, the labor tax is too high and the capital tax too low.

Inequality (9) is sufficient for the tax system *being biased against labor and in favor of capital.*[8] An important implication of such a biased tax structure is that there is too little

[7]Even if the government is allowed to incur debt or accumulate assets, the result that the optimal tax system should not simultaneously impose significant taxes on labor and zero (or small) taxes on capital extends to an infinite-horizon setting provided that the long-run elasticity of capital supply, $\varepsilon^k(k)$, is not infinite. Straub & Werning (2020) show that, in a representative household economy where preferences are not time-additive separable and the tax system is not constrained by other consideration, optimal taxes on both capital and labor should converge to zero. We prove in the Appendix that if in addition there are labor market distortions, then optimal long-run taxes are lower on labor than capital.

[8]The government budget constraint implies that both taxes cannot be too high or too low at the same

employment relative to the optimal allocation in Proposition 1, and thus marginal increases in employment will have first-order positive effects on welfare. We exploit this insight in the next proposition, where we take the tax system as given and consider a marginal change in automation. To do this in the simplest way, we relax the government budget constraint, (6), and value changes in revenue at the social value of public funds given by the multiplier μ.

Proposition 2 ***(When reducing automation improves welfare)*** *Suppose that the tax system $\{\tau^k, \tau^\ell\}$ satisfies inequality (9) (and is thus biased against labor and in favor of capital). Welfare (inclusive of fiscal costs and benefits) increases following a small reduction in θ below $\theta^m(k,\ell)$. A small reduction in θ also increases net output provided that $\varepsilon^\ell(\ell) > \varepsilon^k(k)$ and government revenue provided that $\tau^\ell \cdot (1+\varepsilon^\ell(\ell)) > \tau^k \cdot (1+\varepsilon^k(k))$.*

This result shows that, in contrast with Proposition 1, when taxes are not optimal and are biased against labor (in the sense that inequality (9) holds), it is welfare improving to restrict automation below its equilibrium level. This result is intuitive in light of the observation in Corollary 2 that employment is below the socially optimal level. Specifically, a small reduction in automation will create a first-order welfare gain by shifting demand from capital to labor. Distorting automation is costly, but starting from the equilibrium level of automation, $\theta^m(k,\ell)$, this cost is second-order (since $f_\theta(k,\ell;\theta^m(k,\ell)) = 0$), and hence, a small reduction in automation is welfare improving. This intuition relates Proposition 2 to the notion of "so-so (automation) technologies" proposed in Acemoglu & Restrepo (2019a,b): automation is not beneficial to labor when it only increases productivity by a small amount, while still creating the usual displacement of workers as tasks are reallocated from them to capital. The equilibrium condition $f_\theta(k,\ell;\theta^m(k,\ell)) = 0$ implies that automation technologies adopted at marginal tasks are, by definition, so-so. The planner is therefore happy to sacrifice these so-so technologies in order to help labor.[9]

As we will see in Section 4, the US tax system lies within the range that satisfies inequality (9), so that there are *prima facie* reasons for suspecting that the level of automation may be excessively high in the US economy, as in this proposition.

One common intuition is that when confronted with a tax system with distortions, $\{\tau^k, \tau^\ell\}$, the best policy is to redress these tax distortions directly. We next show that this is not always the case. In particular, if for other reasons taxes on labor cannot be

time (provided that we are below the peak of the Laffer curve, meaning that tax revenues cannot be increased by lowering both taxes). Thus, inequality (9) is sufficient for $\tau^\ell > \tau^{\ell,r}$ and $\tau^{k,r} > \tau^k$.

[9]If automation decisions were constrained by available technology (i.e., θ had to be less than some $\bar{\theta} < 1$ as in Acemoglu & Restrepo, 2018, 2019a), we could have that $f_\theta(k,\ell;\theta^m(k,\ell)) > 0$ if $\theta^m(k,\ell) = \bar{\theta}$. In this case, productivity gains from automating marginal tasks could be positive. If they were sufficiently large, then automation would no longer be a so-so-technology and Proposition 2 would not apply.

reduced below a certain threshold (which we denote by $\bar{\tau}^\ell$), then the tax system satisfies inequality (9) and is biased against labor, but this does not necessarily imply that capital taxes should be increased. Rather, constrained optimal policy calls for a reduction in the equilibrium level of automation and may even involve a *lower* tax on capital. Before presenting this result, let us note that in this case we are imposing $\tau^\ell \geq \bar{\tau}^\ell$, which can be expressed as an additional constraint on the Ramsey problem (7) of the form

$$\nu'(\ell) \leq (1-\bar{\tau}^\ell)\cdot(1-\varrho)\cdot f_\ell, \tag{10}$$

where the lower bound on labor taxes translates into an upper bound on the marginal disutility from work. In the next proposition, we denote the multiplier on this constraint by $\gamma^\ell \cdot \ell \geq 0$ (where the ℓ simply normalizes the multiplier and makes it easier to interpret).

Proposition 3 ***(Excessive automation with tax distortions)*** *Consider the constrained Ramsey problem of maximizing (7) subject to the additional constraint $\tau^\ell \geq \bar{\tau}^\ell$, and suppose that in the solution to this problem (10) binds. Then the constrained optimal taxes and allocations are given by*

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

$$\frac{\tau^{k,c}}{1-\tau^{k,c}} = \frac{\mu}{1+\mu}\frac{1}{\varepsilon^k(k)} - \frac{\gamma^\ell}{1+\mu}\cdot(1-\bar{\tau}^\ell)\cdot(1-\varrho)\frac{f_{\ell k}\cdot\ell}{u'-1}, \tag{11}$$

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

Before discussing the implications of this proposition, we explain the meaning of constraint (10). The fact that this constraint is binding means that the planner would have chosen a tax rate on labor $\tau^{\ell,r}$ below $\bar{\tau}^\ell$, but the constraint forces the planner to set a higher tax on labor of $\bar{\tau}^\ell$, which results in a tax system biased against labor and in favor of capital (or in other words, inequality (9) holds). This also implies that the level of employment is below what the planner would have chosen in the unconstrained Ramsey problem.

Given this biased tax system, the planner wants automation to be less than its equilibrium level. The intuition is identical to that in Proposition 2: the reduction in automation creates a second-order productivity cost but a first-order gain via its impact on increased employment. Importantly, this holds even when capital taxes can be freely adjusted.

Moreover, the optimal capital tax formula in (11) has an additional negative term on the right-hand side relative to (8). This negative term can lead not just to lower capital taxes than in the unconstrained Ramsey problem in Proposition 1, but even to capital subsidies.[10] The combination of lower capital taxes and limiting the set of tasks that are automated ensures that capital gets used intensively in tasks that are (and should be) automated, while avoiding the displacement of workers from marginal tasks. Both of these margins contribute to raising employment, which increases welfare when the tax system is biased against labor. This is related to the discussion of deepening of automation in Acemoglu & Restrepo (2019a): deepening of automation, which means an increase in the use or productivity of capital in tasks that are already automated, is always beneficial for labor. What is potentially damaging to labor is the extensive margin of automation—because this displaces workers from tasks they were previously performing. Proposition 3 builds on this logic: the planner would like to reduce the range of tasks that are automated by reallocating marginal tasks back to labor and may also want to reduce capital taxes or even subsidize capital at the same time, so that automated tasks can use capital more intensively.

Proposition 3 focused on the case with a lower bound on labor taxes. An equally plausible case is one where, because of political influence of capital owners or because of concerns about capital flight, there is an upper bound on capital taxes.[11] Proposition A.2 in the Appendix establishes that in this case too the planner would like to reduce automation below its market level, even if taxes on labor can be adjusted. The intuition is similar: the upper bound on capital taxation leads to a tax system biased in favor of capital and against labor, and this makes the displacement of labor by capital in marginal tasks socially costly.

2.5 Implementation

To ease exposition, we have so far assumed that the planner can directly control θ. We now discuss how the desired level of θ can be implemented via taxes. Recall that $k(x)$ is the capital used in task x, and so far we have assumed that all types of capital are taxed at the same uniform rate, τ^k. In practice, taxes vary by type of capital (e.g., equipment, software, structures) and industry (because of differential depreciation allowances). In the context of our model, this can be viewed as a task-specific capital tax rate of $\tau^k(x)$. The next proposition establishes when such task-specific capital tax rates are useful and in the process further clarifies the nature of optimal policy interventions.

[10] This might at first appear surprising, especially because the program in Proposition 1 is convex, so moving in the direction of the unconstrained optimum should be beneficial. However, convexity is in the space of allocations and does not imply convexity in the space of taxes. Therefore, increasing the tax rate on capital towards $\tau^{k,r}$ is not necessarily welfare-improving.

[11] A similar constraint on capital taxation is used in the optimal taxation literature (Chamley, 1987; Judd 1999; Straub & Werning, 2020).

Proposition 4 ***(Automation tax)*** *Suppose the planner can set task-specific capital taxes and cannot directly control automation decisions. Then:*

1. *Under the conditions of Proposition 1, the planner sets a uniform capital tax rate, i.e.,* $\tau^k(x) = \tau^k$.

2. *Under the conditions of Proposition 3, the planner prefers to depart from uniform capital taxation. In particular, she can implement the level of automation* $\theta^c < \theta^m(k, \ell)$ *with the following tax scheme:*

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

   where $\tau^A > 0$ *is a task-specific automation tax.*

The reason (unconstrained) optimal policy has no use for task-specific taxes is intuitive: in the unconstrained Ramsey problem, there is no need to distort equilibrium automation decisions. However, in the presence of additional constraints, the planner would like to reduce automation to $\theta^c < \theta^m(k, \ell)$, and she can achieve this by imposing an incremental tax to capital used in tasks above θ^c. By design, these incremental taxes encourage the use of capital in tasks where capital has a comparative advantage (which helps labor via complementarities across tasks) and discourages the automation of marginal tasks (which also benefits labor by preventing its displacement). In what follows, we refer to the incremental tax on capital τ^A as an automation tax.

3 The US Tax System

In this section, we first introduce the notion of *effective taxes* on capital and labor. Effective taxes summarize the average distortion that the US tax system introduces in the use of capital and labor. We then provide formulae for effective taxes that take into account the various elements of the US tax code and their interaction with the type of financing and ownership structure of the firm making investment decisions.

3.1 Defining Effective Taxes on Capital

In our framework, τ^k is the effective tax on (the use of) capital. It is defined as the wedge that the tax system introduces between the internal rate of return for a firm investing in capital and the after-tax rate of return paid to investors. The US tax system includes several taxes, not just a single effective tax on the use of capital. We have personal income taxes on capital income, corporate income taxes, depreciation allowances and many other instruments that contribute to taxes on different types of capital. Moreover, these taxes vary by form of organization (C-corporation vs. passthrough) and type of financing (equity vs. debt).[12]

We start by providing formulae for effective taxes on the use of capital by type of asset, j, form of organization and type of financing. To simplify the exposition, we assume the economy is in steady state—the capital-labor ratio remains constant, the tax system is not expected to change, the price of capital goods changes at a constant rate $\pi^j = q_t^j / q_{t-1}^j$ and the capital stock of type j depreciates at a constant rate $\delta^j > 0$.

The internal rate of return of investing one dollar in equipment j at time $t-1$ is given by

$$r^{f,j} = \mathrm{mpk}^j - \tilde{\delta}^j,$$

where mpk^j is the marginal product of investing one dollar in asset j and $\tilde{\delta}^j = 1 - \pi^j \cdot (1-\delta)$ denotes the total depreciation of the asset (inclusive of investment price changes). Let us denote the after-tax steady-state rate of return to investors by r. The effective tax rate on capital of type j, $\tau^{k,j}$, can then be defined as

$$\frac{1}{1-\tau^{k,j}} = \frac{r^{f,j}}{r} = \frac{\mathrm{mpk}^j - \tilde{\delta}^j}{r}. \tag{12}$$

This formula aligns closely with the effective capital taxes in our conceptual framework in the previous section. In particular, in equation (4), $\frac{1}{1-\tau^k}$ is equal to the wedge (ratio) between the return to the firm from using capital ($\mathrm{mpk}^j - \tilde{\delta}^j$ here and given by $f_k - \delta$ in equation (4)) and the return demanded by investors (r here and $u'(\bar{y}-k) - 1$ in equation (4)).[13]

The computation of effective tax rates requires measuring the marginal product of capital. We follow Hall and Jorgenson (1967) and back out the marginal product of capital using a representative firm's first-order condition for investment. We need to distinguish between C-corporations and passthrough businesses as well as the source of financing, since each of these combinations implies a different first-order condition for investment as well as a different set of taxes on the income generated from capital.

[12]Passthrough organizations include both S-corporations and other passthroughs, such as sole proprietor businesses and partnerships, and are subject to different tax rules as we explain below.

[13]An alternative is to use a formula for effective taxes based on gross returns: $\frac{1}{1-\tau^{k,j}_{\text{gross}}} = \frac{\mathrm{mpk}^j}{r+\tilde{\delta}^j}$. All of our results can be expressed in terms of gross returns, but this would require adjusting the empirical estimates of capital supply elasticities, which are in terms of net returns.

For C-corporations that finance their investment with equity, the first-order condition is

$$\text{mpk}^j = \frac{1 - \alpha^j \cdot \tau^c}{1 - \tau^c} \cdot \left(r^e + \tilde{\delta}^j \right), \tag{13}$$

where τ^c is the corporate income tax rate and $\alpha^j \in [0,1]$ are discounts from depreciation allowances, which reduce taxable income and are discussed in the next subsection. In the absence of corporate income taxes, this expression is identical to the standard user cost formula. In addition, r^e is the pre-tax return to equity holders. This implies that $r = r^e \cdot (1 - \tau^{e,c})$, where $\tau^{e,c}$ is the tax rate on income resulting from ownership of public equity.

Combining the formula for effective taxes in (12) with the first-order condition for investment in equation (13), the effective tax rate for an equity financed C-corporation is

$$\frac{1}{1 - \tau^{k,j}_{\text{c-corp,equity}}} = \frac{1}{1 - \tau^{e,c}} \cdot \left(\frac{r^e + \tilde{\delta}^j}{r^e} \frac{1 - \alpha^j \cdot \tau^c}{1 - \tau^c} - \frac{\tilde{\delta}^j}{r^e} \right). \tag{14}$$

The formula shows that the effective tax on capital depends on the taxation of capital income of equity owners, corporate income tax rates and depreciation allowances. It reiterates that depreciation allowances can significantly offset corporate taxes. For example, with full (immediate) expensing, which corresponds to $\alpha^j = 1$, we would have $\tau^{k,j}_{\text{c-corp,equity}} = \tau^{e,c}$.

