Optimal Taxation with Automation: Navigating Capital and Labor's Complicated Relationship*

Eric Donald†
University of Pittsburgh

December 17, 2025

Abstract

The advent of new AI tools has caused many to worry about the displacement effect of automation. The common reaction to this problem follows a Pigouvian intuition: automating capital harms workers, so it should be taxed. This paper argues that this Pigouvian intuition is misguided. Capital harms workers at the extensive margin of automation, but at the intensive margin, more capital in a task that has already been automated raises overall output via capital deepening. I show that if the Planner can stipulate how much more expensive labor must be than capital before automation can occur, then capital taxes should be zero, recovering the celebrated result of Atkinson and Stiglitz (1976) in spite of the dependence of relative wages on equilibrium automation. Calibrating occupation-level exposure to automation to match aggregate substitution between capital and labor, I find that the optimal response to the US tax system is to require labor to be 31.8% more expensive than capital before automation can occur, but this reduces to 10.8% with an optimized non-linear income tax. Simulating the introduction of AI, I find that the optimal threshold rule is sharply increasing in the expected productivity growth of capital.

*I thank Doug Hanley, Bob King, David Lagakos, Ben Lockwood, Theodore Papageorgiou, Pascual Restrepo, and Stephen Terry as well as seminar participants at Boston University, Green Line Macro Meeting, SEA Annual Meeting, and TCU for their helpful comments.

†Email: eric.donald@pitt.edu

1 Introduction

Automation is one of the leading theories for why the wages of some American workers have stagnated or even declined (Acemoglu and Restrepo, 2022b), and recent advances in artificial intelligence have led to widespread concern about further displacement and deep inequality (Eloundou et al., 2024). The reaction to this problem, by both lay people and the economics literature, follows a Pigouvian intuition: if robots harm workers, they should be taxed. Indeed, in an influential 2017 interview, Bill Gates claimed "the robot that takes your job should pay taxes". More formally, the literature on the subject, such as Guerreiro et al. (2022) and Costinot and Werning (2023), takes as a starting point that, in the absence of perfect redistributive instruments, taxation of automating capital is the natural policy response to automation.

In this paper, I argue that this Pigouvian intuition fails to distinguish between the extensive and intensive margin of automation. One of the key findings of the recent literature modeling automation using the task framework is that the displacement of labor from tasks – the extensive margin – is what adversely affects the wages of exposed workers (Acemoglu and Restrepo, 2018). Once capital is performing a given task, more capital producing that task – the intensive margin – raises overall output via capital deepening. To illustrate, imagine a checkout cashier who is being replaced by a kiosk. This initial act of displacement harms the cashier as the loss of tasks to capital lowers their wages. However, once the task of "checking out customers" has been automated, further investment in kiosks raises demand for the cashier in their remaining tasks. A tax on automating capital is unable to distinguish between these two pecuniary externalities that are opposite in sign.

To study the problem of worker displacement from automation, I consider an optimal taxation problem with a variety of worker types who have differing exposure to automation. Workers make both a labor supply decision and a savings decision. Each type of worker produces their own interval of tasks where they must compete with capital. The equilibrium threshold of automation, above which tasks are allocated to labor, determines wages as a function of workers' comparative advantage across tasks as well as the relative price of capital. The problem of this economy is not one of efficiency, but of equity. Put differently, the economy will land on the Pareto frontier, but because the Planner does not have access to idiosyncratic lump sum transfers à la the Second Welfare Theorem, they cannot necessarily pick the point along the frontier that is socially optimal. Instead, I allow for distortionary redistribution through a general, nonlinear income tax that can tax both capital and labor income.

The supply of capital affects the level of automation in equilibrium, so in principle, the Planner may want to tax capital so as to change the distribution of wages and slacken their incentive compatibility constraints. This is precisely the Naito (1999) critique of the classic Atkinson and Stiglitz (1976) zero capital tax result applied to the case of automation. However, leveraging the insights from the task framework, I consider an alternative policy instrument,

which I call the *threshold rule*, that stipulates how much more expensive labor must be than capital before a given task may be automated. This introduces a wedge in the extensive margin decision to automate a task. In laissez-faire, this number is zero: tasks are automated as soon as labor is more expensive. But with a positive threshold rule, marginal tasks where capital is only slightly cheaper than labor stay with the workers.

I show via an envelope argument that when the threshold rule is set optimally, capital taxes should be zero. This is because changes in the level of automation from changing the threshold rule are proportional to those from changing the supply of capital. Therefore, when the Planner has access to the threshold rule, it is "as-if" the conditions of Atkinson and Stiglitz (1976) hold; any effect on relative wages from changing the supply of capital will be second-order. Intuitively, the threshold rule is a less distortionary way of changing the level of automation. By shifting the marginal task back to labor, the threshold rule generates a first-order increase in wages with only a second-order reduction in output. And it does so without generating a wedge in the capital accumulation decision for tasks that remain automated. In contrast, reducing the supply of capital with a tax will also shift tasks back to labor, but at the cost of generating a wedge in inframarginal tasks as well. A capital tax therefore harms workers by reducing capital deepening at the same time it helps them by slowing automation, whereas the threshold rule is able to focus on the latter without distorting the former.

This novel zero capital tax result does not depend on the welfare weights placed on each type of worker or the productivity schedule for capital and labor. That is, the result holds whether the capital in question is a steam engine, office computer, or superintelligent AGI. Put differently, astonishing progress in artificial intelligence does not, by itself, justify capital taxation.

However, the optimal size of the threshold rule depends on the Planner's concern for those that are displaced by automation and their degree of exposure. To this end, I use a perturbation argument to derive a condition which describes the optimal threshold rule. This condition equalizes the marginal efficiency loss from distorting the task allocation with the marginal equity gain and fiscal impact of reducing automation. The equity gain is given by the covariance of welfare weights and equilibrium changes in income when automation is reduced. Therefore, a higher threshold rule will be desirable when it tends to increase the income of those with relatively high social value of additional consumption.

Finally, I consider the case where the Planner is unable to use the threshold rule. In that case, the Pigouvian intuition prevails: the capital tax should be used to influence the equilibrium level of automation. Moreover, even when nonlinear taxation is available, the capital tax should be linear because the Planner wants all households to internalize the same pecuniary externality of capital on wages. A similar perturbation formula determines the optimal capital tax in this case, but now it must equate the equity benefit of changing income with the direct cost of taxing capital, given by the covariance between welfare weights and

capital ownership. Thus, taxing capital can help to address displacement from automation, but it should be thought of as an indirect and costly way of targeting the real problem: the loss of workers' tasks to capital.

I extend the model in a variety of ways to show the robustness of my extensive vs. intensive margin argument. This includes heterogeneity in skill within types, workers deriving intrinsic value from being a useful part of the production process, and the possibility that workers can switch their type (including a switch to non-employment). In each of these cases, the threshold rule continues to dominate the capital tax as the tool for addressing automation. Of particular interest is the setting where the direction of innovation is endogenous. In a world where the Planner can push the direction of innovation to be more labor- or capital-augmenting, it could be that such intervention is another useful tool for reducing displacement from automation. To the contrary, I show that there is no need for intervention in the innovation side of the economy when the threshold rule is optimal. This is analogous to the Pigou Principle where, for example, optimal carbon pricing implies that innovation policy does not need to pay attention to whether a technology produces carbon emissions (Golosov et al., 2014; Acemoglu et al., 2016; Donald, 2023). One may imagine a similar result holds for an optimal capital tax, but in fact, the opposite is true. The higher the optimal capital tax, the more the Planner should intervene to push innovation towards labor-augmenting technology.

A natural question is how to implement the threshold rule in practice. The value of the threshold rule comes from the fact that it only affects the extensive margin decision to automate, leaving the intensive margin undistorted following automation. This makes it similar to a fixed cost of automation, which would only be worth paying when the gains from using capital are sufficiently high. Building on this intuition, I show that the threshold rule can be implemented via a tax or subsidy on fixed costs for automation. By influencing fixed costs, policy makers can use standard policy instruments to exclusively focus on the extensive margin. In reality, policy makers may not be able to perfectly discern fixed and marginal costs, but the goal of this paper is to argue that policymakers should strive to precisely target the extensive margin of automation whenever possible. While unconventional, there are some precedents for such targeting. For instance, the European Commission considered a policy where Chinese goods may only be imported if they are 10% cheaper than their domestic counterpart.[1] My paper argues that this form of intervention is the right way to deal with displacement from automation.