The main difference for passthrough businesses is that these organizations do not pay the corporate income tax and are only subject to personal income taxation. Depreciation allowances in this case lower personal income tax obligations for business-owners. The formula for the effective tax on the use of capital for a passthrough business that is financing its investment with (private) equity is

$$\frac{1}{1 - \tau^{k,j}_{\text{passthrough,equity}}} = \left(\frac{r^e + \tilde{\delta}^j}{r^e} \frac{1 - \alpha^j \cdot \tau^{o,p}}{1 - \tau^{o,p}} - \frac{\tilde{\delta}^j}{r^e} \right), \tag{15}$$

where $\tau^{o,p}$ denotes the individual tax rate on the income of owners of passthrough businesses. Note again that with immediate expensing ($\alpha^j = 1$), we have $\tau^{k,j}_{\text{passthrough,equity}} = 0$.

We next turn to debt-financed investments, which allow a further tax discount by subtracting interest payments from taxable income. The presence of these additional tax discounts modifies the first-order condition for investment to

$$\text{mpk}^j = \frac{1 - \alpha^j \cdot \tau^c}{1 - \tau^c} \cdot \left(r^b \cdot (1 - \tau^c) + \tilde{\delta}^j \right), \tag{16}$$

where r^b is the return offered to bond-holders and $r^b \cdot (1 - \tau^c)$ incorporates the lower tax liabilities (which is multiplied by the corporate income tax rate τ^c faced by C-corporations). Note that the after-tax return to households that own bonds is $r = r^b \cdot (1 - \tau^{b,c})$, where $\tau^{b,c}$ is the personal income tax rate for capital income from C-corporation bonds.

Combining the formula for effective taxes in equation (12) with the first-order condition for investment in equation (16), the effective tax rate for a debt-financed C-corporation is

$$\frac{1}{1 - \tau^{k,j}_{\text{c-corp,debt}}} = \frac{1}{1 - \tau^{b,c}} \cdot \left(\frac{r^b \cdot (1 - \tau^c) + \tilde{\delta}^j}{r^b} \cdot \frac{1 - \alpha^j \cdot \tau^c}{1 - \tau^c} - \frac{\tilde{\delta}^j}{r^b} \right). \tag{17}$$

The effective tax on capital again depends on the personal income tax rate of bond-holders, corporate income tax rates, interest rate deductions and depreciation allowances. The additional tax discounts can easily lead to a net subsidy to the use of capital. In particular, with full expensing ($\alpha^j = 1$), we have $\tau^{k,j}_{\text{c-corp,debt}} \approx \tau^{b,c} - \tau^c$, which is negative if bond-holders face lower individual tax rates than corporations.

Owners of passthrough businesses can also subtract their interest payments on debt from their taxable income. However, if they issue bonds, payments to bond-holders are subject to personal income taxation. The formula for the effective tax on the use of capital for a passthrough business that is financing its investment with debt is similar to that of a C-corporation and given by

$$\frac{1}{1 - \tau^{k,j}_{\text{passthrough,debt}}} = \frac{1}{1 - \tau^{b,p}} \cdot \left(\frac{r^b \cdot (1 - \tau^{o,p}) + \tilde{\delta}^j}{r^b} \frac{1 - \alpha^j \cdot \tau^{o,p}}{1 - \tau^{o,p}} - \frac{\tilde{\delta}^j}{r^b} \right), \tag{18}$$

where $\tau^{b,p}$ denotes the individual income tax rate applying to holders of passthroughs' bonds. As before, with full expensing ($\alpha^j = 1$), we would have $\tau^{k,j}_{\text{passthrough,debt}} \approx \tau^{b,p} - \tau^{o,p}$, which is negative if bond-holders face lower income taxes than owners of passthrough businesses.

3.2 Computing Effective Taxes on Capital

We compute effective taxes for equipment, software and structures separately. For each type of capital good, we calculate effective taxes by form of organization and type of financing, and aggregate these taxes into a single effective tax rate for the relevant type of capital using investment shares as weights. The Appendix details the sources and numbers used in our calculations. Here we outline the computations of the key objects in our formulae for effective taxes on capital: depreciation allowances, α^j; corporate income taxes and taxes on owners of equity and passthroughs; and interest rates, economic depreciation and investment prices.

Depreciation allowances: The tax discount term, α^j, is equal to the present discounted value of depreciation allowances associated with one unit of capital purchased at time t, which can be computed as

$$\alpha^j = d_0^j + \sum_{s=0}^{\infty} d_{s+1}^j \cdot \prod_{\tau=0}^{s} \frac{1 - d_\tau^j}{1 + r}, \tag{19}$$

where d_s^j denotes the fraction of the investment that a firm is allowed to subtract from its tax liabilities s years after the purchase.

One useful benchmark is when firms can subtract the economic depreciation of their capital goods each period. In the above formula, this means $d_0^j = 0$ and a constant depreciation rate of δ^j from there on, which adds up to an allowance of $\tilde{\alpha}^j = \delta^j / (\delta^j + r) < 1$.

The IRS and the US tax code handle depreciation allowances quite differently from this benchmark, however. The way in which depreciation allowances are determined is specified in IRS Publication 946. The current system places each type of capital under a specific *class life*—the number of years that a new unit of capital lasts for tax purposes—based on its characteristics and sector. The first reason why tax discounts α^j differ from the one given by constant economic depreciation, $\tilde{\alpha}^j$, is that the depreciation rate implied by a class life is different from the economic depreciation rate.

A second source of an additional tax discount is that the tax code requires taxpayers to follow specific depreciation schedules and enables front-loading of allowances. When computing their tax discount, firms may use a combination of *straight-line* and *declining-balance* methods that yields the highest possible discount. The straight-line method allows firms to expense a constant fraction of their *initial* investment (or undepreciated investment in the *initial* year in which the method is applied) for each year of remaining tax life. The declining-balance method can be used for assets with a class life below 20 years, and allows firms to front-load their depreciation allowances by expensing a decreasing fraction of their initial investment each year. Assets in a class life of 10 years or less can be depreciated using a 200% declining-balance rule, which allows firms to expense their undepreciated investment at two times the rate prescribed by the straight-line method (2×10% for an asset in a class life of 10 years). Firms can then switch to the straight-line method near the end of the asset life to maximize their allowances.[14] Assets with a class life between 10 and 20 years, on the

[14] As an example, consider the allowances generated by the purchase of a machine with a class life of 10 years. Suppose the purchase takes place in the middle of the year. The straight-line method allows a deduction of 5% of the cost in the first year, 10% for the following nine years, and 5% on the eleventh year. The 200% declining balance method gives an allowance of 10% in the first year (two times the straight-line rate of 5%), 18% in the second year (two times the straight-line rate of 10% times the undepreciated stock, 90%), 14.4% in the third year (two times the straight-line rate of 10% times the undepreciated stock, 72%).

other hand, can be depreciated using a 150% declining-balance rule, while assets with a class life of more than 20 years adhere to the straigh-line method.

The third and final source of large discounts from depreciation allowances is recent changes in legislation, passed as part of economic stimulus plans, which introduced *bonus depreciation*.[15] Under current bonus depreciation provisions, most capital with a class life below 20 years enjoys a 100% bonus depreciation, meaning that investors can immediately expense their capital purchases as current costs. This full expensing yields the maximum discount of $\alpha^j = 1$.[16]

We compute α_t^j for 1980–2018 for each type of capital taking into account changes in the treatment of depreciation allowances and bonus depreciation programs (excluding the further reductions in effective capital taxes generated by the 2017 tax reforms, since these did not affect the automation decisions throughout the 2010s). When computing α_t^j, we assume that firms anticipate no future changes in the tax code, so that they expect current rates to apply in the future.[17]

Figure A.1 in the Appendix plots $\tilde{\alpha}^j$ and α^j for software, equipment, and non-residential structures. The figure show that α^j typically exceeds $\tilde{\alpha}^j$ for software and equipment, and that recent bonus depreciation provisions generated an increase in allowances bringing α^j close to 1 for software and equipment in the 2010s.

Tax rates on corporations and capital owners: Effective taxes on capital also depend on taxes on corporations and the households who own capital. We approximate the average marginal corporate income tax rate τ_t^c for each year as the average tax paid by

This continues up to year 7, where the method prescribes an allowance of 5.89%, which is below the straight-line method allowance of 6.55% computed on the undepreciated stock of capital and 4.5 years of useful life left. Therefore, the schedule for 10-year property follows the 200% declining-balance method until year 7 and switches to a constant allowance of 6.55% of the undepreciated cost for the remaining 4.5 years. For further discussion and examples, see the Appendix in House and Shapiro (2008).

[15] In particular, the "Job Creation and Worker Assistance Act of 2002" (JCWAA) introduced a 30% bonus depreciation for 2002-2003; the "Jobs and Growth Tax Relief Reconciliation Act of 2003" (JGTRRA) raised the bonus to 50% for 2004; the "Economic Stimulus Act of 2008" introduced a 50% bonus, extended until 2017 by successive bills; the "Tax Relief, Unemployment Insurance Reauthorization, and Job Creation Act of 2010" temporarily raised the bonus to 100% (full expensing) between September 2010 and the end of 2011. Finally, the "Tax Cuts and Jobs Act of 2017" raised bonus depreciation to 100% for 2018-2022.

[16] A 100% bonus depreciation corresponds to $d_0 = 1$ and $d_s = 0$ for all $s > 0$ in equation (19). As stated above, capital allowances are generally set by the schedules in Publication 946, which give a specific d_s^j for all s, j, such that $\sum_{s=0}^{T_j} d_s^j = 1$, for each investment type j, and where T_j is the class life for the capital type j. When bonus depreciation is $\gamma < 1$, the taxpayer obtains a first-year bonus allowance equal to γ and then follows the schedules for depreciation allowances for the undepreciated capital stock. Therefore, the bonus allowance series, $\tilde{d}_s^j$, has $\tilde{d}_s^j = (1-\gamma)d_s^j$, for all $s \geq 1$, and $\tilde{d}_s^j = \gamma + (1-\gamma)d_0^j$ in the initial period.

[17] Anticipated tax reforms create a "reevaluation" effect for capital that is already installed.

C-corporations:

$$\tau_t^c = \frac{\text{corporate tax revenue}}{\text{net operating surplus of C-corporations}}.$$

The corporate tax revenues are obtained from NIPA Tables. The computation of the tax base is presented in the Appendix. We start with operating surplus from corporations and subtract depreciation allowances. We then allocate a fraction of these profits to C-corporations using data from the IRS on profits by type of corporation. The remaining share is accounted for by S-corporations which do not pay corporate income taxes, and this share is not included as part of the tax base in the above calculation. The share of corporate profits generated by C-corporations has fallen over time from 93% in 1980 to 61% in 2018, in line with the findings Smith et al. (2019). Our calculations show that once we account for this changing share, the average tax rate on C-corporations increased from 25% in 1981 to 35% in 2000, and then declined to 17.5% in 2018.

Note that we are computing corporate income taxes as an average of the taxes paid, rather than by using the statutory rate (46% in 1981, 35% in the intervening years, and 21% in 2018). This is because many corporations pay less corporate income tax than implied by the statutory rate. Throughout, we interpret average taxes as averages of marginal tax rates faced by different types of firms.

Besides taxes paid by corporations, taxes paid by households on their capital income from equity and lending also contribute to the effective tax on the use of capital (the terms $\tau^{e,c}$, $\tau^{b,c}$ and $\tau^{b,p}$ in equations (14), (17) and (18)). We compute $\tau^{e,c}$ as the average tax rate paid by owners of equity on their dividends and capital gains. We start by computing the share of corporate equity that is directly held by US households and is thus subject to taxation. Using data from the Board of Governors of the Federal Reserve System, we approximate this as the share of corporate equity owned by US households and nonprofit organizations serving these households, which has fallen from 58% in 1981 to 37% in 2018. We follow CBO (2014) and assume that the remaining share is owned by funds or kept in accounts that are not subject to additional taxation.

Taxes paid by households depend on how corporate profits are realized. Qualified dividends or capital gains are taxed at a *maximum capital gain tax rate* specified by the IRS.[18] These include dividends on stocks held by more than 61 days, or capital gains on stocks owned for over a year. Ordinary (non-qualified) dividends or capital gains apply to stock owned over shorter periods and are taxed at the same rate as individual income. The re-

[18]The maximum capital gain tax rate is specified in IRS publication 550. In 2018, taxpayers facing a marginal tax rate below 15% had a maximum capital gain rate of 0%. Taxpayers facing a marginal tax rate between 22% and 35% had a maximum capital gain tax rate of 15%. Finally, taxpayers facing a marginal tax rate of 35% faced a maximum capital gain tax rate of 20%.

maining profits are for stocks held until death, whose capital gains are never realized and thus face no taxation. We compute the share of profits realized through ordinary dividends and short-term capital gains by using data from the IRS Individual Complete Report (Publication 1304, Table 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. Short-term dividends and ordinary capital gains account for the bulk of realized profits from C-corporations (about 60% in recent years). The remaining share of profits is accounted for by long-term qualified gains and dividends, or by stocks held until death whose capital gains are never realized. We assume that each of these two forms makes up an equal share of profits, which aligns with what the CBO reports for 2011.

The average tax rate on profits derived from C-corporation profits (after paying corporate taxes) is thus given by

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

Here, 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. Both average taxes are computed using data from the Office for Tax Analysis for 1980–2014. In recent years, the average tax rate on ordinary short-term gains and dividends was $\tau_t^o = 24\%$ and the average tax rate on long-term qualified capital gains and dividends was $\tau_t^q = 18\%$. Our estimates show that $\tau_t^{e,c}$ has hovered around a historical average of 15% and experienced a temporary reduction to 12.5% during the 2000s.

Turning to taxation of rental income for bond-holders, the CBO estimates that 52.3% of C-corporation bonds are held directly by households, 14.9% generate income that is temporarily deferred for tax purposes, and the rest is held by funds or kept in accounts that are not subject to additional taxation. For passthrough entities, the share owned by households is larger, 76.3%, and the share deferred is 10.1%. Moreover, the CBO reports that the rental income owned by households is subject to personal income taxes at the average rate 27.4% in 2014. Supposing that temporarily deferred income is subsequently taxed at the same rate as the rest of rental income, we estimate the average tax paid by bond-holders on their rental income from C-corporations and passthroughs, respectively, as $\tau^{b,c} = 16.84\%$ and $\tau^{b,p} = 23.25\%$, and assume that these rates have remained constant over time.