Ultimately, the degree to which the threshold rule outperforms the more blunt option – a general capital tax – is a quantitative question. Therefore, I conclude the paper with a quantitative exercise showing the gains from using the threshold rule with both optimal non-

[1] For a discussion of this policy proposal, see The Economist: "The Battle for Europe's Economic Soul", March 2023. The trade analog of the policy prescription of this paper would be that, conditional on Chinese goods satisfying this condition, they do not face any tariffs.

linear taxation and the US status quo tax system (Heathcote et al., 2017). What matters for this exercise is the heterogeneity in automation exposure across workers, so to calibrate this heterogeneity, I consider 300 different occupations using the scheme from Deming (2017). I then take the exposure score for each occupation from Webb (2020), which measures the textual similarity between O*NET task descriptions and patents for robots and software. Finally, I calibrate the mapping of this measure to occupation-specific comparative advantage using the estimate of the aggregate elasticity of substitution across capital and labor from Karabarbounis and Neiman (2014). This exercise implies occupation-level elasticities of substitution which are consistent with those measured by Caunedo et al. (2023).

With a fully optimal non-linear tax system, the threshold rule should be set equal to 10.8%, replacing an optimal capital income tax of 10.9% when the threshold rule is absent. This switch eliminates the distortion of household saving decisions and increases utilitarian welfare by a consumption equivalent 0.05%. With the US status quo tax system, the optimal threshold rule rises substantially to 31.8%. This change reflects the greater desirability of distorting automation when the remainder of the economy – mainly capital ownership – is more unequal. In the non-linear tax allocation, the initial endowment is fully taxed and redistributed, allowing the benefits of capital income to be shared more widely. But in the status quo tax system, capital ownership starts out highly concentrated, so reducing automation brings a greater equity benefit. As a result, introducing the optimal threshold rule in the status quo raises welfare by a much larger consumption equivalent 1.4%.

In the final section of the paper, I simulate the introduction of AI and compute the optimal threshold rule in various scenarios of capital-augmenting productivity growth. Occupation-level exposure to AI may differ from previous rounds of automation, so to address this, I modify comparative advantage schedules by including the Webb (2020) exposure score to AI. Unlike robots and software (the scores used in the previous exercise), exposure scores to AI are positively correlated with wages, reflecting the potential for AI tools to displace cognitive white-collar tasks. For various assumptions of how labor-augmenting productivity grows in tandem, the optimal threshold rule rises sharply with capital-augmenting productivity growth. In particular, with the US status quo tax system, doubling capital-augmenting productivity implies the threshold rule should rise to 95.8%. The consequences for welfare are similarly stark, with the consumption equivalent welfare gains from introducing the threshold rule also sharply increasing with capital-augmenting productivity. For instance, introducing the optimal threshold rule in the doubling scenario raises consumption equivalent welfare by an impressive 5%. While somewhat speculative, this exercise implies that the more we expect AI to transform the economy, the more it matters that we get automation policy right.

1.1 Related Literature

This paper builds on the insights of the task framework and applies them to the question of optimal policy. Some of the seminal contributions modeling automation as a reallocation of tasks from labor to capital include Zeira (1998), Acemoglu and Autor (2011), and Acemoglu and Restrepo (2018). Empirical studies arguing for the importance of automation in determining wage inequality include Autor et al. (2003) and Acemoglu and Restrepo (2020, 2022b).[2] An advantage of the task framework from the perspective of optimal policy is that it highlights the distinction between the extensive and intensive margin of automation. In particular, the reasoning behind my main result draws heavily on the notion of "so-so" automation developed in Acemoglu and Restrepo (2018). The margin of automation is, by definition, the place where capital's comparative advantage is the smallest, so distorting the task allocation near this margin has little efficiency cost. Automating these tasks is "so-so" in the sense that it elicits small productivity improvements while still reducing wages of exposed workers. Labor may still be highly productive in these tasks and so wages may be greatly affected by their loss; it is only the productivity difference between capital and labor that is small.

This paper's most significant contribution is to the still fairly young literature on the optimal policy response to automation.[3] As stated above, the existing papers in the literature all take the taxation of automating capital as a starting point, including Guerreiro et al. (2022), Thuemmel (2023), and Costinot and Werning (2023).[4] All three papers argue for the taxation of capital as a form of redistribution over and above already redistributive income taxation. My results clarify the role of capital taxation in the optimal policy response to automation and show that such taxation is an imprecise tool to help workers exposed to automation. Acemoglu et al. (2020) also study an optimal tax problem in a task model, but in stark contrast to my main result, they find that the optimal tax system involves production efficiency: capital and labor are taxed (at similar rates) and the allocation of tasks is left undistorted. The source of this difference is the number of agents. Acemoglu et al. (2020) consider a single type of labor, so they satisfy the conditions of the Diamond and Mirrlees (1971a,b) production efficiency result: all goods and factors are taxed at their own linear rate. In a heterogeneous-agents model, the assumption of a uniform labor tax schedule violates this condition. Thus, production inefficiency can be optimal when there are multiple types of workers. Also related is the work of Lehr and Restrepo (2022) and Beraja and Zorzi (2024), who both focus on the frictional reallocation of displaced workers. These papers argue for the slowing down of automation, but

[2] See also Krusell et al. (2000), Karabarbounis and Neiman (2014), Burstein et al. (2019), Hémous and Olsen (2022), and Moll et al. (2022) for related work on the effect of capital prices and automation on the labor share and inequality.

[3] There are also a number of papers focusing on risk from AI (Guerreiro et al., 2023; Acemoglu and Lensman, 2024; Jones, 2024, 2025; Lehr and Restrepo, 2025). I study labor market disruption rather than extreme risks, but Section 6.2 considers scenarios where highly productive AI technologies can create high levels of automation.

[4] Ales et al. (2015), Hosseini and Shourideh (2020), Schulz et al. (2023), Bond and Kremens (2024), and Atolia et al. (2025) also study the role of income taxation in responding to disruption from technological change.

they do not distinguish between policy instruments that affect the extensive versus intensive margin of automation. I take a complementary view that does not focus on adjustment frictions or unemployment but instead considers how to best offset wage inequality driven by differential exposure to automation across worker groups.[5]

Another important, more developed, literature with which this paper interacts is that on optimal capital taxation (Chamley, 1986; Golosov et al., 2003; Farhi et al., 2012; Saez and Stantcheva, 2018; Straub and Werning, 2020). The main result of this paper, that automation concerns need not affect the capital tax when the threshold rule has been set optimally, is complementary to past capital taxation results that did not consider an automation margin. If the extensive margin of automation can be controlled separately, our past thinking about capital taxation can again become applicable even in the face of automation concerns. It is only in the absence of such targeting that one needs to update their views of the capital tax to account for automation.

The remainder of the paper is organized as follows. Section 2 lays out the task framework that I will use throughout the paper. Section 3 describes my main result on the role, or lack thereof, for capital taxation in responding to automation. Section 4 shows that my main result holds in a variety of extensions. In Sections 5 and 6, I simulate optimal policy within a calibrated version of my model. Section 7 concludes.

2 Model

This section lays out the task model used throughout the paper. In the main text, I consider a two period model, but I show in Section 4.1 that all of the insights carry over to an infinite horizon setting as well. There are J-many worker types, which I will index with i or j. To keep things simple, I will imagine that income in the first period is given by a fixed, type-specific endowment and workers make a consumption/savings decision over this income. In the second period, they make a labor supply decision and consume their capital and labor income. Factor prices will be endogenously determined by the allocation of capital and labor to tasks, where each worker type can have their own comparative advantage schedule with respect to capital.

2.1 Preferences & Welfare

Household preferences are given by a utility function that is additively separable across consumption and labor:

$$V_j = u(c_{j,0}) + \beta(u(c_{j,1}) - h_j(l_j)), \tag{1}$$

[5] Automation and unemployment are often conflated, but in this setting, I am focusing exclusively on automation, as defined by capital displacing labor from productive tasks. Whether the resulting reductions in the marginal product of labor lead to *unemployment* (workers that would like to supply labor at prevailing wages but cannot find a buyer) depends on whether there are frictions in the labor market. Here I am abstracting from such frictions and just focusing on differential exposure to task displacement.

where β is the discount factor common across households.

Social welfare is given by some function over population-weighted, type-specific utility:

$$W = \mathcal{W}(\{n_j V_j\}) \tag{2}$$

where n_j is the population size of type j households. I will normalize the total size of the population to one: $\sum_j n_j = 1$. The generality of the welfare function implies that one can also view the optimal policy problem described below as sketching the constrained Pareto frontier; one would simply need to vary the welfare weights placed on each household type.

2.2 Tasks & Production

Final output is produced in the second period using a CES aggregator over task-level production:

$$Y = \Big(\sum_j \int_{\mathcal{X}_j} y(x)^{\frac{\sigma-1}{\sigma}} dx \Big)^{\frac{\sigma}{\sigma-1}}, \tag{3}$$

where σ is the elasticity of substitution across tasks, and x indexes over the tasks in the economy. $\mathcal{X}_j$ is the set of tasks that is specific to type j workers. Production at the task-level is perfect substitutes across capital and labor:

$$y(x) = A_k \alpha_k(x) k(x) + A_j \alpha_j(x) \ell_j(x) \quad x \in \mathcal{X}_j, \tag{4}$$

where A_k, A_j are factor-augmenting productivities and $\alpha_k(x), \alpha_j(x)$ are task-level productivities for capital and labor, respectively. This setup assumes that worker types do not compete with each other; workers only compete with capital.

Total factor demand must equal supply, so

$$L_j = \int_{\mathcal{X}_j} \ell_j(x) dx, \tag{5}$$

where $L_j \equiv n_j l_j$ is the total amount of labor supplied by type j households. Similarly for capital, we have

$$K = \sum_j \int_{\mathcal{X}_j} k(x) dx, \tag{6}$$

where K is the total supply of capital.

In the first period, there is a fixed endowment of final goods $\bar{Y}$, which can be either consumed or transformed into capital. One could think about the first period as taking place after production when the only remaining decision is how much to consume and save. Therefore,

the first period resource constraint requires

$$\sum_j n_j c_{j,0} + K = \bar{Y}. \tag{7}$$

Production takes place in the second period, when all final goods and remaining capital are consumed. Thus, the second period resource constraint requires

$$\sum_j n_j c_{j,1} = Y + (1-\delta)K, \tag{8}$$

where δ is the depreciation rate of capital.

2.3 Automation & Equilibrium

The equilibrium allocation of tasks to factors is determined by the profit maximization of a competitive, representative firm. This firms solves

$$\max_{\{k(x),\{\ell_j(x)\}\}} Y - \sum_j w_j \int_{\mathcal{X}_j} \ell_j(x)dx - r\sum_j \int_{\mathcal{X}_j} k(x)dx, \tag{9}$$

where w_j is the type-specific wage, and r is the rental rate on capital. I assume that the final good is the numeraire. Because task-level production is perfect substitutes, the firms would like to produce each task with the lowest marginal cost factor. The threshold rule θ introduces a wedge in this decision and stipulates that tasks may only be automated when the marginal cost of labor exceeds that of capital by factor $1+\theta$. Therefore, task allocations follow

$$\Omega_j = \Big\{ x \in \mathcal{X}_j \Big| \frac{w_j}{A_j \alpha_j(x)} < \frac{r(1+\theta)}{A_k \alpha_k(x)} \Big\} \tag{10}$$

$$\Omega_k = \Big\{ x \in \cup_j \mathcal{X}_j \Big| \frac{w_j}{A_j \alpha_j(x)} \geq \frac{r(1+\theta)}{A_k \alpha_k(x)} \Big\}. \tag{11}$$

These sets describe the equilibrium set of tasks that is assigned to each factor in production.[6] The threshold rule introduces *production inefficiency* because failing to use the cheapest factor in each task implies the production function does not maximize output for a given set of factor inputs. To add some additional structure to the task allocation, I will assume that task sets are simply an interval: $\mathcal{X}_j = [0, \bar{x}]$. Furthermore, with the assumption that $\frac{\alpha_j(x)}{\alpha_k(x)}$ is strictly increasing in x (i.e. labor has a comparative advantage in higher order tasks), we will have a threshold of automation x_j below which tasks are automated and above which tasks are produced with labor. This threshold task must then satisfy the following marginal cost

[6] I am assuming that the marginal task is allocated to capital, but this is not relevant as this task is measure zero.

equalization condition:

$$\frac{w_j}{A_j\alpha_j(x_j)} = \frac{r(1+\theta)}{A_k\alpha_k(x_j)}. \tag{12}$$

This condition allows us to describe each type's comparative advantage schedule across tasks. We have that the changes in automation must follow

$$dx_j = \Big(\frac{\partial \ln\left(\alpha_j(x_j)/\alpha_k(x_j)\right)}{\partial x_j}\Big)^{-1} d\ln\Big(\frac{w_j}{r(1+\theta)}\Big), \tag{13}$$

That is, the semi-elasticity of relative productivity provides a measure of a type's comparative advantage with respect to capital because it tells us how much their automation threshold shifts in response to price changes. If $\frac{\alpha_j(x)}{\alpha_k(x)}$ is flat, small changes in prices can generate large changes in automation, but if $\frac{\alpha_j(x)}{\alpha_k(x)}$ is steep, then automation is largely insensitive to price changes.

Given its importance for optimal policy, some additional discussion of the threshold rule is warranted. Figure 1 provides an illustration of the introduction of the threshold rule (holding factor prices fixed). The left panel shows the task allocation in laissez-faire, where tasks are produced with the cheapest factor. This is the task allocation that maximizes output. With a positive threshold rule, some marginal tasks are given back to labor, namely those where the marginal cost of producing with labor is less than $1+\theta$ times greater than that of producing with capital. This generates a distortion, as illustrated by the red deadweight loss triangle, but this distortion only stems from the marginal cost difference between capital and labor in the marginal tasks that are given back to labor. That is, this is an extensive margin distortion that is initially second order because there is no cost difference between capital and labor at the marginal task. The inframarginal tasks, where capital is sufficiently cheaper than labor, are left undistorted.

Following a standard argument in the task literature, I show in Appendix A.1 that final output can be represented as a CES production function with endogenous shares:

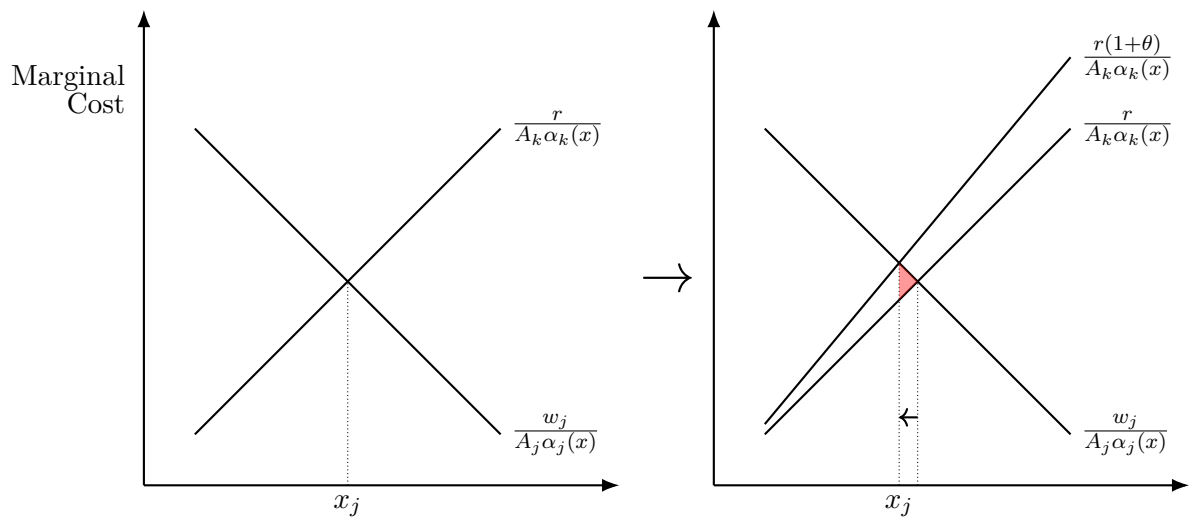
$$F(X,L,K) = \Big(\Lambda_k^{\frac{1}{\sigma}}\left(A_kK\right)^{\frac{\sigma-1}{\sigma}} + \sum_j \Lambda_j^{\frac{1}{\sigma}}\left(A_jL_j\right)^{\frac{\sigma-1}{\sigma}}\Big)^{\frac{\sigma}{\sigma-1}}, \tag{14}$$

where $\Lambda_k \equiv \int_{\Omega_k}\alpha_k(x)^{\sigma-1}dx$ and $\Lambda_j \equiv \int_{\Omega_j}\alpha_j(x)^{\sigma-1}dx$, and $X \equiv [x_j]$ is the vector of automation thresholds. Therefore, one should think about the task allocation as determining the CES shares Λ_j assigned to each factor, and as a result, factor prices in equilibrium. Factor prices follow the familiar formulas:

$$w_j = \Lambda_j^{\frac{1}{\sigma}}A_j^{\frac{\sigma-1}{\sigma}}\Big(\frac{Y}{L_j}\Big)^{\frac{1}{\sigma}} \tag{15}$$

$$r = \Lambda_k^{\frac{1}{\sigma}}A_k^{\frac{\sigma-1}{\sigma}}\Big(\frac{Y}{K}\Big)^{\frac{1}{\sigma}}. \tag{16}$$

Figure 1: Task Allocation with Threshold Rule



From this, we can see the two pecuniary externalities that capital imposes on labor. If we write the automation thresholds x_j as endogenous equilibrium objects that depend on factor supplies and the threshold rule, we can write

$$\frac{\partial \ln(w_j)}{\partial \ln(K)} = \underbrace{\frac{1}{\sigma}\frac{\partial \ln(\Lambda_j)}{\partial x_j}\frac{\partial x_j}{\partial \ln(K)}}_{\text{Automation (-)}} + \underbrace{\frac{1}{\sigma}\frac{\partial \ln(Y)}{\partial \ln(K)}}_{\text{Capital Deepening (+)}} . \tag{17}$$

As we can see, increasing the capital stock lowers wages by expanding the set of tasks that are automated. That is, more abundant capital lowers the rental rate and makes it cheaper to produce some marginal tasks with capital. Conversely, this increase in the capital stock also raises wages through capital deepening. With CES production, factor prices scale with output, so increases in capital raise demand for workers by simply expanding the economy. Note that the automation margin varies by type as it will depend on the comparative advantage schedule of a given type across tasks. The capital deepening margin, on the other hand, has the same proportional effect on all wages. This implies that the effect of capital on relative wages only operates through differential exposure to automation. In effect, I am ruling out other forms of capital-skill complementarity that are not related to automation (such as those studied in Slavík and Yazici (2014)) so as to keep focus on automation as a source of inequality.