The final item required for our calculations is the tax rate paid by owners of passthroughs, which we separate into S-corporations and other passthroughs (sole proprietor businesses and partnerships). Profits from S-corporations are taxed at the individual income rate of the owners. We assume that the average tax rate paid by owners of S-corporations is the same as the average tax paid by individuals earning ordinary short-term dividends and capital gains, τ_t^o.[19] In economic terms, this requires owners of S-corporations to have a similar income profile as investors in public equity. In addition, part of the profits generated by S-corporations accrue only when the company is sold, and these profits are taxed at the maximum qualified rate, τ_t^q. Thus, we measure the average tax paid by owners of S-corporations on their profits as

$$\tau^{o,s} = \tau_t^o - \text{share capital gains} \cdot (\tau_t^o - \tau_t^q).$$

Using data on sales of passthrough businesses reported by the IRS for 1990–2000, we estimate the average share of capital gains in S-corporation profits as 25%, and assume it has remained at this level over time. Our estimates imply that $\tau_t^{o,s}$ has been roughly constant as well, at about 27%, reaching 28% in 2018. Since self-proprietors' and partnerships income is reported as personal income, we have no data on the tax rate faced by owners on profits, and so we assume that they face the average tax rate on income (obtained from the IRS, SOI Tax Stats), which has been approximately 14.6% in recent years.

Overall, our estimates imply that in 2011 the average corporate income tax was 26.4% (with equity holders paying an additional 11.8% on top of this), the average tax rate paid by S-corporation owners was 23%, and the average tax rate paid by owners of other passthroughs was 14.6%. These numbers align closely with those by the CBO and Cooper et al. (2016).[20]

Interest rates, depreciation and investment prices: We assume a constant interest rate, a constant growth rate for investment prices and a constant rate of economic depreciation for each asset that match historical averages from 1981 to 2017. We use a constant value of $r^b = 4.21\%$ per annum for bond-holders, given by the average of the Moody's Seasoned AAA Corporate Bond Yield minus realized inflation between 1981 and 2017. Likewise, we use a constant value of $r^e = 4.36\%$ per annum for equity-holders, which is the historical average of the real rate of return on the S&P 500 over 1957–2018. The constant growth rate for

[19] Profits from S-corporations are also taxed as corporate income by some states. To account for these taxes, we add the average state and local tax rate on businesses, which we compute by dividing state and local revenues from business taxes by the net operating surplus of corporations. State and local taxes on businesses are small in practice, with an average value near 3% in recent years.

[20] Using IRS data, Cooper et al. estimate that in 2016 C-corporations paid an average tax rate 23% (plus 8.25% on the household side), S-corporations paid an average tax rate of 25% and other passthroughs paid an average tax rate of about 14.7%.

investment prices is estimated from the average change of investment price indices by type of capital from the BEA Fixed Asset Tables (FAT) between 1981 and 2017. These imply an annual average growth rate of prices equal to -1.6% for software, -1% for equipment, and 2% for non-residential structures. The economic depreciation rates, the δ_t^j's, are taken directly from the BEA FAT as the averages for 1981-2017 (the average economic depreciation rate per annum is 23.4% for software, 13.9% for equipment, and 2.6% for non-residential structures).

3.3 Effective Taxes on Labor

In our model, τ^ℓ is the effective tax on (the use of) labor. However, as with capital, there is no single tax on labor in the US tax code. Instead, labor income is subject to a number of different taxes both at the federal and local level. Means-tested public programs may generate additional implicit taxes on labor. The effective tax on labor is given by the wedge that the tax system introduces between the marginal product of labor and the before-tax wage, mpl^f. The representative firm will demand labor until the marginal product of labor, mpl^f, equals the cost of one unit of labor given by total compensation. That is,

$$\text{mpl}^f = \text{compensation} = \text{salary} + \text{benefits}.$$

Wage income is subject to personal income tax at a rate τ^h, and payroll taxes at a rate τ^p. Benefits are not taxed, but might be imperfectly valued by workers, which we capture by converting them to an income-equivalent amount by multiplying it with $\varphi \in [0,1]$. Consequently, the after-tax return to work for the household is given by

$$w = \text{salary} \cdot (1 - \tau^h - \tau^p) + \text{benefits} \cdot \varphi.$$

The effective tax rate on labor is defined, analogously to the effective tax on capital, as

$$\frac{1}{1-\tau^\ell} = \frac{\text{mpl}^f}{w} \Rightarrow \tau^\ell = \frac{\text{salary} \cdot (\tau^h + \tau^p) + \text{benefits} \cdot (1-\varphi)}{\text{compensation}}.$$

We measure the terms in this expression as follows. From national accounts we obtain data on salaries and total compensation for the corporate sector. We treat employers' contributions to pensions and health insurance as part of the benefits since these are not taxed. We assume that workers outside the corporate sector receive a similar split between benefits and salaries and are therefore subject to the same effective taxes. We use a payroll tax rate of 15.3%, which is the statutory rate that applies to all earners with an income below $132,900 dollars in 2018 (a level that roughly matches the 95th percentile of income). Since the vast majority of jobs at-risk of automation are performed by workers in the middle of the income distribution, the payroll tax of 15.3% is relevant for automation decisions and is incorporated in our effective tax rate on labor. We measure the personal income tax rate τ^h, consistently with our treatment of payroll taxes, as the average income tax paid by earners below the 95th percentile. This is computed from publicly available data from the IRS for 1986–2017. The estimate for τ^h has been stable in recent years at a level close to 10%.[21] Finally, we use a value of $\varphi = 0.65$ building on estimates from Gruber and Krueger (1991), Goldman et al. (2005) and Lennon (2019), which suggest that one dollar of spending on benefits is valued on average at 65 cents by households. This increases our estimates for τ^ℓ by 3%.

Besides our baseline estimate for τ^ℓ described above, in the Appendix we present results using an estimate for the effective tax on labor which incorporates the implications of means-tested welfare programs. In particular, there is a range of programs, including cash transfers and credits, that are phased out as individual income increases, and various social programs (such as disability insurance and unemployment insurance) in which individuals participate less when labor demand is high (see for instance the evidence in Autor & Duggan, 2003; Autor, Dorn & Hanson, 2013; and Acemoglu & Restrepo, 2020a). As a consequence, transfers decline as labor demand rises, and this acts as an additional implicit tax on labor, τ^d. Austin, Glaeser and Summers (2018) estimate that the public expenditures resulting from a person going into non-employment was of \$4,900 per year between 2010–2016 (\$6,300 for those in long-term non-employment and \$2,300 for the short-term unemployed). This is roughly 8% of the average yearly worker compensation during this period, suggesting that social expenditure and disability insurance add an extra 8% tax to labor.

3.4 Effective Tax Rates in the US

Figure A.2 in the Appendix depicts the evolution of the average personal income tax and average capital tax rates for C-corporations (including both corporate income taxes and personal income taxation) and for S-corporations (whose owners only pay personal income taxes and some state-level taxes). Taxes on C-corporations' profits decline significantly from 2000 onwards, reflecting declines in the statutory corporate income tax rate over time. Taxes on passthrough profits have remained stable around 25% and the average individual income tax has remained close to 15%.

Figure 1 presents our estimates for the effective tax rates on labor and different types of capital (in turn computed from effective tax rates on capital and depreciation allowances for C-corporations, S-corporations and other passthrough businesses, and the differential

[21] If we were to use the average payroll tax (about 10% in recent years) and the average income tax (about 14.6% in recent years), we would end up with a very similar effective tax rate on labor.

taxation of capital financed with debt and equity). The solid lines show the effective taxes on software, equipment, non-residential structures and labor.

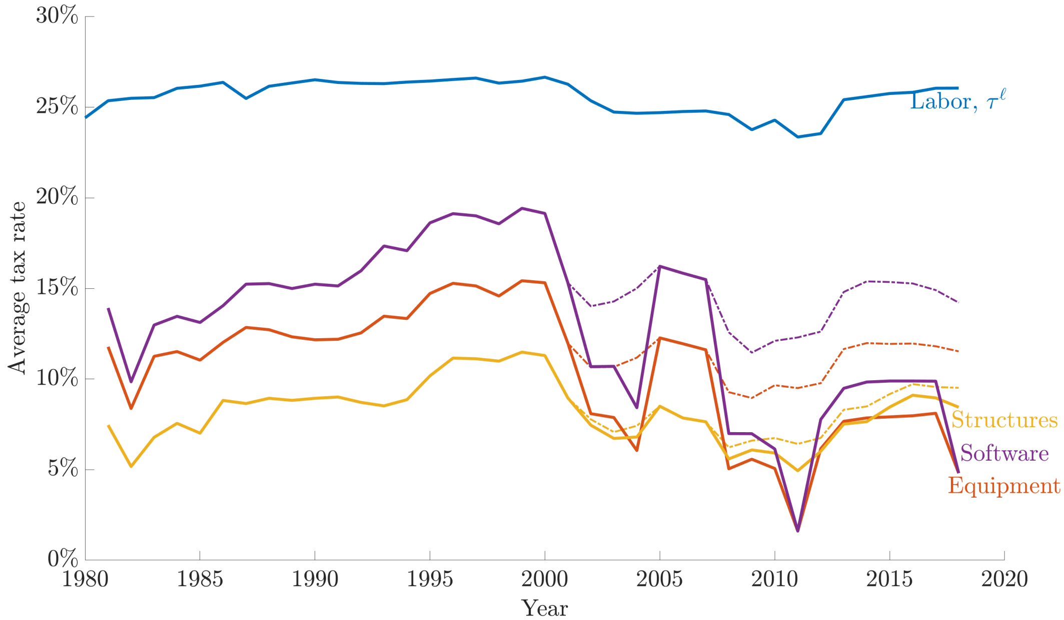


Figure 1: Effective tax rates on labor, software capital, equipment, and non-residential structures.

Notes: The solid lines depict the observed effective taxes. The dashed lines present the effective taxes that would result if the treatment of allowances had remained as in the year 2000. See the text for details.

Several points about these effective tax rates are worth noting. First, effective taxes on equipment and software are low compared to the effective taxes on labor. Our benchmark effective tax on labor (which does not include the implicit taxes implied by means-tested programs) hovers around 25.5%.[22] In contrast, effective taxes on both equipment capital and software in the 2010s (and before the tax reform of 2017) are around 10%.[23] Second, effective taxes on equipment and software were higher in the 1990s and early 2000s, and declined significantly thereafter. This decline is mostly because of the reforms summarized in footnote 15, which have increased depreciation allowances. The dashed lines illustrate the contribution of these reforms by plotting the (counterfactual) effective taxes on different types

[22]Our estimates imply that the net tax revenue collected by the government with these instruments is roughly 18.6% of GDP (25.5% × labor income in GDP +10% × net capital income in GDP). This figure matches closely the average share of personal income taxes, corporate taxes and social security contribution in GDP for the period considered in our study (18.7% for 1981–2018 in NIPA Table 3.1).

[23]These effective tax rates are lower than those reported in CBO (2014). Two factors explain the differences. First, and most importantly, the CBO does not incorporate bonus depreciation allowances (based on the argument that these may not be extended in the future). Second, the CBO uses the statutory rate of corporate income tax. As noted above, we do not believe this gives an accurate estimate of the effective tax on capital, since most corporations pay less than the statutory rate.

of capital that would have applied had the treatment of depreciation allowances remained as it was in 2000. They show that about half of the decline in the effective taxes on software and equipment capital is due to the more generous depreciation allowances introduced since 2002. Third, effective taxes on equipment and software decreased further, to about 5%, following the 2017 tax reform, which introduced full expensing of these capital expenditures. Finally, because depreciation allowances for structures are lower, the effective tax on non-residential structures is higher today than tax rates on equipment and software, but in the past the ordering was reversed.

For our purposes, effective tax rates on equipment and software are more relevant, since these are the types of capital that are involved in automation. In what follows, we will summarize the US tax system as an effective tax on labor of $\tau^\ell = 25.5\%$ and an effective tax on capital of $\tau^k = 10\%$ (the level before the 2017 tax reforms). We will also separately discuss the implications of the reforms in the 2000s and the 2017 tax reform.

4 Does the US Tax Code Favor Automation?

In this section, we investigate whether the US tax system is biased against labor and favors excessive automation. We then explore the implications of different tax reforms.

4.1 Parameter Choices

We first review the estimates of the main parameters in our model.

The parameter λ corresponds to the *short-run* elasticity of substitution between capital and labor. This is the elasticity of substitution between capital and labor holding the amount of automation (and more generally the state of technology) constant, and without any compositional changes (for example, between firms with different technologies or between industries). Under the assumption that in the short run the allocation of tasks to factors is fixed, this elasticity can be approximated by the short-run elasticity of substitution within establishments, which is estimated to be $\lambda = 0.5$ in Oberfield and Raval (2014).

The other important building block of the production side of our economy is given by the comparative advantage schedules for labor and capital, $\psi^\ell(x)$ and $\psi^k(x)$. We reduce the dimensions of these functions by assuming that they take iso-elastic forms:

$$\frac{\psi^\ell(x)}{\psi^k(x)} = A \cdot x^\zeta \qquad\qquad \psi^\ell(x) = A \cdot x^{\zeta \upsilon},$$

where $\zeta \geq 0$ controls how the comparative advantage of labor changes across tasks, and υ controls the relationship between the comparative and absolute advantage of labor. We take

$\upsilon = 1$ as our baseline, which implies that labor is more productive at higher-index tasks (where it has a comparative advantage), while capital has a constant productivity across tasks (as in "balanced growth" specification in Acemoglu & Restrepo, 2018). The Appendix explores the opposite case in which $\upsilon = 0$ and labor is less productive in tasks where it has comparative advantage.