Heterogeneity in exposure to automation also implies type-specific elasticities of substitution between capital and labor. In a task model, the elasticity of substitution comes from substitution both across tasks and within tasks (Hubmer and Restrepo, 2021), so each type of labor will have a local elasticity of substitution that depends on their comparative advantage

schedule:

$$\Sigma_j \equiv \frac{\partial \ln\left(L_j/K\right)}{\partial \ln\left(r/w_j\right)} = \sigma + \frac{\frac{\alpha_j(x_j)^{\sigma-1}}{\Lambda_j} + \frac{\alpha_k(x_j)^{\sigma-1}}{\Lambda_k}}{\frac{\partial \ln\left(\alpha_j(x_j)/\alpha_k(x_j)\right)}{\partial x_j}}. \tag{18}$$

The first term is common across types and reflects substitution *across* tasks. That is, if relative prices shift, demand will shift towards tasks produced by cheaper factors. The second term is heterogeneous across types and reflects substitution *within* tasks. As the relative price of capital and labor changes, automation thresholds change according to the comparative advantage of labor across tasks, $\frac{\partial \ln\left(\alpha_j(x_j)/\alpha_k(x_j)\right)}{\partial x_j}$, which shows up in the denominator.

We can then define an aggregate elasticity of substitution by considering how the total share of income captured by capital changes in response to changes in the price of capital:

$$\Sigma_k \equiv 1 - \frac{\partial \ln\left(S_k\right)}{\partial \ln\left(r\right)} = \sum_j S_j\Big(1 - \frac{\partial \ln\left(w_j\right)}{\partial \ln\left(r\right)}\Big)\Sigma_j. \tag{19}$$

where S_k is the capital share and S_j is the income share of each type. This aggregate elasticity of substitution is then a weighted average of the type-specific elasticities.[7] The weights – which must sum to one – are determined by income shares and the sensitivity of wages to changes in the price of capital. Therefore, we have that the aggregate level of substitutability across capital and labor is also given by across and within task substitution, where the influence of each type's heterogeneous exposure to automation scales with the income share of that type.

This economy is efficient, so the role of policy is to address equity. In other words, the allocation will land on the Pareto frontier, but there is no guarantee that market endowments will select the point on the frontier that is optimal from the perspective of social welfare. To address this equity concern, the Planner has access to a nonlinear tax function $\mathcal{T}(R\kappa_j, w_j l_j)$, the key restriction of which is that it can only condition on financial observables (capital and labor income), rather than type. This rules out type-specific lump-sum transfers (like those used in the Second Welfare Theorem) and type-specific linear labor taxes (like those used in Diamond and Mirrlees (1971a,b)). Hence, redistribution requires introducing distortions, and the key question of this paper is how the inequality implications of automation can be addressed with as little distortion as possible.

Households then maximize utility subject to their budget constraint, which follows

$$Rc_{j,0} + c_{j,1} = R\omega_j\bar{Y} + w_j l_j - \mathcal{T}(R\kappa_j, w_j l_j), \tag{20}$$

where $R \equiv 1 + r - \delta$ is the total pre-tax return to savings, $\kappa_j \equiv \omega_j\bar{Y} - c_{j,0}$ is the level of household savings between periods, and ω_j is the per-person ownership share of the initial endowment. One should note that this setup allows for a "universal basic income" (UBI) –

[7] Note that Equation (19) holds for any constant returns production function $F\big(K, \{L_j\}\big)$.

another common policy response to automation – as the intercept of the tax function. I do not put the initial endowment in the tax function to keep focus on distortionary taxation, but one could allow for exogenous taxes on the initial endowment by simply relabeling ω_j. For example, in the Mirrlees problem of Section 3.1, I will assume the Planner fully taxes and uniformly redistributes the initial endowment before then deciding the distortionary, non-linear tax function. Denote by $\tau_{k,j} \equiv \frac{\partial \mathcal{T}_j}{\partial R\kappa}$ and $\tau_{l,j} \equiv \frac{\partial \mathcal{T}_j}{\partial wl}$ the implicit tax wedges on capital and labor, evaluated at type j's bundle.[8] Then, standard household optimization decisions satisfy

$$\frac{h_j'(l_j)}{u'(c_{j,1})} = (1 - \tau_{l,j})w_j, \tag{21}$$

as well as

$$u'(c_{j,0}) = (1 - \tau_{k,j})R\beta u'(c_{j,1}). \tag{22}$$

Finally, the tax schedule must satisfy the government budget constraint:

$$\sum_j n_j \mathcal{T}_j = 0. \tag{23}$$

Given these conditions, the competitive equilibrium is defined as follows:

Definition 1 (Equilibrium)**.** *Given a level of initial endowment $\bar{Y}$, ownership shares $\{\omega_j\}$, population $\{n_j\}$, factor-augmenting productivities $\{A_k, \{A_j\}\}$, and task-level productivities $\{\alpha_k(x), \{\alpha_j(x)\}\}$, an equilibrium consists of a tax schedule $\mathcal{T}(.)$, a threshold rule θ, final output Y, factor demand $\{k(x), \{\ell_j(x)\}\}$, a rental rate r, wages $\{w_j\}$, task allocations and automation thresholds $\{\Omega_k, \{\Omega_j, x_j\}\}$, capital supply K, labor supply $\{l_j\}$, and consumption $\{c_{j,0}, c_{j,1}\}$ such that:*

(i) Prices $\{r, \{w_j\}\}$ and quantities $\{Y, \{k(x), \{\ell_j(x)\}\}\}$ solve the profit maximization problem (9);

(ii) Task allocations and automation thresholds $\{\Omega_k, \{\Omega_j, x_j\}\}$ satisfy (10), (11), and (12);

(iii) Labor supply $\{l_j\}$ and consumption $\{c_{j,0}, c_{j,1}\}$ solve households' utility maximization problem by satisfying their budget constraint (20), labor supply condition (21), and Euler equation (22);

(iv) Factor markets clear (5) and (6);

(v) The resource constraints (7) and (8) and government budget constraint (23) hold.

[8] It would be more accurate to call $\tau_{k,j}$ a wealth tax, but this distinction does not matter in a model where all savings earn the same return.

3 Automation & Capital Taxation

This section describes the optimal policy response to automation. Optimal choice of the threshold rule implies that the capital tax should be zero. Instead of taxing capital, the Planner prefers to distort the extensive margin of automation with the threshold rule, and I derive a formula which characterizes the optimal threshold rule. However, in the absence of the threshold rule, the Planner wants to linearly tax capital so as to cause savers to internalize the pecuniary externality of automation, and I derive a formula to characterize the optimal capital tax in this case.

3.1 Non-Linear Taxation

I start with a mechanism design formulation of the Planner's optimal taxation problem, where the Planner must satisfy incentive compatibility constraints which reflect that an agent's type is private information (Mirrlees, 1971).

Definition 2 (Non-Linear Tax Planning Problem)**.** *The Planner solves*

$$\max_{\{c_{j,0},c_{j,1},l_j\},K,\theta} \mathcal{W}(\{n_j V_j\}) \;\; s.t. \;\; (7),(8), \tag{24}$$
$$u(c_{i,0}) + \beta\big(u(c_{i,1}) - h_i(l_i)\big) \geq u(c_{j,0}) + \beta\big(u(c_{j,1}) - h_i(w_j l_j/w_i)\big) \quad \forall i,j.$$

That is, the Planner must pick an allocation where it is incentive compatible for households to truthfully report their type, and we can back out the properties of the tax function that would implement this allocation as a competitive equilibrium. The role of automation in this problem is in the determination of relative wages. Because the Planner can only observe income, rather than type, a household of type i that is imitating type j will need to supply $w_j l_j/w_i$ units of labor to achieve type j's income. Therefore, influencing the equilibrium level of automation may slacken the incentive compatibility constraints by compressing the wage distribution and reducing the temptation for one type to imitate another.

The only policy variable that explicitly shows up in the Planner's problem is the threshold rule as this policy instrument allows them to influence wages through automation decisions. Wages are determined by automation thresholds and factor supply, and the automation thresholds in turn depend on factor supply and the threshold rule. Therefore, we can write an implicit function that characterizes the equilibrium dependence of automation thresholds on factor supply and the threshold rule and use this implicit function to consider the full impact of changes in these variables on wages. The following Lemma characterizes the key feature of this implicit function.

Lemma 1. *The equilibrium change in automation thresholds from raising capital supply is*

proportional to that from changing the threshold rule. That is,

$$\frac{\partial x_j}{\partial K} \propto \frac{\partial x_j}{\partial \theta}. \tag{25}$$

The proof of Lemma 1 is in Appendix A.3. Lemma 1 shows that there is an equivalency between influencing the equilibrium level of automation by changing the supply of capital and by changing the threshold rule, and this implies the main result of the paper.

Proposition 1. *When the threshold rule is set optimally, capital taxes should be zero:*

$$\tau_{k,j} = 0 \quad \forall j. \tag{26}$$

The proof of Proposition 1 can be found in Appendix A.4.[9] This result tells us that capital taxation is not the right tool for dealing with automation when the threshold rule is available. In principle, the Planner may want to influence the supply of capital so as to change relative wages through automation. However, Lemma 1 implies that any impact on the Planner's objective from changing automation through the supply of capital will envelope out when the threshold rule is chosen optimally. Thus, it is "as if" Atkinson and Stiglitz (1976) holds. With automation targeted by the threshold rule, there is no reason to introduce a wedge between households' intertemporal marginal rate of substitution and the marginal rate of transformation of final goods across time; any slackening of incentive compatibility by changing equilibrium automation will already have been achieved by choice of the threshold rule. Put differently, every household should satisfy an undistorted Euler equation.[10]

This is a qualitative result in the sense that it holds independent of the parameter values of the economy. Namely, it does not depend on the details of the social welfare function, such as the level of inequality aversion, and it does not depend on the productivity schedule of factors across tasks. Therefore, the result holds whether the capital in question is a primitive tool or a science fiction robot capable of producing all tasks. The model can encompass the latter case as a vast increase in $\alpha_k(x)$ for tasks that are typically associated with labor. The main point is that it is better to explicitly target the extensive margin of automation. We will now turn to characterizing the optimal threshold rule.

3.2 Distorting Automation

Turning now to the optimal choice of threshold rule, I will make use of the other great tradition of optimal taxation: the perturbation approach popularized by the seminal work of Saez (2001). This approach allows me to derive a formula for the optimal threshold rule in terms

[9] Note that everything would continue to hold with the more general utility specification $V_j = \mathcal{V}_j\big(\mathcal{U}(c_{j,0}, c_{j,1}), l_j\big)$.

[10] This also implies that there should be no taxation of capital in any subset of the automated tasks as this would still generate a wedge in households' Euler equations.

of a perturbation around the optimum.[11] For this result, covariance and expectations operators will make use of population size $\{n_j\}$ as a probability distribution, and I will denote by $\lambda_j \equiv \partial_j \mathcal{W} u'_{j,1} / \mathbb{E}(\partial_i \mathcal{W} u'_{i,1})$ the welfare weight on type j households. The following proposition describes the optimal choice of the threshold rule.

Proposition 2. *For any tax function that allows for a freely chosen lump-sum, the optimal threshold rule is set to satisfy*

$$Y \sum_j \Big(\frac{r}{A_k \alpha_k(x_j)} \Big)^{1-\sigma} \frac{(1+\theta)^{1-\sigma} - 1}{1-\sigma} (-dx_j) = Cov(\lambda_j, (1-\tau_{l,j}) w_j l_j d\ln(w_j) + (1-\tau_{k,j}) R\kappa_j d\ln(R)) + \mathbb{E}(\tau_{l,j} w_j l_j d\ln(l_j) + \tau_{k,j} R\kappa_j d\ln(\kappa_j)) \tag{27}$$

for perturbations generated by changes in the threshold rule.

The proof of Proposition 2 can be found in Appendix A.5, which follows a similar argument to that found in Costinot and Werning (2023). The left-hand side of Equation (27) expresses the efficiency cost of distorting the extensive margin of automation. This depends on the income share of each marginal task as well as the scale of the distortion generated by the threshold rule. The former – given by the marginal cost of capital at the threshold of automation raised to one minus the cross-task elasticity of substitution – measures the importance of the marginal task in production à la Hulten's Theorem (Domar, 1961; Hulten, 1978). The latter measures the opportunity cost of using the "wrong" factor near the marginal task, which scales with the size of θ. In fact, the scale of the distortion has the same shape as a CRRA utility function with curvature given by the cross-task elasticity of substitution. As I show formally in Appendix A.8, this gives the natural interpretation that the scale of distortion is logarithmic when task aggregation is Cobb-Douglas. Both of these terms are then scaled by the equilibrium reduction in automation $-dx_j$ and output Y.

The right-hand side is the equity and fiscal benefit of perturbing the threshold rule. The equity benefit is the covariance between a type's welfare weight and the total equilibrium change in their after-tax income from the perturbation. This term will be high if the Planner values the consumption of those whose income rises relative to those whose income falls. As I will argue below, the distribution of equilibrium changes in income is closely connected to heterogeneity in exposure to automation. The fiscal benefit is the effect on government revenue of behavioral changes brought about by the perturbation. The supply of capital and labor may adjust in response to changes in the threshold rule, and these factor supply changes will effect the government insofar as those factors are taxed. Therefore, the perturbation is also beneficial

[11] Alternatively, Appendix A.4 derives a formula for the optimal threshold rule within the mechanism design framework, equating the marginal production loss from distortion with the marginal value of slackening the incentive compatibility constraints by compressing the wage distribution.

insofar as it slackens the government's budget constraint. Note that these behavioral effects do not affect welfare directly due to the envelope theorem.

The formula of Proposition 2 does not depend on the tax schedule having been chosen optimally. It holds for any tax function that allows for a freely chosen lump-sum tax or transfer which does not need to condition on type. This is required to guarantee that perturbations remain fiscally feasible. Proposition 2 therefore allows for the introduction of a threshold rule without needing to stipulate the remainder of the tax function. I use this property in Section 6 to study the welfare gain from introducing the threshold rule in the US economy while maintaining the status quo tax system.

The perturbations of factor prices considered here are general equilibrium responses, taking into consideration the full, optimal response of every agent in the economy to a perturbation of the threshold rule $d\theta$. But to gain intuition as to how wages change for different types, we can note that the comparative advantage of a type across tasks $\frac{\partial \ln(\alpha_j(x_j)/\alpha_k(x_j))}{\partial x_j}$ will determine the size of the direct reduction in automation in response to increases in the threshold rule. Again, if a type's relative productivity schedule across tasks is steep, then there will be few tasks that are near the automation margin. Therefore, increases in the threshold rule will not do much to raise that type's wages directly. Conversely, if a type's relative productivity schedule across tasks is flat, many of their tasks will be close to the margin of automation, implying that small changes in the threshold rule can return to them a large chunk of tasks. And importantly, these will be tasks where they were almost as cheap as capital, minimizing the production losses.

The threshold rule therefore leverages differences in comparative advantage across worker types to move the wage distribution. As shown by Proposition 2, the desirability of using the threshold rule then depends on to what degree this differential exposure – and the resulting general equilibrium changes to income – covary with the relative social desirability of increasing a given type's consumption. This raises two questions. First, there is no apriori guarantee that this covariance will be positive. This is a quantitative question that will vary with the structural parameters of the economy. As I argue in Section 6, occupations that have high exposure to automation from their "routiness" tend to also have high (utilitarian) welfare weights, justifying a positive threshold rule, but there is no guarantee that this positive relationship between automation exposure and welfare weights will hold in every episode of automation.

Second, why is moving the wage distribution with the threshold rule a better way to redistribute income than the income tax? Moving tasks away from the cheapest factor generates production inefficiency, so could we not simply redistribute the productivity gain from eliminating the threshold rule in a way that compensates the losers? This is the line of reasoning typically used to rule out commodity taxation (Laroque, 2005; Kaplow, 2006). The answer to these questions comes from the fact that the variation in the income distribution that one can bring about with the threshold rule cannot, in general, be replicated with the income tax. Put

differently, these two instruments span different variation in income space, so compensating the losers of automation will not always be budget feasible. The following corollary formalizes this point.