The parameter of comparative advantage ζ (together with λ) shapes the long-run substitution possibilities between capital and labor. In the long run, changes in factor prices will lead to endogenous development and adoption of automation technologies, and as the allocation of tasks to factors changes, there will be greater substitution between capital and labor than implied by λ. The extent of this greater substitution is shaped by the comparative advantage of labor across tasks. In particular, since $\lambda = 0.5$, a lower user cost of capital will increase the labor share of national income in the short run (because capital and labor are gross complements given θ), but as automation adjusts, the labor share could end up lower than it was before the change. Karabarbounis and Neiman (2014) estimate that a 10% reduction in the user cost of capital lowers the labor share by 0.83-1.67 percentage points in the long run. The midpoint of this range implies $\zeta = 2.12$ in the context of our model.[24]

Turning to labor market imperfections, recall that the wedge ϱ captures the difference between the wage earned by workers and their opportunity cost. This motivates measuring ϱ as the (average) permanent earning loss from job separation. The majority of the estimates of these earning losses in the labor literature are within the range 5%-25% with a midpoint of 15%.[25] Motivated by this evidence, we choose a baseline value of $\varrho = 0.15$.[26]

The remaining key parameters of our framework are the Hicksian elasticities of labor and capital supply (Hicksian elasticities are the relevant ones in our context because we are focusing on permanent tax reforms). We adopt the following functional forms for utility: $u(\bar{y} - k) = -B \cdot k^{1+1/\varepsilon^k}/(1 + 1/\varepsilon^k) - k$ and $\nu(\ell) = \ell^{1+1/\varepsilon^\ell}/(1 + 1/\varepsilon^\ell)$, so that the two Hicksian elasticities, $\varepsilon^k \geq 0$ and $\varepsilon^\ell \geq 0$, are constant. The parameters A and B are calibrated to match

[24] More specifically, these authors use a constant elasticity of substitution aggregate production function without automation or reallocation of tasks, and show that their estimates correspond to a long-run elasticity of substitution in the $1.2 - 1.5$ range.

[25] Couch and Placzek (2010) survey this literature and present their own estimates, suggesting long-run earning declines from separations of 5%. Jacobson, Lalonde, and Sullivan (1993) find long-run earning declines of about 25%. Davis and von Wachter (2011) report long-run earning losses of 10% in normal times and 20% in recessions.

[26] Some of the earning losses may be due to loss of firm-specific human capital. If productivity gains from firm-specific human capital are shared equally between firms and workers, these would also create a wedge identical in reduced-form to our ϱ.

We also note that there are other factors that would act like a wedge, generating additional incentives to raise employment. These include negative spillovers from non-employment on family, friends and communities and on political behavior (see Austin, Glaeser and Summers, 2019). Because quantifying these effects is more difficult, we are ignoring them in the current paper.

an aggregate labor share of 56% and a net capital share of 26%, with the depreciation rate fixed at 5.5% per year.

Because our model does not distinguish between the intensive (hours conditional on employment) and extensive margin (employment), we use the combined elasticity for total hours of work. Chetty et al. (2011) report *micro elasticity* estimates, obtained from differences in tax rates and wages across regions and demographic groups within a country, in the range 0.46–0.76 (of which 0.33 comes from the intensive margin and 0.13–0.45 comes from the extensive margin). These numbers are close to *macro elasticity* estimates obtained from tax differences across countries, which are also around 0.7. Because there might be non-linearities in supply elasticities (see, for example, Mui and Schoefer, 2019), and there is uncertainty about the exact supply elasticities, we explore the implications of labor supply elasticities between 0.46 and 1 in our robustness checks.[27]

The parameter ε^k corresponds to the long-run elasticity with which the supply of capital responds to changes in net returns $d \ln k / d \ln r$ or the keep rate from net capital taxes $d \ln k / d \ln(1 - \tau^k)$ (and is thus different from the "demand-side" elasticities that are informative about how much investment or capital at the firm level will respond to the user cost of capital). Although there is much uncertainty about this elasticity and many theoretical analyses assume it to be infinite (for example by imposing time-additive, discounted utility), a number of recent papers estimate it to be much smaller. These studies exploit reforms that change taxes on wealth for different groups of households and find medium-run elasticities that range from 0.2 to 0.65 over 4–8 year periods (see Zoutman; 2018, Duran-Cabré et al., 2019; Jakobsen et al., 2020).[28] Using a calibrated life-cycle model and assuming a net after-tax return of $r = 5\%$, Jakobsen et al. show that their medium-run estimates are consistent with long-run elasticities ranging from 0.58 for the wealthy to 1.15 for the very wealthy. With a lower-tax net return of 4% (in line with the numbers used in our computation of net effective taxes), long-run capital supply elasticities would be even lower, and conversely, with an after-tax rate of return of 7%, these elasticities would range between 1 for wealthy households, and 1.9 for the very wealthy (see Table III in Jakobsen et al., 2020). We set our baseline capital supply elasticity to 0.65, which is the average elasticity for the wealthy in Jakobsen et al.'s preferred scenario with $r = 5\%$ and lies at the upper end of the

[27]In the presence of some types of labor market frictions, the extensive margin changes in employment may take place off the labor supply curve, while intensive margin changes are on the labor supply curve. In Table A.1 in the Appendix we show that our main conclusions are robust if we reduce ϱ to 0.075, so that labor market frictions apply only to the extensive margin changes in employment (which make up about half of the variation given the elasticities reported in the text).

[28]These estimates are from small and fairly open economies, such as Denmark, the Netherlands and Catalunya, and thus presumably include the response due to the international mobility of capital.

medium-run elasticities reported above.[29] We explore the robustness of our results to using a higher elasticity of capital supply in the Appendix.

4.2 Is the US Tax System Biased against Labor?

We first verify that the US tax system (with $\tau^\ell = 25.5\%$ and $\tau^k = 10\%$) is biased against labor. The estimated US taxes comfortably satisfy inequality (9) when we use the elasticity estimates in the previous subsection, $\varepsilon^\ell = 0.7$ and $\varepsilon^k = 0.65$.

Inequality (9) implies that current US taxes on labor are too high and US taxes on capital too low relative to the optimum. In fact, the formulae in Proposition 1 for our baseline choice of parameters, imply that optimal taxes should be $\tau^{k,r} = 26.65\%$ and $\tau^{\ell,r} = 18.22\%$, which contrast with the observed taxes of $\tau^k = 10\%$ and $\tau^\ell = 25.5\%$. The optimal tax on labor is lower than on capital because the supply elasticities for the two factors are similar, while there is an additional wedge for labor ($\varrho = 0.15$), which the optimal tax system corrects for.

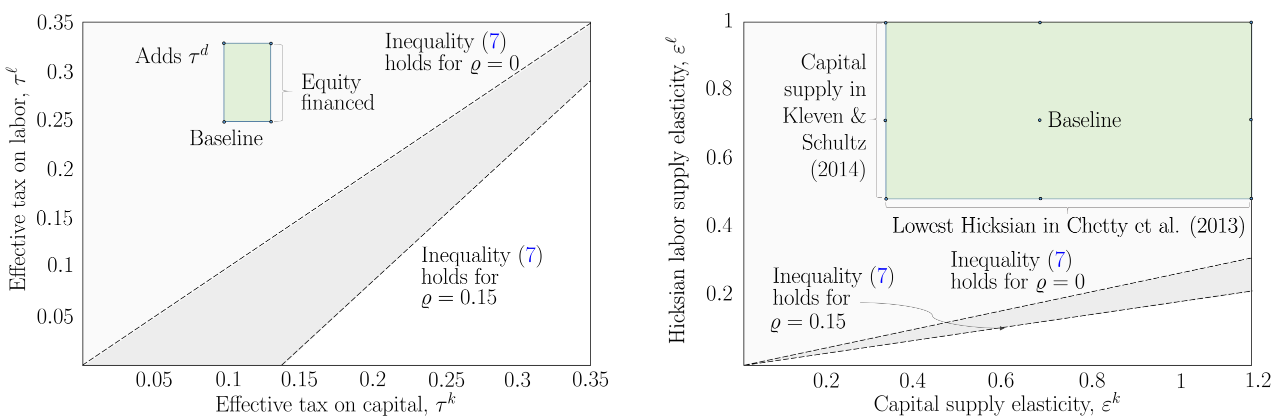


Figure 2: Contour plots of taxes and elasticities that verify inequality (9).
Notes: Panel A shows contour plots for estimates of the current US tax system and Panel B depicts contour plots for labor and capital supply elasticities to verify the robustness of the claim that the US tax system is biased against labor. Green boxes represent the range of estimates we consider in our robustness checks and in each case we separately mark our baseline estimates. Inequality (9) is satisfied for $\varrho = 0$ in the light gray area and for $\varrho = 0.15$ in both the light and the dark gray areas. See the text for details.

[29] We view our choice as conservative given other estimates in the literature. Brülhart et al. (2019) estimate the elasticity of capital to after-tax returns using variation across Swiss Cantons. They find an elasticity of 1.05 but also show that about a quarter of the effects are driven by migration across cantons and do not involve a change in savings—which is the relevant margin for optimal taxation. In their concluding remarks, they argue that once this response is accounted for, their numbers are comparable to the medium-run estimates of Jakobsen et al., (2020). Kleven and Schultz (2014) estimate an elasticity of capital supply with respect to one minus the tax rate on capital income of 0.3, which would imply an even more inelastic response of capital, reinforcing our results. Finally, a related literature finds small elasticities of savings to one minus the estate tax rate, typically about 0.09–0.16 (see Joulfaian, 2006 and Kopczuck and Slemrod, 2001), which also imply less elastic responses of the supply of capital.

The conclusion that the US tax system is biased against labor is robust to variations in our measurement of effective taxes and our estimates of the elasticities of the supply of capital and labor. Panel A of Figure 2 documents that variations in how we compute effective taxes on capital and labor do not change this conclusion. It depicts two contour plots for τ^ℓ and τ^k that satisfy inequality (9) for the baseline values of the remaining parameters ($\varepsilon^\ell = 0.7; \varepsilon^k = 0.65$) and for $\varrho = 0.15$ (the solid line) and $\varrho = 0$ (the dotted line). All of our tax estimates lie within these sets and thus comfortably satisfy inequality (9) regardless of the value of ϱ.

Panel B of Figure 2 documents that the US tax system remains biased against labor when we vary the elasticities for the supply of capital and labor. The figure presents contour plots for combinations of elasticities ε^ℓ and ε^k that satisfy inequality (9) for our baseline estimates of the US tax system ($\tau^\ell = 25.5\%; \tau^k = 10\%$), and again separately for $\varrho = 0.15$ and $\varrho = 0$. We find that even if the capital supply had a unitary elasticity, the US tax system would satisfy inequality (9) and would continue to be biased against labor.

4.3 Implications of the US Tax System for Automation and Employment

As discussed in our theory section, the bias against labor in the US tax system will generate excessive automation and lead to lower employment than socially optimal. We now return to our baseline parameters and investigate the implications of the pro-capital bias of the US tax system for automation, employment, the labor share and welfare.

As a first step, we compare the implied equilibrium level of automation under the tax system in the 2010s (before the 2017 tax reform), $\tau^\ell = 25.5\%$ and $\tau^k = 10\%$, to the equilibrium with optimal taxes, $\tau^{\ell,r} = 18.22\%$ and $\tau^{k,r} = 26.65\%$. Columns 1 and 2 of Table 1 present this comparison. Because the optimal tax system encourages the use of labor in production (relative to the US system in the 2010s), it leads to a lower level of automationthan currently. Under the optimal tax system, θ declines by 4.1% from its equilibrium value in the 2010s.[30] This lower level of automation would also increase the labor share by 0.78 percentage points, and together with the lower labor tax, increase employment by 4.02%. Finally, welfare would be higher by 0.38% in consumption-equivalent terms (meaning that the welfare gains are equivalent to increasing consumption in period 1 by 0.38%). Although this increase in welfare appears small (relative to the change in employment), this is due to the usual intuition related to "Harberger triangles": because changes in welfare are second-order near the optimum, they tend to be smaller than changes in quantities unless we are very far away

[30] Though the magnitude of a change in θ is not directly interpretable, we can compute the share of employment that would be displaced with the higher level of θ. Given our parametrization of λ, $\psi^\ell(x)$ and $\psi^k(x)$, reducing θ from 0.276 to 0.265 results in 3.3% fewer workers displaced by automation.

from this optimum.

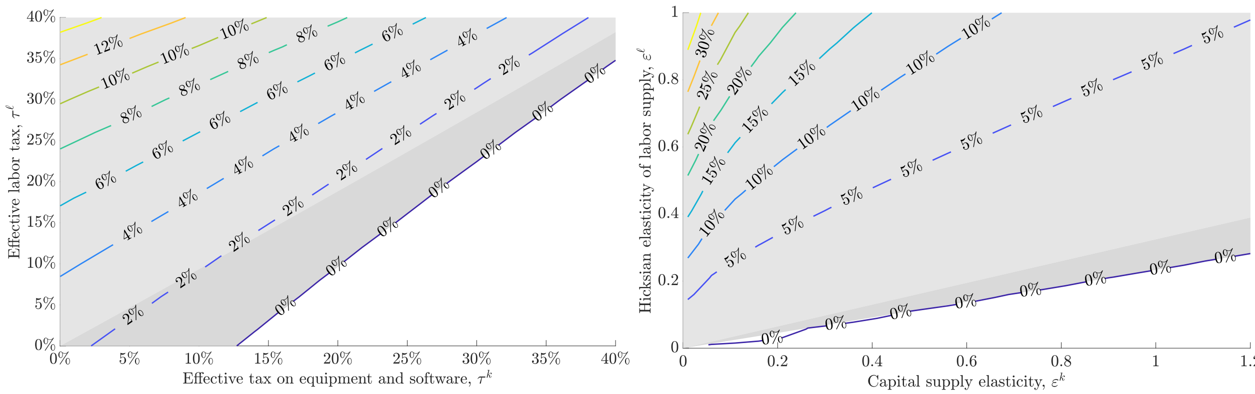
TABLE 1: Equilibrium under the current tax system and under other potential scenarios.

	Current System (1)	Ramsey Solution (2)	Distorting θ (3)	Distorting θ and changing τ^k (4)	Distorting θ and changing τ^ℓ (5)
Tax system:					
τ^k	10.00%	26.65%	10.00%	8.39%	10.00%
τ^ℓ	25.50%	18.22%	25.50%	25.50%	24.89%
θ	0.276	0.265	0.267	0.265	0.264
τ^A	0.00%	0.00%	10.15%	12.90%	13.07%
Aggregates:					
Employment	.	+4.02%	+1.14%	+1.59%	+1.96%
Labor Share	56.00%	56.78%	57.93%	58.44%	58.54%
Net Output	.	+0.44%	−0.10%	+0.16%	+0.20%
C.E. welfare change	.	0.38%	0.09%	0.14%	0.18%
Revenue	.	0.00%	+1.41%	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.

In Table 1, we used an effective tax rate on labor of $\tau^\ell = 25.5\%$, which does not include the additional implicit tax on labor implied by means-tested programs. Table A.2 in the Appendix shows that when we incorporate this additional implicit tax on labor supply and set $\tau^\ell = 33.5\%$, the employment and welfare gains from changing the current system are amplified. Moving to optimal taxes now increases employment by 6.07%, the labor share by 1.09 percentage points and welfare by 0.81%.

The conclusion that the we can achieve higher welfare through tax reforms that raise employment and reduce automation is robust to variations in parameters and the measurement of taxes. Figure 3 considers the same range of taxes and parameters as in the two panels of Figure 2. The contours in this figure correspond to combinations of current tax rates (Panel A) and elasticities (Panel B) that give the same employment response when we switch from the current tax system to optimal taxes. For a wide range of parameters, optimal taxes induce levels of employment that are 2-10% larger than in the current system.



Panel A Panel B

FIGURE 3: CONTOUR PLOTS FOR THE PERCENT CHANGE IN EMPLOYMENT RESULTING FROM A MOVE FROM THE CURRENT TAX SYSTEM TO THE OPTIMAL TAX SYSTEM.
Notes: Panel A is for different measurements of current US taxes, and Panel B is for different combinations of estimates of labor and capital supply elasticities. See the text for definitions and details.

Recall from Proposition 2 that, when the tax system is biased against labor, the level of automation is not only greater than in the Ramsey solution, but it is also excessively high compared to what would be socially optimal *given* the tax system. Column 3 quantifies this inefficiency by computing the level of automation that would maximize welfare taking the current capital and labor taxes as given.[31] The level of automation that maximizes welfare is $\theta = 0.267$, which is 3.3% lower than equilibrium automation. In line with Proposition 4, this lower level of automation can be implemented with an automation tax of 10.15%, so that a task will be automated only if replacing labor with capital reduces the cost of producing that task by more than 10.15%. The automation tax raises employment by 1.14%—partially correcting for some of the inefficiencies in the current system and raising welfare—and the labor share by 1.93 percentage points. Even though equilibrium automation decisions are being distorted, aggregate net output remains essentially unchanged (it declines by 0.10%). As already noted, this is because marginal tasks automated under a biased tax system do not increase productivity much (or the automation technology being used in these tasks is "so-so").

Column 3 allows the planner to change θ, but without modifying the effective tax on capital, τ^k. We next verify that, as implied by Proposition 3, if the planner could also modify τ^k (but could not reduce labor taxes), she would complement any reform with an

[31]The alternative is to follow Proposition 2 and maximize the sum of the representative household's utility plus the change in revenue valued at μ (which is the social value of government funds). Here, we simply maximize welfare—given by the representative household's utility—to make the results in this column comparable to the rest of the table. Valuing additional revenues with the multiplier μ leads to higher automation taxes, since reductions in θ have the additional benefit of generating higher labor tax revenue.

automation tax to reduce automation below its market level. This is illustrated in column 4, which shows that in this case the planner achieves higher welfare through a combination of lower capital taxes (τ^k decreases to 8.39%) and an automation tax of 12.9%, which further reduces θ to 0.265. This alternative tax system would lead to a 2.44 percentage points higher labor share and 1.59% more employment.[32]

Finally, column 5 turns to a setting where the planner can reduce taxes on labor and distort θ, but cannot increase taxes on capital (as mentioned, this scenario may be relevant due to political constraints or fear of capital flight). In this case, the planner would combine a lower labor tax with an automation tax of 13.07%, reducing automation again to $\theta = 0.264$, and increasing employment by 1.96% and the labor share by 2.54 percentage points.[33]

In summary, our quantitative results show that the current tax system inefficiently favors automation and leads to an employment level that is below the social optimum. The best policy would be to set taxes at their optimal levels, which does not require any further distortions to automation. But if optimal taxes were infeasible, then reducing automation, with or without accompanying changes in other taxes, could reverse some of the inefficiencies in the current tax system and increase employment by 1.14–1.96% and the labor share in national income by 1.93–2.54 percentage points.

4.4 Recent Reforms and Effective Stimulus

As described in footnote 15, a series of reforms enacted between 2000 and the mid 2010s significantly reduced effective taxes on equipment and software (from about 20% in the year 2000 to about 10%). The tax reform of 2017, which came into effect in 2018, further reduced effective taxes on equipment and software to about 5%. These reforms aimed to raise employment by stimulating investment and overall economic activity. In this subsection, we use our calibrated model to study the effectiveness of these reforms and their implications for automation.

Our main finding is that, although all of these reforms increased employment (because they reduced effective taxes), their effects were limited and did so at a large fiscal costs per

[32] As mentioned above, reducing capital taxes may be optimal because the use of capital in tasks in which it has a comparative advantage benefits labor due to complementarity between tasks. In practice, capital might also complement labor in labor-intensive tasks. To capture this possibility, the task production function could be changed to

$$y(x) = \psi^\ell(x) \cdot \ell(x)^\alpha \cdot \tilde{k}(x)^{1-\alpha} + \psi^k(x) \cdot k(x),$$

where $\tilde{k}(x)$ is the capital used to complement labor within tasks. Table A.3 in the Appendix shows that allowing for direct complementarities in this way (with $\alpha = 0.75$) does not change our main findings.

[33] Importantly, this can be implemented without raising *any* capital taxes. In particular, a tax on automation can also be implemented via a subsidy to labor of $\tau^A = 13.07\%$ combined with a tax of τ^A on the output of tasks above $\theta^c = 0.264$. This alternative implementation is discussed in Proposition A.1 in the Appendix.

job created, in large part because they encouraged additional automation. In contrast, we show that alternative reforms reducing labor taxes or combining lower capital taxes with an automation tax could have increased employment by more and at a much lower cost per job.

Column 1 of Table 2 reports the market equilibrium for the capital and labor taxes in 2000—$\tau^\ell = 25.5\%$ and $\tau^k = 20\%$. Column 2 then documents the impact of the tax cuts on capital enacted between 2000 and the mid 2010s, which reduced the effective tax on software and equipment to 10% and reduced government revenue by 10.49%. Our model implies that these tax cuts raised employment by a modest 1.01%, and did so at a substantial fiscal cost of $162,851 per job. As our theoretical analysis highlights, the lackluster employment response was in part because the lower taxes on capital encouraged greater automation, as shown by the increase in θ. Column 3 turns to the most recent (2017) tax cuts on capital. These are predicted to reduce government revenue by an additional 5.51% (or 16% relative to the revenue collected in 2000) and encourage further automation, with θ rising to 0.278. The resulting employment gain is again small, 1.47% relative to 2000 (or 0.46% relative to the mid-2010s), and comes at a fiscal cost per additional job of $169,857.

Columns 4-6 turn to alternative tax reforms that would have cost the same revenue as the capital tax cuts implemented between 2000 and the mid-2010s (10.49% of the year 2000 revenue). In column 4, we consider the implications of reducing labor taxes (for example, with a payroll tax cut) to $\tau^\ell = 21.09\%$ and keeping $\tau^k = 20\%$ as in 2000. This alternative reform would have increased employment by 3.56%, and would achieve this at 1/4th of the cost of one additional job in column 2. Part of the reason why reducing payroll taxes is much more effective in stimulating employment than cutting capital taxes is that lower payroll taxes reduce automation (θ falls to 0.269) whereas lower capital taxes further increase automation (θ increases from 0.272 to 0.276 between columns 1 and 2).

Column 5 considers another reform, this time combining lower capital taxes with an automation tax (again chosen to cost the same revenue as the tax cuts enacted between 2000 and the mid 2010s). This reform would have also raised employment by more than the reforms of the 2010s, increasing it by 2.43%, and would have cost $67,316 per job, which is less than half the cost per job in column 2. Notably, this policy combination involves an even larger tax cut for capital—from 20% to 8.58%. But crucially, the automation tax simultaneously rolls back any automation that the capital tax cut would have otherwise induced.[34]

[34] A policy of reducing taxes on capital and at the same time taxing automation is equivalent to lowering the tax on capital by 11.42%, but only at tasks below $\theta = 0.266$. This exceeds the 10% tax cut from 2000 to the 2010s. These targeted tax cuts for capital at tasks in which it has a strong comparative advantage allow policy makers to give even larger subsidies to capital accumulation without triggering excessive automation.

TABLE 2: Comparison of observed tax reforms and reforms costing the same revenue.

		Observed reforms		Alternative reforms		
	System in 2000 with $\tau^k = 20\%$	System in 2010s: reform to $\tau^k = 10\%$	System in 2018: reform to $\tau^k = 5\%$	Labor tax reform	Capital tax reform with automation taxation	Capital and labor tax reform
	(1)	(2)	(3)	(4)	(5)	(6)
Tax system:						
τ^k	20.00%	10.00%	5.00%	20.00%	8.58%	26.65%
τ^ℓ	25.50%	25.50%	25.50%	21.09%	25.50%	18.22%
θ	0.271	0.276	0.278	0.269	0.266	0.265
τ^A	0.00%	0.00%	0.00%	0.00%	11.42%	0.00%
Aggregates:						
Employment	.	+1.01%	+1.47%	+3.56%	+2.43%	+5.06%
Capital	.	+5.64%	+8.34%	+1.14%	+3.07%	−2.19%
Labor Share	56.30%	56.00%	55.86%	56.46%	58.15%	56.77%
Net Output	.	+2.38%	+3.49%	+2.83%	+2.55%	+2.83%
Cost/Revenue per job	.	$162,851	$169,857	$45,954	$67,316	$32,378
Revenue	.	−10.49%	−16.00%	−10.49%	−10.49%	−10.49%

Notes: This table shows the effective capital and labor taxes and the level of automation for different tax reforms. Column 1 presents the equilibrium under the tax system of the year 2000. Columns 2 and 3 present the resulting changes from the capital tax cuts enacted up to the mid 2010s and then the subsequent capital tax cuts enacted in 2017. Columns 4–6 then show the effects of three alternative reforms that would have cost the same as the capital tax cuts enacted between 2000 and the mid 2010s. Column 4 considers cutting the effective labor tax. Column 5 considers a combination of capital tax cuts and a tax to automation. Column 6 considers a combination of lower labor taxes and higher capital taxes.

Finally, column 6 considers a reform that changes both capital and labor taxes in a welfare maximizing way and costs the same revenue as the reform in column 2. By definition, this reform coincides with the Ramsey solution in column 2 of Table 1 and would have raised the effective capital tax rates to 26.65%, and reduce the labor tax to 18.22% (eliminating the payroll tax almost entirely). We include it in this table to show that, in addition to the 5.06% additional increase in employment, such a reform would have had a much smaller cost per job—only $32,378, or about 1/5th of the cost per job generated by the capital tax cuts since 2000.

Overall, this discussion shows that, because automation responds to the cost of capital, reducing capital taxes uniformly (via generous depreciation allowances or reductions in corporate taxes) is not an effective way of stimulating employment. Reforms over the last two decades that reduced capital taxes achieved only a modest increase in employment and instead encouraged further automation. Moving forward, reducing payroll taxes or accompanying tax cuts for capital with a tax on automation can more powerfully stimulate economic activity and achieve greater increases in employment at lower fiscal costs.

4.5 Capital Distortions

Our analysis so far incorporates labor market imperfections via the wedge ϱ, but ignores capital distortions. This is motivated by two considerations. First, our starting point is that, because of labor market imperfections such as bargaining, search or efficiency wages, even without any taxes, the level of employment would be too low; the wedge used in our model introduces this property in a simple way. Second, while earning losses from worker displacement provide a natural way of identifying the labor market wedge, there is no simple method for ascertaining whether there are capital wedges and how large they may be.[35] Nevertheless, we have carried out a number of exercises to verify that our conclusions are not unduly affected by this asymmetry in the treatment of capital and labor.

First, if equity finance is not subject to an additional distortion, then the deductions of interest rate payments from taxes in the case of debt finance more than undo any capital market distortions. This is because the interest rate on corporate loans is an upper bound on the capital wedge and is deducted from taxes. Therefore, we can conservatively use effective tax rates on equipment and software that would apply with full equity financing (without any of the reductions in effective capital taxes that come with debt finance). Table A.4 in the Appendix provides analogous results to Table 1 in this case. The effective capital taxes are now $\tau^k = 12\%$ but this has minimal effects on our results. Second, Table A.5 in the Appendix repeats our main exercise but now assuming a capital wedge of $\varrho^k = 0.15$—the same as the labor wedge. The employment and welfare gains from moving to optimal taxes are still non-trivial even if about half as large as our baseline estimates. We conclude that our results are not driven by the assumption that there are no capital wedges or the asymmetric treatment of capital and labor.

5 Extensions

In this section, we discuss two extensions that generalize our model and reinforce our main conclusion that the US tax code favors capital and promotes excessive automation.

[35] For example, large corporations that have significant cash at hand should not be using a different external rate of return than their internal rate of return, and their behavior should not be affected by a capital wedge, even if they use external funds. Smaller corporations may face a higher rate of return when borrowing funds, but if investment in these and larger corporations are highly substitutable, this may not correspond to an aggregate capital wedge.

5.1 Human Capital Investments

The asymmetric treatment of capital and labor may further distort investments in human capital, which may interact with automation decisions. To incorporate this possibility, suppose that the efficiency units of labor services provided by a worker is augmented by his or her human capital. Assume also that all workers have the same amount of human capital h, so that the efficiency units of labor are now $\ell_h = h \cdot \ell$.[36] The cost of investing in human capital h for ℓ workers is $\frac{\ell}{1+1/\varepsilon^h} \cdot h^{1+1/\varepsilon^h}$ in terms of the final good of the economy, and $\varepsilon^h > 0$. This parameter will be the elasticity of investment in human capital with respect to changes in wages. Likewise, we take the iso-elastic specification of $\nu(\ell)$ used in our quantitative section, so that ε^ℓ is the constant Hicksian elasticity of labor supply.

Incorporating human capital into the labor market-clearing condition, we obtain

$$f_{\ell_h} \cdot (1-\tau^\ell) \cdot (1-\varrho) = \ell_h{}^{1/(\varepsilon^\ell + \varepsilon^h + \varepsilon^\ell \cdot \varepsilon^h)}.$$

The relevant elasticity for the supply of efficiency units of labor has now been replaced by $\varepsilon^\ell + \varepsilon^h + \varepsilon^\ell \cdot \varepsilon^h$, which incorporates the elastic response of human capital and is thus always greater than ε^ℓ. Intuitively, efficiency units of labor can be increased not just by supplying labor, but by investing in human capital as well.