Corollary 1. *Production efficiency $\theta = 0$ is optimal if a budget-feasible tax reform can reproduce the threshold rule perturbations to household income before behavioral effects. This requires*

$$\mathbb{E}\Big(\left[\tau_{l,j} w_j l_j \varepsilon^{h,j}_{l,w} + \tau_{k,j} R \kappa_j \varepsilon^{h,j}_{\kappa,w}\right] \left[d_\theta \ln\left(w_j\right) - d_\mathcal{T} \ln\left((1-\tau_{l,j}) w_j\right)\right] \\ + \left[\tau_{l,j} w_j l_j \varepsilon^{h,j}_{l,R} + \tau_{k,j} R \kappa_j \varepsilon^{h,j}_{\kappa,R}\right] \left[d_\theta \ln\left(R\right) - d_\mathcal{T} \ln\left((1-\tau_{k,j}) R\right)\right] \Big) = 0, \tag{28}$$

where $\varepsilon^{h,j}_{q,p}$ is type j's Hicksian elasticity of factor q with respect to after-tax price p, and $d_\theta, d_\mathcal{T}$ denote perturbations generated by the threshold rule and tax reform, respectively.

The proof of Corollary 1 can be found in Appendix A.6. It tells us what must be true if we are to compensate the losers with the income tax. Imagine constructing a simultaneous perturbation of the threshold rule and income tax that leaves everyone's utility unchanged. That is, the first-order impact of this joint perturbation on income before behavioral effects is zero because the general equilibrium price changes from changing both instruments are offset by changes in taxes, by construction. The question then is whether this joint perturbation is feasible for the government's budget.[12] For this to be true, we need to consider the behavioral effects. By construction, there are no income effects, so for this joint perturbation to be revenue neutral, it must be that all changes in tax revenue from substitution effects must sum to zero. This is why Corollary 1 is written in terms of Hicksian elasticities.

The difficulty of meeting this condition comes from the fact that all types must face the same income tax schedule. Automation exposure varies by type, so each type needs their income perturbed by the tax reform in a way that is equal and opposite to their exposure. But the income tax cannot treat people with the same income level differently even when the labor generating that income has very different levels of exposure to automation. And because one type's marginal tax is another type's inframarginal tax, there is no reason to expect that the adjustments to the tax system that compensate the losers will also adjust marginal taxes in such a way that behavioral effects are revenue neutral.

It is also helpful to consider a case where we are guaranteed to satisfy this condition: type-specific linear taxes. In that case, each type's tax can be changed so as to exactly offset the changes in factor prices from the joint perturbation. This kills two birds with one stone in that it both does not change income (keeping utility constant) and does not change after-tax prices (eliminating behavioral effects). As expected, this implies the famous production efficiency result of Diamond and Mirrlees (1971a,b).

[12] Schulz et al. (2023) consider a similar question and also derive conditions for when economic shocks can be compensated with changes in the income tax.

The reasoning of this section has relied on the distinction between the extensive and intensive margin of automation. This distinction is a key feature of the task model and shows up prominently in the task-based literature. Discussing its empirical implications, Restrepo (2023) argues that different empirical settings will give different answers about the effects of capital on workers in so far as they focus more on one of these margins than the other. This includes Aghion et al. (2022), who find that reductions in the cost of investment lead to *greater* demand for labor, underlining that blunt increases in the cost of capital may be harmful to workers on net. de Souza and Li (2023) uses natural language processes to distinguish between capital that can replace workers (robots) and that which can complement workers (tools), finding very different effects on employment and wages between the two. Finally, Jones and Tonetti (2025) consider the growth effects of automation and find that the largest TFP gains from historical episodes of automation have come from cumulative gains on the intensive margin, rather than the initial switch from labor to capital.

The task-base literature, including the papers just cited, has predominantly focused on the *descriptive* implications of the extensive vs. intensive margin distinction, but here I am considering the *normative* implication of this distinction that the threshold rule dominates the capital tax. As the evidence shows, capital can affect wages, employment, and TFP very differently if it operates on the extensive or intensive margin, and policy should take on board this difference by focusing on the extensive margin. Nonetheless, this may not always be possible, so it is informative to consider the optimal capital tax whenever the threshold rule is unavailable. In that case, the Naito (1999) critique of Atkinson and Stiglitz (1976) begins to bite, leaving room for the Pigouvian intuition to argue for capital taxation.

Proposition 3. *When the threshold rule is unavailable, capital taxes should be linear and satisfy*

$$\begin{aligned} Cov(\lambda_j, R\kappa_j) = Cov(\lambda_j, (1-\tau_{l,j})w_j l_j d\ln(w_j) + (1-\tau_k)R\kappa_j d\ln(R)) \\ + \mathbb{E}(\tau_{l,j} w_j l_j d\ln(l_j) + \tau_k R\kappa_j d\ln(\kappa_j)). \end{aligned} \tag{29}$$

The proof of Proposition 3 can be found in Appendix A.7. The first substantive feature of this result is that the capital tax should be linear, reducing the returns to capital for every household by the same amount. This is consistent with a Pigouvian interpretation where the goal of the instrument is not redistribution per se but to make households internalize the pecuniary externality of capital on relative wages. With this established, a similar perturbation argument to before gives us a condition that the optimal capital tax must satisfy.

At the optimum, the capital tax equalizes the covariance of capital income and welfare weights with, as before, equity effects from general equilibrium price changes and fiscal effects. Now the left hand side represents the direct welfare cost (net of fiscal benefits) of taxing capital income, which will be high in so far as capital ownership positively covaries with welfare weights.

This is analogous to the production efficiency cost of Equation (27). The right hand side has the same interpretation as before: it tells us the benefit of perturbing the capital tax from both the welfare-weighted impact of price changes on each type's income as well as the fiscal impact of behavioral effects. Automation exposure again shows up in the equity term as raising the cost of capital will reduce automation and all else equal send more tasks back to those with the lowest comparative advantage. However, this will now be set against a reduction in capital deepening in the tasks that remain automated, which reduces wages for every type.

4 Extensions

This section lays out how the main argument of this paper – that automation policy should target the extensive margin of automation rather than tax capital – carries over to a wide range of environments. Furthermore, I show that the threshold rule can be implemented as a tax (or subsidy) on fixed costs to automate tasks.

4.1 Infinite Horizon

First, I consider an infinite horizon setting. I have thus far considered a two period setting for the sake of expositional simplicity, but I show in Appendix B.1 that all of the insights of the two period setting carry over to a richer dynamic environment. In particular, there should not be any wedges in savings decisions (i.e. capital taxes) whenever the threshold rule is set optimally at each point in time. Furthermore, the formulas of Propositions 2 and 3 continue to hold with minor adjustment. I also allow the task set of each type $\mathcal{X}_{jt}$ to (exogenously) change over time, so the much discussed possibility that workers can "find new tasks" is a separate consideration from the arguments of this paper (Acemoglu and Restrepo, 2018; Autor et al., 2024).

4.2 Within-Type Heterogeneity

Next, I allow for there to be heterogeneity in the endowment of efficiency units of labor within types. To keep things simple, my baseline model assumed that all workers of a given type are the same, but of course, this is not the case in reality. To model within-type heterogeneity, I make the standard assumption that workers can vary in their skill in the sense that a unit of labor can imply different amounts of efficiency units of labor for different workers of a given type. Thus, a worker's type determines their wage per efficiency unit of labor – and therefore their exposure to automation – while their skill determines the return to their labor effort. For the Planner, both of these attributes are private information.

As I show in Appendix B.2, allowing for heterogeneity does not change the core argument. It is still the case that capital taxation is dominated by the threshold rule, and the formulas

of Propositions 2 and 3 just need to be adjusted to also integrate over the skill distribution within each type. Therefore, the lessons of the baseline model were not an artifact of assuming away within-type heterogeneity.

4.3 Intrinsic Value from Work

A common concern in discussions of automation is that people derive value from work, and so losing tasks to capital may reduce the meaning people achieve in their lives. This is an interesting but highly speculative concern, quite different from the standard practice of modeling work as an unambiguous bad that one sells in exchange for consumption. I will not be able to do justice to the full range of possible ways in which this concern could play out, but in Appendix B.3, I explore how my findings may interact with such concerns. In particular, I allow workers to value being useful, as proxied by their marginal product. This could come from both intrinsic motivation (people simply like being useful) or instrumental motivation (political power is in some way related to one's own economic importance).

In this setting, the envelope argument underlying Proposition 1 continues to hold, so capital taxation is not the right instrument to address the distributional effects of automation. Interestingly, under the assumption that the value from one's marginal product is additively separable from the disutility of labor effort, we have that it is optimal to *subsidize* capital. This is because capital deepening raises marginal products, creating a positive externality through the value people derive from their work. Therefore, the concern that automation may rob people of an important source of meaning does not necessarily create a justification for addressing automation with a capital tax. In fact, the opposite may be the case.