The next proposition characterizes optimal taxes in the presence of human capital and shows that labor taxes need to be adjusted to take into account the greater elasticity with which labor services respond to taxation. This pushes in the direction of (relatively) lower labor taxes, and conversely, higher capital taxes.

Proposition 5 ***(Optimal taxes with endogenous human capital)*** *The solution to the Ramsey problem in an environment with human capital satisfies* $\theta^r = \theta^m(k,\ell)$ *and*

$$\frac{\tau^{k,r}}{1-\tau^{k,r}} = \frac{\mu}{1+\mu}\frac{1}{\varepsilon^k(k)} \qquad \frac{\tau^{\ell,r}}{1-\tau^{\ell,r}} = \frac{\mu}{1+\mu}\frac{1}{\varepsilon^\ell + \varepsilon^h + \varepsilon^\ell \cdot \varepsilon^h} - \frac{\varrho}{1+\mu}.$$

Moreover, if an economy has too low a tax on capital and excessive automation without human capital (in the sense of Proposition 2), it will a fortiori have too low a tax on capital and excessive automation when there is an elastic response of human capital.

We next provide a back of the envelope quantification of the effect of human capital investments on optimal policy. To do this, we augment our analysis in the previous section

[36] This formulation ignores the fact that high-human capital workers may be employed in tasks that are not automated or are complementary to automation technologies. The impact of automation on the employment and wages of different types of workers is explored in Autor, Levy & Murnane (2003) and Acemoglu & Restrepo (2020b).

with an estimate for the elasticity of human capital, ε^h. We set the elasticity of human capital supply, ε^h, to 0.092. This value is in the mid-range of estimates from the literature on high-school completion (Jensen, 2010; Kuka et al., 2018) and college major choice (Wiswall & Zafar, 2015; Beffy et al., 2012).[37] This increases the supply elasticity of efficiency units of labor to 0.86, and as a result, the optimal labor tax is now lower, τ^ℓ = 16.90%, and the optimal capital tax is modestly higher, τ^k = 29.21% (see Table A.7 in the Appendix). Replacing the current system with optimal taxes leads to more pronounced changes: 1.06 percentage point higher labor share, 5.73% increase in employment, and 0.59% increase in welfare in consumption-equivalent terms.

5.2 Endogenous Technology

In our baseline model, increases in θ represent both the development and the adoption of automation technologies. In principle, these two decisions are distinct, even if related. Unless automation technologies are developed, they cannot be adopted. If they are expected to be adopted, then there are greater incentives to develop them. Moreover, as emphasized in Acemoglu & Restrepo (2018), new automation technologies may come at the expense of other technological changes with different implications for capital and labor. For instance, more resources devoted to automation typically imply less effort towards the introduction of new tasks that tend to increase the labor share and demand for labor. If so, a tax structure that favors capital may distort the direction of technological change in a way that disadvantages labor. In this subsection, we provide a simple model to highlight these ideas and show that, with endogenous technology, optimal policy may also need to redirect the direction of technological change, and this is the case even when capital and labor taxes are set optimally.

For brevity, we borrow from the formulation of endogenous technology in Acemoglu (2007, 2010), whereby a (competitive) production sector decides how much capital and labor to use and which technology, from a menu of available technologies, to utilize, while a monopolistically competitive (or simply monopolistic) technology sector decides which menu of technologies to develop and offer to firms.

We consider a menu of technologies consisting of both automation techniques and technologies that increase the productivity and the set of tasks performed by labor (such as

[37] Jensen's (2010) experimental results imply a 0.097 high-school completion elasticity in response to perceived returns. Kuka et al. (2018) estimate a high-school completion elasticity of 0.019-0.086 in response to actual returns, and 0.014-0.17 in response to perceived returns. Wiswall & Zafar (2015) estimate elasticities in the range of 0.036-0.062 from the response of college major choice to changes in relative wage premium. Previous estimates in Beffy et al. (2012) put the same elasticity in the range 0.09–0.12. Taken together, these studies imply values for ε^h in the range 0.014–0.17.

the introduction of new tasks considered in Acemoglu & Restrepo, 2018). We summarize this menu by Θ with the convention that a higher Θ means a menu that is more biased towards automation technologies. Given menu Θ, firms choose their level of automation θ and their utilization of other technologies ω subject to the feasibility constraint $G(\theta, \omega; \Theta) \leq 0$. Therefore, the index of technologies Θ determines what combinations of automation and other technologies are feasible for final good producers. We denote the production function given θ and ω by $f(k, \ell; \theta, \omega)$, and assume that f_k/f_ℓ is increasing in θ as in our baseline model and decreasing in ω. We further assume that when Θ increases, the set $G(\theta, \omega; \Theta) \leq 0$ includes higher values of θ and lower values of ω, so that a higher Θ enables more adoption of automation technologies and less adoption of other technologies.

The profit-maximizing adoption decision solves the following problem:

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

The assumptions on $G(\theta, \omega; \Theta)$ imply that $\omega^m(k, \ell; \Theta)$ is decreasing in Θ and $\theta^m(k, \ell; \Theta)$ is increasing in Θ—so that a higher Θ means a menu of technologies that is more biased towards automation. Hence, as the menu of available technologies becomes more biased towards automation, it crowds out the adoption of non-automation technologies (such as new tasks or others that increase human productivity).

Finally, we assume that the technology sector charges markups for the use of technologies by final good producers and via this, captures a constant fraction $\kappa \in (0, 1)$ of the output of these producers (this could be micro-founded as in Acemoglu (2007, 2010), by assuming that the technology sector sells machines embedding the new technology with a constant markup). We denote the cost of choosing a menu of technologies Θ by $\Gamma(\Theta)$. Thus, the maximization problem of the technology sector that determines the equilibrium bias of technology is

$$\max_{\Theta} \kappa \cdot f(k, \ell; \theta^m(k, \ell; \Theta), \omega^m(k, \ell; \Theta)) - \Gamma(\Theta). \tag{20}$$

We make the following assumptions on $\Gamma(\Theta)$:

- $\Gamma(\Theta)$ has a minimum at $\bar{\Theta} \in (0, 1)$. This assumption means that there exists a baseline bias of technology $\bar{\Theta}$, such that deviations from this baseline involve increasing costs. More specifically, deviations from $\bar{\Theta}$ can come in the direction of further automation or further effort devoted to creating new tasks (and thus less automation). Both of these will be more costly than continuing with $\bar{\Theta}$. In the dynamic framework of Acemoglu & Restrepo (2018), $\bar{\Theta}$ corresponds to the state of technology inherited from the past.

- $\Gamma(\Theta)$ is convex, which captures diminishing returns in research directed at changing the bias of technology away from the baseline level $\bar{\Theta}$.

In addition to capital and labor taxes, we allow for subsidies to the use of automation and other technologies in the final good sector to undo the effects of the markup κ and for taxes on the profits of the technology monopolist. Our results do not depend on whether such additional taxes and subsidies exist, but their presence simplifies the expressions and makes them much more closely connected to those in our baseline model in Section 2.

A market equilibrium satisfies the same market-clearing conditions as in our benchmark economy, but is augmented to include the fact that technology adoption decisions of final good producers are given by $\theta^m(k, \ell, \Theta)$ and $\omega^m(k, \ell; \Theta)$, and the equilibrium bias of technology Θ maximizes (20). We assume that a market equilibrium exists and is unique, and that the solution to (20) always involves some interior $\Theta \in (0, 1)$.

We next characterize the solution to the Ramsey problem as in Proposition 1. As in our baseline model, we assume that the planner directly controls the development and adoption of technologies (these choices can be implemented with additional taxes as in Section 2.5).

PROPOSITION 6 ***(Optimal taxes and automation with endogenous technology)*** *The solution to the Ramsey problem with endogenous technology involves capital and labor taxes given as in (8) and undistorted adoption decisions (conditional on Θ) given by $\theta^r(k, \ell; \Theta) = \theta^m(k, \ell; \Theta)$ and $\omega^r(k, \ell; \Theta) = \omega^m(k, \ell; \Theta)$. However, if $\Theta^r \lesseqgtr \bar{\Theta}$, the optimal bias of technology satisfies $\Theta^r \lesseqgtr \Theta^m$ (i.e., the optimal and market bias of technology are the same if and only if $\Theta^r = \bar{\Theta}$).*

The most important implication of this proposition is that, even with optimal taxes on capital and labor, the planner might wish to discourage the development of automation technology. This will be the case when the baseline level of technology is more geared towards automation than what the planner would like to achieve. Put differently, if the economy in question has already gone in the direction of excessively developing automation technologies (which may be a consequence of past distortions or other factors influencing the direction of past technological change), then the planner should intervene by distorting the direction of innovation. The reason for this is that the technology sector does not fully internalize the social surplus created by its technology choices (because of the presence of the term $\kappa < 1$ in (20)), and thus will not develop the right type of technologies. This result has a close connection to one of the key insights in Acemoglu et al. (2012), which established, in the context of optimal climate change policy, that if the economy starts with relatively advanced carbon-emitting, dirty technologies and relatively backward low-carbon, clean technologies, then it is not sufficient to impose Pigouvian taxes; rather, optimal policy additionally calls for direct subsidies to the development of clean technologies.[38]

This result is important in our context because, if as our results in Section 4 suggest, past US tax policy has favored capital and automation, then it is not sufficient to redress the distortions in the current tax system. Because these policies have likely led to excessive development of automation technologies, optimal policy may need to intervene to redirect technological change by subsidizing the creation of new tasks and temporarily discourage further effort towards automation innovations at the margin. We leave a quantitative exploration of the implications of endogenous technology development to future work.

6 Concluding Remarks

Automation is transforming labor markets and the structure of work in many economies around the world, not least in the United States. The number of robots in industrial applications, the use of specialized software, artificial intelligence and several other automation technologies have increased rapidly in the US economy over the last few decades. There has been a concomitant decline in the labor share of national income, wages have stagnated and low-skill workers have seen their real wages decline. Many experts believe that these labor market trends are, at least in part, related to automation.

The general intuition among economists (and many policy-makers) is that even if automation may have some adverse distributional and employment consequences, policy should not slow down (and certainly not prevent) the adoption of automation technologies, because these technologies are contributing to productivity. According to this perspective, policy should instead focus on fiscal redistribution, education and training to ensure more equally-distributed gains and more opportunities for social mobility. But what if automation is excessive from a social point of view?

This paper has argued that the US tax system is likely to be encouraging excessive automation and if so, reducing the extent of automation (or more plausibly, slowing down the development and adoption of new automation technologies) may be welfare-improving. We have developed this argument in three steps.

First, we revisited the theory of optimal capital and labor taxation in a task-based framework where there is an explicit decision of firms to automate tasks. We also introduced, albeit in a reduced-form manner, labor market imperfections. Consistent with the classical theory of public finance, if capital and labor taxes are set optimally, automation decisions are opti-

[38]Note in addition that, once the planner can influence the direction of automation technology and set optimal taxes on capital and labor, there is no need to distort the adoption of automation technologies.

mal in equilibrium. However, away from optimal capital and labor taxes or in the presence of additional constraints on tax decisions, this is no longer the case. Exploiting the structure of our task-based framework, we establish that when the tax system is already biased against labor, it is generally optimal to distort equilibrium automation. The economics of this result is simple but informative: marginal tasks that are automated bring little productivity gains (or in the terminology of Acemoglu & Restrepo, 2019a,b, they are "so-so automation technologies"), and as a result, the cost of reducing automation at the margin is second-order. When the tax system is biased against labor, the gain from reducing automation and preventing the displacement of labor is first-order because it increases employment. In fact, it may even be optimal to reduce automation while at the same time cutting capital taxes (even though the tax system is biased against labor and in favor of capital), because, in contrast to automation, the use of capital in tasks in which capital has a strong comparative advantage is complementary to workers employed in labor-intensive tasks.

Second, we delved into a detailed evaluation of the US tax system in order to map the complex tax code into effective capital and labor taxes. Our numbers suggest that the US system taxes labor heavily and favors capital significantly. While labor is taxed at an effective rate between 25.5% and 33.5%, capital faces an effective tax rate of about 5% (down from 10% in the 2010s and 20% in the 1990s and early 2000s).

Third, we compared the US tax system to the optimal taxes implied by our theoretical analysis. This exercise confirmed that the US tax system is biased against labor and in favor of capital. As a result, we found that moving from the current US tax system and level of automation to optimal taxation of factors and the optimal level of automation would raise employment by 4.02%, the labor share by 0.78 percentage point and overall welfare by 0.38% in consumption-equivalent terms. If optimal taxes can be implemented, there is no need for distorting or taxing automation. If, on the other hand, optimal taxes are infeasible, more modest reforms involving a tax on automation can undo some of the inefficiencies in the current system and increase employment by 1.14–1.96% and the labor share by 1.93–2.54 percentage points. In this case, the constrained optimal policy always involves an automation tax in order to discourage the automation of marginal tasks which bring little productivity benefits but significant displacement of labor.

We also showed that a range of realistic generalizations (absent from our baseline framework) reinforce our conclusions and call for even more extensive changes in automation and capital taxation, and under some conditions, it may be optimal to redirect new innovations away from automation.

To simplify the analysis, we focused on an economy with a single type of labor. As noted in the Introduction, automation is also associated with increases in inequality (Autor, Levy & Murnane, 2003; Acemoglu & Autor, 2011; Acemoglu & Restrepo, 2020a,b). Consequently, slowing down automation may generate additional distributional benefits. These issues are discussed in Guerreiro, Rebelo & Teles (2017), Thuemmel (2018) and Costinot & Werning (2018). A natural next step is to augment these analyses with the possibility that certain aspects of the tax system may be encouraging excessive automation.

In practice, there are many potential sources of excessive automation. Our objective in this paper has been narrow: to focus on tax reasons for excessive automation. Our companion paper Acemoglu, Manera & Restrepo (2020) shows that, even absent tax-related distortions, the market economy tends to generate excessive automation because bargaining power and efficiency wage considerations vary across tasks and this tends to create incentives for firms to automate beyond what is socially beneficial in order to improve their share of rents. Furthermore, as we have already noted, automation-driven job loss may generate negative spillovers on communities and political and social behavior. There may additionally be social factors and corporate strategies concerning the direction of innovation and research (the best minds in many fields being attracted to automation technologies and the most influential companies favoring automation) that further contribute to excessive automation. Quantifying the extent of these other factors is an important area for future research, especially because they have major implications for policy.