4.4 Type Switching

In my baseline model, a worker's type is immutable. In reality, people can change their occupation, and the possibility (or lack thereof) of doing so is a common feature of discussions about the displacement effect of automation. To model the possibility that a worker may change their type, I show in Appendix B.4 how the discrete choice framework that is commonly used to model dynamic migration decisions can be applied to my setting (Caliendo et al., 2019; Donald et al., 2025a).[13] Namely, I assume that workers are assigned an occupation in the initial period, but they also receive an array of preference shocks for each occupation and pick at the conclusion of the initial period the occupation that maximizes their continuation utility inclusive of this shock. I do not make any assumptions over the distribution of these preference shocks, so they can include a common component reflecting the difficulty of moving from one occupation to another as well as an idiosyncratic component that causes individuals to differ in their occupational choice. They also allow for arbitrary elasticities of movement

[13] Kurnaz et al. (2025) use this framework to study the long-run effect of taxes on worker's choice of occupation.

across occupations, in part in response to differing levels of exposure to automation. I also assume that preference shocks are private information. The only substantive assumptions I make are that the preference shocks are additive, and following Donald et al. (2025b), that welfare is utilitarian within types but across shocks.

With this setup, Proposition 1 continues to hold: there should be no distortion of savings with a capital tax whenever automation concerns can be addressed with the threshold rule. It is worth noting that one occupation can be interpreted as exiting the labor market, so this setup nests the case where automation leads people to no longer work. There are no frictions in the labor market,[14] so these are voluntary exits, rather than unemployment per se. Therefore, in a world where workers can transition to other occupations (or out of the labor market), the threshold rule still dominates capital taxation as a tool for dealing with the displacement effect of automation.

4.5 Endogenous Innovation

My baseline model assumes that innovation is exogenous: the productivity schedule of capital and labor in the second period is simply taken as given. However, a common question in the automation literature is how factor prices and other incentives affect the direction of innovation (Acemoglu and Restrepo, 2018, 2022a; Hémous et al., 2025). Namely, it is important to know how policy affects the choice to innovate in technologies that augment labor or capital and whether additional intervention in the innovation side of the economy is warranted to influence the direction of innovation. In Appendix B.5, I modify the directed innovation model of Donald (2023) to study these questions.

My main finding is that policy does *not* need to influence the direction of innovation when the optimal threshold rule is in place. The threshold rule already accounts for the distributional consequences of automation, and absent other reasons for intervention such as knowledge spillovers,[15] market incentives select the correct mix of innovation across capital and labor. This reasoning is analogous to the Pigou Principle, where instruments should separately solve the problem they are meant to target. For example, in the climate literature, setting an optimal carbon price implies that innovation policy can ignore whether a technology is "clean" or "dirty" (Golosov et al., 2014; Acemoglu et al., 2016; Donald, 2023; Aghion et al., 2025). The carbon price is already doing the work of punishing pollution, and likewise, the threshold

[14] This setup does imply a form of market incompleteness in that workers would like to ex-ante insure themselves against the risks of preference shocks. See Donald et al. (2025a) for a discussion of the optimal policy response to this problem. Nonetheless, this market failure is distinct from the inability to find a buyer for one's labor at prevailing wages.

[15] I abstract from cross-technology knowledge spillovers to focus policy on the distributional question of augmenting labor or capital. Allowing for such spillovers would introduce an additional justification for intervening in the direction of innovation that is unrelated to the displacement of workers in tasks. For optimal policy with cross-technology spillovers, see Liu and Ma (2021) and Donald (2023).

rule is already addressing the distributional consequences of automation.[16]

This raises the question: does an optimal capital tax give the same result? In the absence of the threshold rule, perhaps capital taxation can also allow the Planner to avoid intervening in the direction of innovation. Surprisingly, the opposite is true. In fact, when taxing capital, the Planner pushes the direction of innovation towards labor-augmenting technology in proportion to the size of the optimal capital tax. This is because changing automation with technology is proportional to changing it with the capital supply, so the social desirability of pushing innovation towards labor increases with the optimal capital tax. In this case, the Planner wants to use innovation subsidies as another tool to address the displacement effect of automation.

Thus, allowing for endogenous innovation enriches the argument in favor of the threshold rule. By addressing the displacement effect of automation, the threshold rule allows prices to generate optimal innovation incentives as well, whereas the opposite is true for the capital tax.

4.6 Implementation of Threshold Rule with Fixed Costs

I have argued for the threshold rule and it's ability to achieve superior targeting relative to capital taxation. This mathematical fact raises a practical question: how would a policy authority implement the threshold rule? The key feature of the threshold rule is that it only introduces a wedge in the initial decision to automate; once a task is automated, the accumulation of capital should be left alone. This makes it more akin to a fixed cost than a tax,[17] and indeed, I show in Appendix B.6 that a policymaker can recreate the threshold rule with a tax (or subsidy) on fixed costs of automation.

To that end, I modify my baseline model so that automating a task requires the payment of a fixed cost. Allowing for such fixed costs is a common feature of task models that is consistent with firm-level evidence on adoption decisions (Hubmer and Restrepo, 2021; Acemoglu et al., 2024). By using policy to modify the fixed cost schedule, the Planner can shift the task allocation so that marginal tasks stay with labor. This is because firms are only willing to pay the fixed cost to automate a task when capital is productive enough to generate sufficient profit. By modifying fixed costs with a tax or subsidy, the Planner can select the degree to which capital must be productive enough to justify paying the fixed cost, allowing them to reproduce the threshold rule. The lesson here is that automation policy should focus on modifying fixed costs so as to prevent automation in marginal tasks, rather than uniformly taxing capital. In reality, policymakers will have difficulty perfectly discerning between fixed and variable costs,

[16] This is also analogous to the results of Lehr and Restrepo (2022), Costinot and Werning (2023), and Donald et al. (2025b) where optimally set second-best policy instruments imply that innovation can be valued with a first-best formula.

[17] The threshold rule could also be implemented as an choke price tax on tasks where capital is only slightly cheaper than labor, which by construction, no producer would pay in equilibrium. This makes the threshold rule more akin to a regulation in that it simply prohibits the use of capital in a marginal set of tasks without taxing any task where capital is still in use.

but the arguments of this paper provide an economic rationale for why such discernment is useful.

5 Calibration

In this section, I describe how I take my model to the data for the purpose of simulating optimal policy. My calibrated environment corresponds to the US economy in 2016. The main object of interest is occupational differences in exposure to automation, which I calibrate using the occupation-level exposure scores of Webb (2020).

5.1 Households

I start by describing the households of the model. Different household types correspond to the occupational classification scheme developed by Deming (2017). I compute occupation-level wages and labor supply using household-level data from the 2016 American Community Survey (ACS), the details of which are described in Appendix C.1. In one of my validation exercises below, I also make use of data from the 1980 Census.

For household preferences, consumption utility is CRRA:

$$u(c) = \frac{c^{1-\vartheta} - 1}{1 - \vartheta}, \tag{30}$$

where ϑ describes the curvature of marginal utility. I set ϑ to the standard value of one, so consumption utility is logarithmic. The disutility of labor takes an isoelastic form:

$$h_j(l) = \frac{\varphi_j}{1 + 1/\varepsilon} l^{1+1/\varepsilon}, \tag{31}$$

where φ_j is an occupation-specific scale term, and ε is the Frisch elasticity of labor supply. I set $\varepsilon = 0.75$ using the estimate from Chetty et al. (2011).

I will assume that social welfare is utilitarian, so we have

$$\mathcal{W}(\{n_j V_j\}) = \sum_j n_j V_j. \tag{32}$$

That is, the weight on the utility of each household type is just given by its population size.

To determine ownership shares of the initial endowment, I will assume that the initial endowment is split between an inherited capital stock and production. Production income is then split between capital and labor, with a share $S_k = 0.4$ going to capital. As I describe Appendix C.1, the total labor share is split amongst the occupations according to their share of total wage income in the ACS. I then need to assign wealth shares to each occupation to determine who owns the initial capital stock and its associated income. To do so, I assume a

mapping from labor shares S_j to wealth shares S_j^κ which follows

$$S_j^\kappa = \frac{S_j^\Theta}{\sum_i S_i^\Theta}, \tag{33}$$

where Θ determines the convexity of the wealth distribution, relative to the labor income distribution. For example, if $\Theta = 1$, the wealth distribution would be the same as the labor income distribution. But as Θ increases above one, we have that the ownership of wealth becomes more and more concentrated relative to the ownership of labor income.