Finally, we should briefly comment on how our results relate to two popular policy proposals: wealth taxes and "robot taxes". Although our framework suggests that it may be beneficial to increase taxes on capital, wealth taxes on high wealth individuals may not be the most direct way of achieving this, because they would not necessarily increase the effective tax on the use of capital. Increasing corporate income taxes and eliminating or lowering depreciation allowances may be more straightforward ways of implementing higher effective taxes on capital (provided that there are no other distributional or political benefits from wealth taxes). Moreover, our framework highlights that it is not always beneficial to increase taxes on capital: when it is not feasible to implement optimal taxes, reducing automation becomes a central objective (and may even need to be combined with lower taxes on capital). Our automation tax is also different from taxes on robots for the same reasons: it is not a uniform tax on all automation technologies; rather, it is applied to technologies automating tasks above a certain threshold (which are tasks in which humans still have a significant comparative advantage). In fact, our results clarify that, instead of taxing all automation technologies, optimal policy often involves subsidizing capital in tasks in which machines have a strong comparative advantage. Lastly, our analysis also clarifies that if the tax system is reformed so that it is no longer biased against labor and in favor of capital, then employment and welfare can be increased without an automation tax.

References

Acemoglu, Daron. 2007. "Equilibrium Bias of Technology." *Econometrica* 75 (5): 1371–1409.

———. 2010. "When Does Labor Scarcity Encourage Innovation?" *Journal of Political Economy* 118 (6): 1037–78.

Acemoglu, Daron, Philippe Aghion, Leonardo Bursztyn, and David Hemous. 2012. "The Environment and Directed Technical Change." *American Economic Review* 102 (1): 131–66.

Acemoglu, Daron & David Autor. 2011. "Skills, Tasks and Technologies: Implications for Employment and Earnings." In *Handbook of Labor Economics*, 4:1043–1171. Elsevier.

Acemoglu, Daron & Pascual Restrepo. 2018. "The Race between Man and Machine: Implications of Technology for Growth, Factor Shares, and Employment." *American Economic Review* 108 (6): 1488–1542.

———. 2019a. "Artificial Intelligence, Automation, and Work." Chapter in *The Economics of Artificial Intelligence: An Agenda*, edited by Agrawal, Gans and Goldfarb.

———. 2019b. "Automation and New Tasks: How Technology Displaces and Reinstates Labor." *Journal of Economic Perspectives* 33 (2): 3–30.

———. 2020a. "Robots and Jobs: Evidence from US Labor Markets." *Journal of Political Economy*

———. 2020b. "The Labor Share, Displacement and Wage Inequality." Mimeo. Massachusetts Institute of Technology.

Acemoglu, Daron, David Autor, Joe Hazell & Pascual Restrepo 2020. "AI and Jobs: Evidence from Online Vacancies." Mimeo. Massachusetts Institute of Technology.

Acemoglu, Daron, Andrea Manera & Pascual Restrepo. 2020. "Automation in Imperfect Labor Markets." Mimeo. Massachusetts Institute of Technology.

Atkinson, A. B. & J. E. Stiglitz. 1972. "The Structure of Indirect Taxation and Economic Efficiency." *Journal of Public Economics* 1 (1): 97–119.

Austin, Benjamin, Edward Glaeser & Lawrence Summers. 2018. "Jobs for the Heartland: Place-Based Policies in 21st-Century America." *Brookings Papers on Economic Activity*, 151–232.

Autor, David H., David Dorn & Gordon H. Hanson. 2013. "The China Syndrome: Local Labor Market Effects of Import Competition in the United States." *American Economic*

Review 103 (6): 2121–68.

Autor, David H. & Mark G. Duggan. 2003. "The Rise in the Disability Rolls and the Decline in Unemployment." *The Quarterly Journal of Economics* 118 (1): 157–206.

Autor, David H., Frank Levy & Richard J. Murnane. 2003. "The Skill Content of Recent Technological Change: An Empirical Exploration." *The Quarterly Journal of Economics* 118 (4): 1279–1333.

Becker, Robert A. and John H. Boyd III. 1993. "Recursive Utility: Discrete Time Theory." *Hitotsubashi Journal of Economics* 34 (Special Issue): 49–98 .

Beffy, Magali, Denis Fougère & Arnaud Maurel. 2012. "Choosing the Field of Study in Postsecondary Education Do Expected Earnings Matter?" *Review of Economics & Statistics* 94 (1): 334–47.

Brülhart, Marius, Jonathan Gruber, Matthias Krapf, and Kurt Schmidheiny. 2016. "Taxing Wealth: Evidence from Switzerland." w22376. Cambridge, MA: National Bureau of Economic Research.

Chamley, Christophe. 1986. "Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives." *Econometrica* 54 (3): 607–22.

Chetty, Raj, Adam Guren, Day Manoli & Andrea Weber. 2011. "Are Micro and Macro Labor Supply Elasticities Consistent? A Review of Evidence on the Intensive and Extensive Margins." *American Economic Review* 101 (3): 471–75.

Congressional Budget Office. 2006. "Computing Effective Tax Rates on Capital Income." Publication 2624.

———. 2014. "Taxing Capital Income: Effective Marginal Tax Rates Under 2014 Law and Selected Policy Options." Publication 49817.

———. 2019. "Marginal Federal Tax Rates on Labor Income: 1962 to 2028." Publication 54911.

Cooper, Michael, John McClelland, James Pearce, Richard Prisinzano, Joseph Sullivan, Danny Yagan, Owen Zidar, and Eric Zwick. 2016. "Chapter 3: Business in the United States: Who Owns It, and How Much Tax Do They Pay?" In *NBER/Tax Policy & the Economy (University of Chicago Press)*, 30:91–128. National Bureau of Economic Research (NBER).

Costinot, Arnaud & Iván Werning. 2018. "Robots, Trade, and Luddism: A Sufficient Statistic Approach to Optimal Technology Regulation." Working Paper 25103. National Bureau of Economic Research.

Couch, Kenneth A & Dana W Placzek. 2010. "Earnings Losses of Displaced Workers Revisited." *American Economic Review* 100 (1): 572–89.

Davis, Steven J., Till Von Wachter, Robert E. Hall & Richard Rogerson. 2011. "Recessions and the Costs of Job Loss." *Brookings Papers on Economic Activity*, 1–72.

Diamond, Peter A. and James E. Mirrlees. 1971. "Optimal Taxation and Public Production I: Production Efficiency" *American Economic Review* 61 (1): 8–27.

Durán-Cabré, José M., Alejandro Esteller-Moré and Mariona Mas-Montserrat. 2019. "Behavioural Responses to the (Re)Introduction of Wealth Taxes. Evidence From Spain". IEB Working Paper 2019/4. Institut d'Economia de Barcelona.

Edgerton, Jesse. 2010. "Investment Incentives and Corporate Tax Asymmetries" *Journal of Public Economics* 94 (11–12): 936–952.

Elsby, Michael W. L., Bart Hobijn & Aysegül Şahin. 2013. "The Decline of the U.S.Labor Share." *Brookings Papers on Economic Activity* 2013 (2): 1–63.

Garrett, Daniel G., Eric Ohrn & Juan Carlos Suarez Serrato. Forthcoming. "Tax Policy and Local Labor Market Behavior." *American Economic Review: Insights*.

Goldman, Dana, Neeraj Sood & Arleen Leibowitz. 2005. "Wage and Benefit Changes in Response to Rising Health Insurance Costs." Working Paper 11063. National Bureau of Economic Research.

Goolsbee, Austan. 1998. "Investment Tax Incentives, Prices, and the Supply of Capital Goods." *The Quarterly Journal of Economics* 113 (1): 121–48.

Gruber, Jonathan & Alan B. Krueger. 1991. "The Incidence of Mandated Employer-Provided Insurance: Lessons from Workers' Compensation Insurance." In *Tax Policy and the Economy*, edited by David Bradford, 5:111–44. The MIT Press.

Guerreiro, Joao, Sergio Rebelo & Pedro Teles. 2017. "Should Robots Be Taxed?" Working Paper 23806. National Bureau of Economic Research.

Hall, Robert E., and Dale W. Jorgenson. 1967. "Tax Policy and Investment Behavior." *The American Economic Review* 57 (3): 391–414.

Hassett, Kevin A. & R. Glenn Hubbard. 2002. "Chapter 20 - Tax Policy and Business Investment." In *Handbook of Public Economics*, edited by Alan J. Auerbach and Martin Feldstein, 3:1293–1343. Elsevier.

House, Christopher L & Matthew D Shapiro. 2008. "Temporary Investment Tax Incentives: Theory with Evidence from Bonus Depreciation." *American Economic Review* 98 (3): 737–68.

Jacobson, Louis S., Robert J. LaLonde & Daniel G. Sullivan. 1993. "Earnings Losses of Displaced Workers." *The American Economic Review* 83 (4): 685–709.

Jakobsen, Katrine, Kristian Jakobsen, Henrik Kleven, and Gabriel Zucman. 2020. "Wealth Taxation and Wealth Accumulation: Theory and Evidence From Denmark." *The Quar-*

terly Journal of Economics 135 (1): 329–88.

Jensen, Robert. 2010. "The (Perceived) Returns to Education and the Demand for Schooling." *The Quarterly Journal of Economics* 125 (2): 515–48.

Joulfaian, David. 2006. "The Behavioral Response of Wealth Accumulation to Estate Taxation: Time Series Evidence." *National Tax Journal* 59 (2): 253–68.

Judd, Kenneth L. 1985. "Redistributive Taxation in a Simple Perfect Foresight Model." *Journal of Public Economics* 28 (1): 59–83.

———. 1999. "Optimal Taxation and Spending in General Competitive Growth Models." *Journal of Public Economics* 71 (1): 1–26.

Karabarbounis, Loukas & Brent Neiman. 2014. "The Global Decline of the Labor Share." *The Quarterly Journal of Economics* 129 (1): 61–103.

Kleven, Henrik Jacobsen & Esben Anton Schultz. 2014. "Estimating Taxable Income Responses Using Danish Tax Reforms." *American Economic Journal: Economic Policy* 6 (4): 271–301.

Kopczuk, Wojciech & Joel Slemrod. 2000. "The Impact of the Estate Tax on the Wealth Accumulation and Avoidance Behavior of Donors." Working Paper 7960. National Bureau of Economic Research.

Kuka, Elira, Na'ama Shenhav & Kevin Shih. 2018. "Do Human Capital Decisions Respond to the Returns to Education? Evidence from DACA." Working Paper 24315. Cambridge, MA: National Bureau of Economic Research.

Lennon, Conor. 2019. "Are the Costs of Employer-Sponsored Health Insurance Passed on to Workers at the Individual Level?" Working Paper.

Mui, Preston and Benjamin Schoefer. 2019. "The Labor Supply Curve at the Extensive Margin: A Reservation Wedge Approach" Working Paper.

Oberfield, Ezra & Devesh Raval. 2014. "Micro Data and Macro Technology." Working Paper 20452. National Bureau of Economic Research.

Smith, Matthew, Danny Yagan, Owen Zidar & Eric Zwick. 2019. "Capitalists in the Twenty-First Century." *The Quarterly Journal of Economics* 134 (4): 1675–1745

Straub, Ludwig & Iván Werning. 2020. "Positive Long-Run Capital Taxation: Chamley-Judd Revisited." *American Economic Review* 110 (1): 86–119.

Suárez Serrato, Juan Carlos & Owen Zidar. 2016. "Who Benefits from State Corporate Tax Cuts? A Local Labor Markets Approach with Heterogeneous Firms." *American Economic Review* 106 (9): 2582–2624.

The U. S. Bureau of Economic. 2018. *NIPA Handbook: Concepts and Methods of the U.S. National Income and Product Accounts.*

Thuemmel, Uwe. 2018. "Optimal Taxation of Robots." 7317. CESifo Working Paper Series. CESifo Group Munich.

Wiswall, M. & B. Zafar. 2015. "Determinants of College Major Choice: Identification Using an Information Experiment." *The Review of Economic Studies* 82 (2): 791–824.

Yagan, Danny. 2015. "Capital Tax Reform and the Real Economy: The Effects of the 2003 Dividend Tax Cut." *American Economic Review* 105 (12): 3531–63.

Zeira, Joseph. 1998. "Workers, Machines, and Economic Growth." *The Quarterly Journal of Economics* 113 (4): 1091–1117.

Zoutman, Floris T. 2018. "The Elasticity of Taxable Wealth: Evidence from the Netherlands." Working Paper.

Zwick, Eric & James Mahon. 2017. "Tax Policy and Heterogeneous Investment Behavior." *American Economic Review* 107 (1): 217–48.

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.

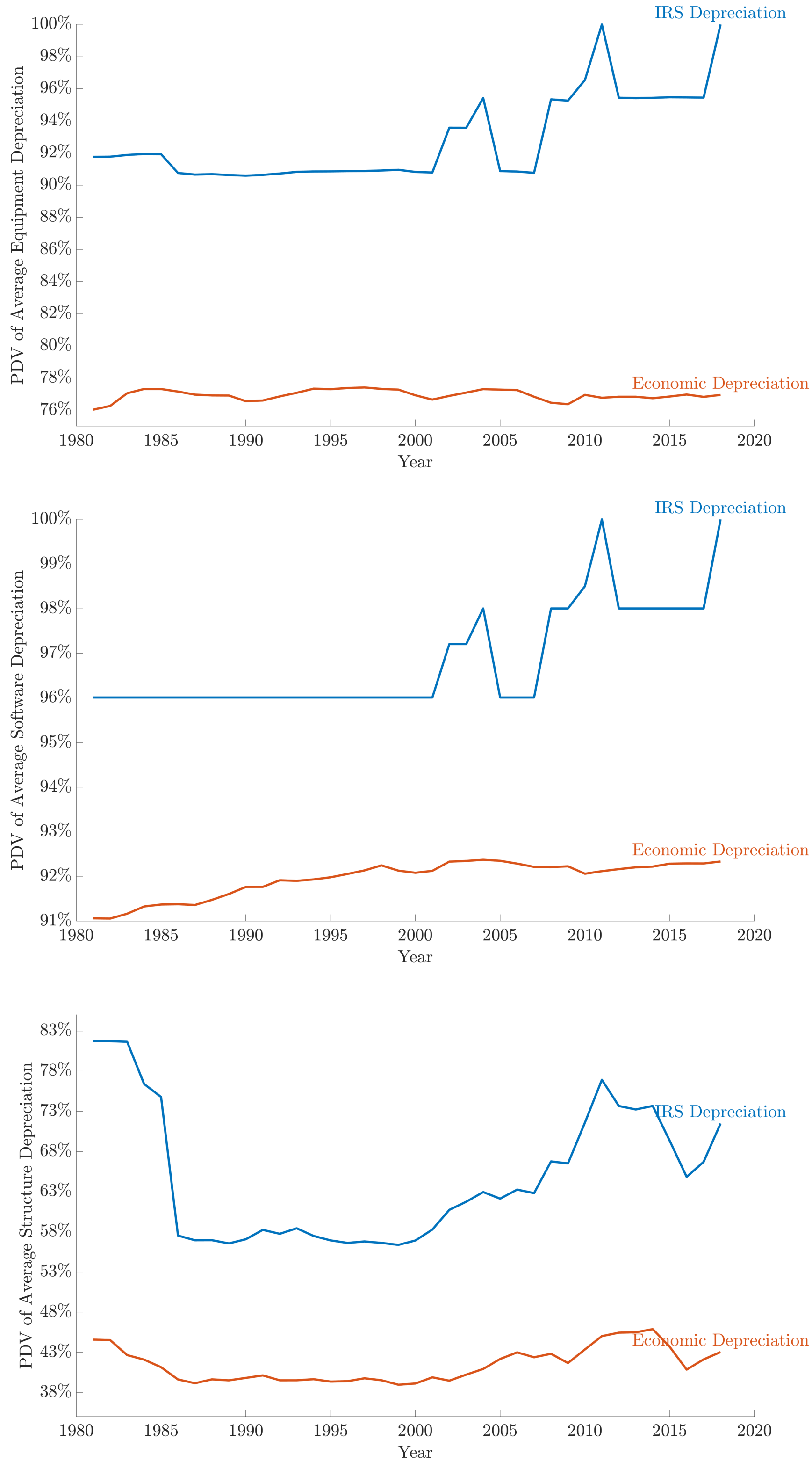


FIGURE A.1: ESTIMATED DEPRECIATION ALLOWANCES OVER TIME FOR EQUIPMENT, SOFTWARE AND NON-RESIDENTIAL STRUCTURES.

Notes: See the text for definitions.

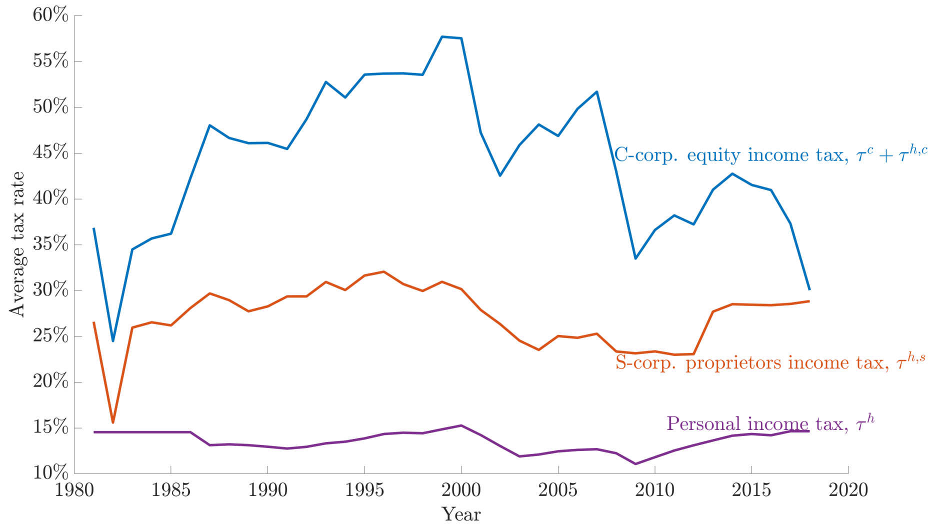


Figure A.2: Average tax rates on capital income, corporate income and personal income, 1981-2018.

Notes: See the text for the definitions and sources.

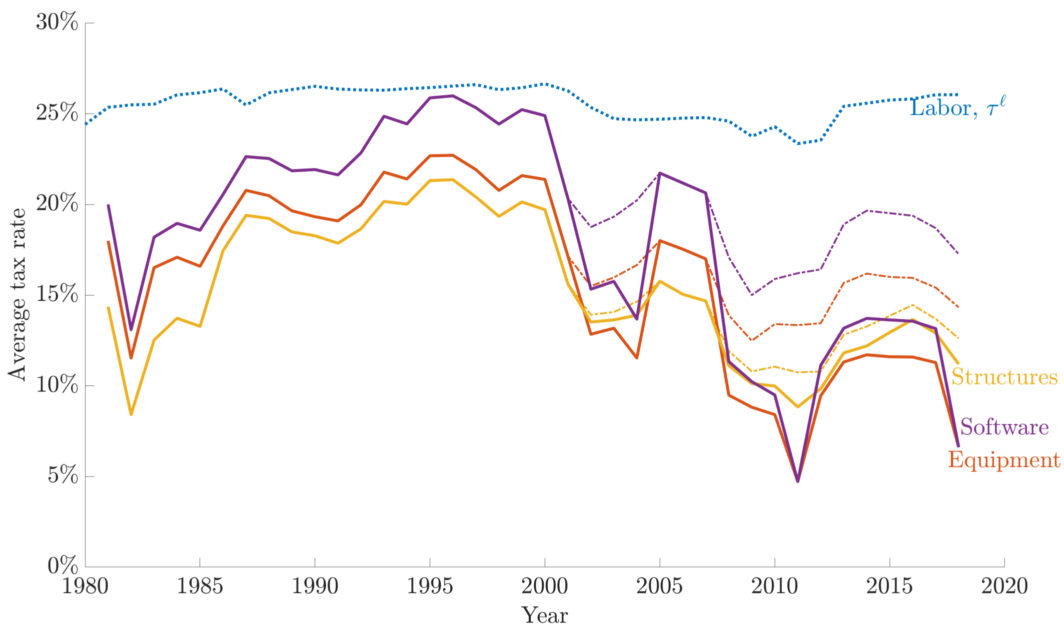


Figure A.3: Effective tax rates on labor, software capital, equipment, and non-residential structures with equity financing.

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 (1)	Ramsey Solution (2)	Distorting θ (3)	Distorting θ and changing τ^k (4)	Distorting θ and changing τ^ℓ (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 (1)	Ramsey Solution (2)	Distorting θ (3)	Distorting θ and changing τ^k (4)	Distorting θ and changing τ^ℓ (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 $x \le \theta$ 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.

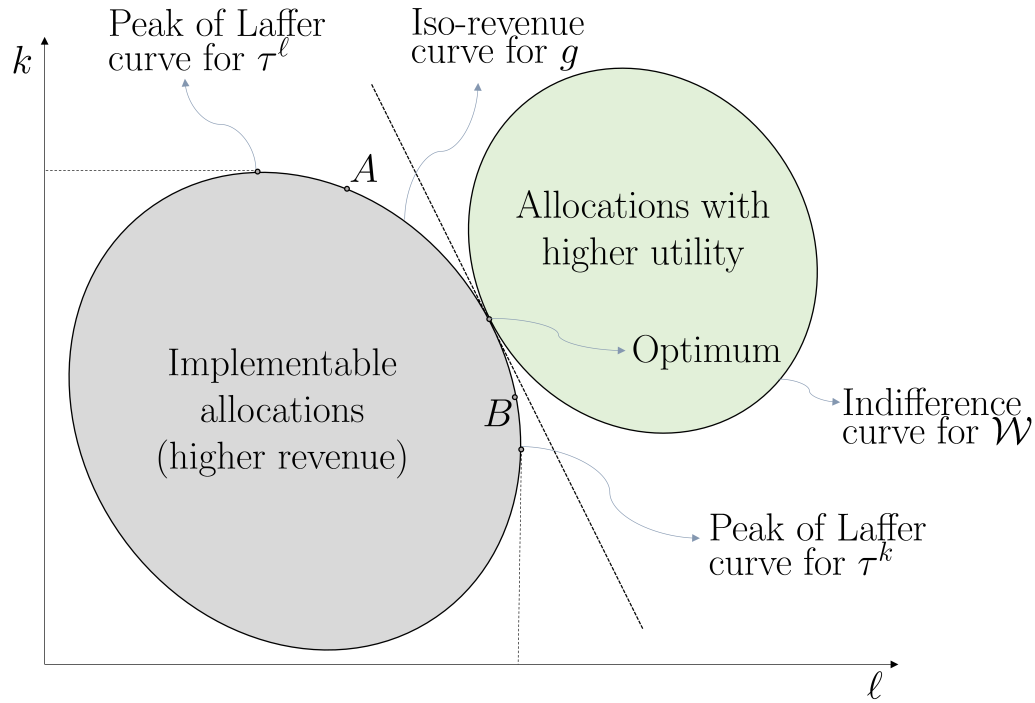


FIGURE A.4: Illustration of optimal policy problem.

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

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

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

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

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

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

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

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

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

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

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

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 $x \le \theta_p$ we have

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

which implies

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

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

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

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

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

Let $\mu > 0$ and $\gamma^k \geq 0$ 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 $\theta \geq \theta^m(k,\ell)$, so that $\theta^c < \theta^m(k,\ell)$. (Note that the inequality $1+\mu-\gamma^k\cdot(1-\bar{\tau}^k) \leq 0$ 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

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

- 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

$$\begin{aligned} &\max_{k, \ell, \theta, \omega, \Theta} \; f(k, \ell; \theta, \omega) + (1 - \delta) \cdot k + u(\bar{y} - k) - \nu(\ell) - \Gamma(\Theta) \\ &\text{subject to: } g \leq f(k, \ell; \theta, \omega) + (1 - \delta) \cdot k - u'(\bar{y} - k) \cdot k - \frac{\nu'(\ell) \cdot \ell}{1 - \varrho} - \Gamma(\Theta) \\ &G(\theta, \omega; \Theta) \leq 0. \end{aligned}$$

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 $\{(c_0, \ell_0), (c_1, \ell_1), \ldots\}$ 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

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

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

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

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

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:

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

- the labor market clears:

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

- the resource constraint holds:

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

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

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

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

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

which is equivalent to having zero taxes on capital. Likewise, equation (A.16) yields $f_\ell^{st} \cdot u_c^{st} + u_\ell^{st} = 0$. 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 $\mu \cdot A^{st} \neq 0$). In this case, (A.16) requires that $f_\ell^{st} \cdot u_c^{st} + u_\ell^{st}$ 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

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

Because ϑ_t is an arithmetic series, the right-hand side in this equation converges to $\bar{M}(V^{st})$, 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]

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

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

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

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 $\mathcal{W}_{uu} \leq 0$. Thus, all aggregators with $\mathcal{W}_{uV} \geq 0$ (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:

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

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 $\theta_t^r = \theta^m(k_t,\ell_t)$—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 $q^{\mathrm{st}} \geq 1/\bar{M}^{\mathrm{st}}$ 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 $\tilde{\mu}^{\mathrm{st}} > 0$. Suppose to obtain a contradiction that $\tilde{\mu}^{\mathrm{st}} = 0$. The first equation of the block implies that $q^{\mathrm{st}} = 1$ and the second equation implies that $\bar{M}^{\mathrm{st}} = f_k^{\mathrm{st}} + 1 - \delta$. The third equation of the block implies $\nu'(\ell^{\mathrm{st}}) = f_\ell^{\mathrm{st}}$. Thus, if $\tilde{\mu}^{\mathrm{st}} = 0$, 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 $\bar{M}^{\mathrm{st}} = f_k^{\mathrm{st}} + 1 - \delta$ by k^{st} and $\nu'(\ell^{\mathrm{st}}) = f_\ell^{\mathrm{st}}$ by ℓ^{st} and add these two equations to obtain

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

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

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

which contradicts the IC constraint.

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

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

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

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

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

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

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

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

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

Suppose next that $\bar{M}'(V^{st}) = 0$, then $q^{st} = 1$ gives the unique solution with $q^{st} > 1/\bar{M}^{st}$ 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

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

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

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

and

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

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

[43] In our system, $a = M_1^{st}/\bar{M}^{st}$ and $b = M_3^{st} \cdot \bar{M}^{st}$. The claim that $a, d \le 0$ is thus equivalent to $M_1^{st}, M_3^{st} \le 0$ 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 $M_{1t} \le 0$. Because $M_3^{st} \cdot \bar{M}^{st} = M_1^{st}$, we have $M_3^{st}, M_1^{st} \le 0$.

Taylor expansion of the first equation of the block gives

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

This equation implies

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

for some constant D, and we can therefore write

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

Substituting this expression into equation (A.23) yields

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

The formula in the proposition follows from the fact that $\mathcal{O}((q^{\mathrm{st}} - 1)^2) = \mathcal{O}({\tau^{k,r}}^2)$, 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

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

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*

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

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

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 $\theta^r_t = \theta^m(k_t, \ell_t)$, and for $t \geq 1$, capital and labor taxes are given by

$$\frac{\tau^{k,r}_t}{1 - \tau^{k,r}_t} = \frac{\tilde{\mu}_t}{1 + \tilde{\mu}_t} \frac{1}{\varepsilon^k_t} \qquad\qquad \frac{\tau^{\ell,r}_t}{1 - \tau^{\ell,r}_t} = \frac{\tilde{\mu}_t}{1 + \tilde{\mu}_t} \frac{1}{\varepsilon^\ell_t} - \frac{1}{1 + \tilde{\mu}_t} \varrho.$$

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 $a^g \leq 0$, 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 $a^g \leq 0$ 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

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

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

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

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, $\tau^{c,\text{SL}}$ is given by the net operating surplus of corporation that can be attributed to C-corporations, $\text{NOSCORP}_t^{\text{C,IRS}}$. 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, $\text{NOSCORP}_t^{\text{C,IRS}}$. 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), CT^{Fed}. Accordingly, the federal tax rate on capital income from C-corporations can be estimated as

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

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

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

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

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

[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^o_t 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^o_t.

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, $\tau^{c,\mathrm{SL}}$, 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 $\left\{d_t^j\right\}$, 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