To calibrate Θ, I use the 2016 Survey of Consumer Finances (SCF). The SCF allows me to capture the salient features of US wealth inequality as it purposely over-samples wealthy households in an attempt to adequately capture the fat right tail of the wealth distribution. As a confirmation of this, the share of wealth owned by the top decile in the SCF is similar to that obtained by Saez and Zucman (2016). I select Θ by first computing decile shares for both labor income and wealth, then minimizing the distance between predicted and observed wealth shares by decile. As I show in Figure D.1, $\Theta = 1.81$ gives a very close match between the model and the data. Unsurprisingly, this level of convexity implies highly concentrated ownership of wealth.

To describe household optimization in the status quo environment, I need a status quo tax function. I will assume that labor taxes take the form described in Heathcote et al. (2017) and that capital income taxes are linear. That is,

$$\mathcal{T}_j = w_j l_j - \Psi(w_j l_j)^{1-\psi} + \tau_k(r-\delta)\kappa_j - D, \tag{34}$$

where ψ determines the progressivity of labor taxes, and Ψ determines the overall level of taxation. Finally, I assume that all tax revenues are rebated back to household with a lump-sum D. I set $\psi = 0.181$ directly from the tax progressivity estimate of Heathcote et al. (2017). For capital taxes, Acemoglu et al. (2020) estimate an effective tax on capital income of $\tau_k = 0.1$ in 2016. Finally, I set $\Psi = 1.14$ to match a government revenue share of total income of 30%.

To increase the quantitative relevance of a two period model, I assume that households make a consumption/savings decision for a permanent steady-state, the details of which I describe in Appendix C.2. With income shares and the status quo tax function established, I set consumption in each period and the household discount factor using the household budget (20) and Euler equation (22), matching an assumed capital-output ratio in both the initial period and steady-state of $\mathcal{K} = 3.5$ and depreciation rate of $\delta = 0.07$. This approach yields a household discount factor of $\beta = 0.98$. Labor disutility φ_j then follows from the household labor supply condition (5).

5.2 Automation Exposure

I will now describe how I calibrate the production side of the economy. The task-level productivity for each occupation will be my main focus because this determines differences in exposure to automation. Appendix C.2 provides details on my calibration strategy, and Table 1 provides a summary of my calibrated parameters.

Table 1: Calibration Summary

Parameter	Description	Value	Source
ϑ	Marginal Utility Curvature	1	Standard
ε	Frisch Elasticity	0.75	Chetty et al. (2011)
β	Discount Factor	0.98	See Text
σ	Task-Level ES	0.5	Humlum (2021)
δ	Depreciation Rate	0.07	Standard
$\mathcal{K}$	Capital-Output Ratio	3.5	FRED
S_k	Capital Income Share	0.4	FRED
Θ	Wealth Share Convexity	1.81	SCF
Γ	Automation Exposure Effect	-0.008	Webb (2020)
Σ_k	Aggregate ES	1.25	Karabarbounis and Neiman (2014)

Following Acemoglu et al. (2020), I assume that task-level productivities take the following form:

$$\alpha_j(x) = x^{\zeta_j \nu_j} \tag{35}$$

$$\alpha_k(x) = x^{\zeta_j(\nu_j - 1)}, \tag{36}$$

where ν_j determines the absolute advantage and ζ_j determines the comparative advantage of labor in higher-ordered tasks, respectively. This implies that $x_j = \left[(w_j/A_j)/((1+\theta)r/A_k)\right]^{1/\zeta_j}$, so we have

$$d\ln(x_j) = \frac{1}{\zeta_j}\left(\left[d\ln(w_j) - d\ln(A_j)\right] - \left[d\ln((1+\theta)r) - d\ln(A_k)\right]\right). \tag{37}$$

That is, $1/\zeta_j$ is the elasticity of automation thresholds with respect to changes in relative marginal costs between capital and labor. If labor has high comparative advantage, increases(decreases) in the cost of labor(capital) will translate into only small advanced in the level of automation. But if labor has low comparative advantage, such changes will induce

large displacement of labor from their marginal tasks.

Therefore, variation in ζ_j is the main reason why different occupations have different exposure to automation. Put differently, cheaper capital will cause greater displacement for those occupations with low comparative advantage. To estimate ζ_j, I will make use of Webb (2020)'s measure of occupation-level exposure to software and robots, which quantifies exposure by measuring the textual similarity between O*NET task descriptions and the patents of a given technology. This corresponds closely to "routineness" measures derived from O*NET that are commonly used in the literature (Autor et al., 2003; Acemoglu and Autor, 2011), and Acemoglu and Restrepo (2022b) find that Webb (2020)'s measure is highly predictive of task displacement from 1980 to 2016. To map from exposure scores to comparative advantage, I assume that $\zeta_j = \exp(\Gamma\chi_j)$, where χ_j is the average exposure percentile across robots and software. Therefore, Γ is the semi-elasticity of comparative advantage with respect to exposure score, which is hypothesized to be negative.

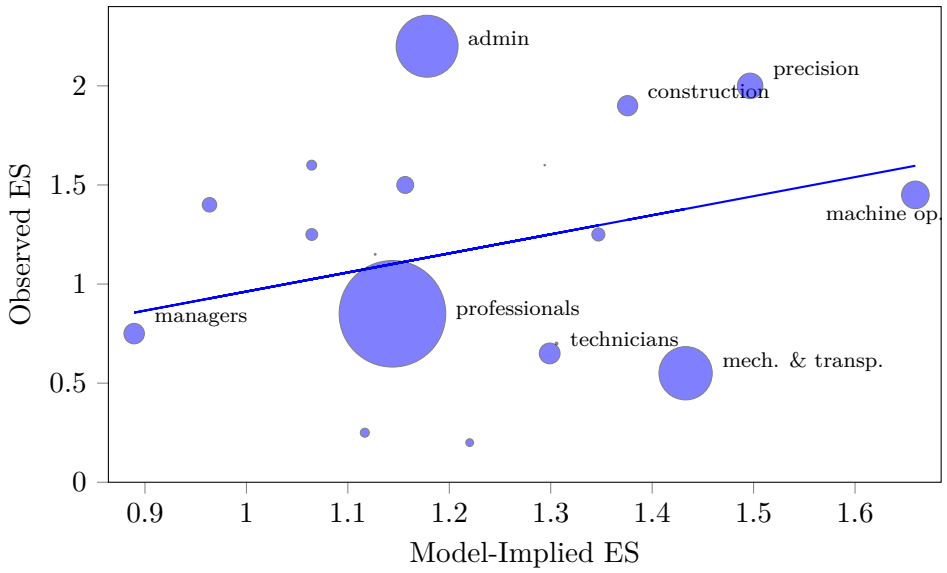
I calibrate Γ by matching an aggregate elasticity of substitution of $\Sigma_k = 1.25$. This is the estimate of Karabarbounis and Neiman (2014), who look at long-run variation in the user cost of capital. As I described in Section 2.3, the aggregate elasticity of substitution Σ_k is a weighted average of occupation-level elasticities of substitution Σ_j, which are in turn determined by the sum of cross-task and within-task substitution. I set the cross-task elasticity of substitution using Humlum (2021)'s estimate of $\sigma = 0.5$, which is also in line with the plant-level elasticity estimates of Oberfield and Raval (2021). Intuitively, I am assuming that the task allocation is fixed in the short run, so short-run variation measures the cross-task elasticity of substitution. However, long-run variation picks up both sources of substitution, so the difference between the two provides information on the degree to which price changes generate displacement from marginal tasks. This allows me to back out an estimate of Γ under the assumption that occupation-level differences in comparative advantage are summarized by Webb (2020)'s exposure measure.

The resulting estimate of this automation exposure effect is $\Gamma = -0.008$, implying that an occupation at the 50th percentile of exposure would have an automation threshold elasticity of $1/\exp(\Gamma \cdot 50) \approx 1.46$. For the remaining parameters, I start by normalizing $\overline{x} = 100$. I then pick labor productivity and absolute advantage $\{A_j, \nu_j\}$ to match labor incomes shares S_j when automation thresholds are set at 20% of the total task interval: $x_j = \varkappa\overline{x}$. This final choice is not without loss, as it determines the absolute degree to which losing tasks lowers wages. To test the empirical reasonableness of this assumption, I compare the model's implied average elasticity of log wages with respect to task displacement with that estimated by Acemoglu and Restrepo (2022b). My choice of $\varkappa = 0.2$ implies an average elasticity of -3.02, which is within their range of estimated values. Finally, I select capital productivity A_k to normalize output to 100 while the capital-output ratio is $\mathcal{K} = 3.5$.

I validate my calibration strategy with two exercises. First, following Acemoglu and Re-

strepo (2022b), I measure how much of the variation across occupations in changes in log wages from 1980 to 2016 can be explained by the task displacement implied by my model. Examining task displacement driven by reductions in the user cost of capital, I find that my model can explain 12.6% of the variation in wage growth across this period. I focus on explaining variation (as the R-squared of a regression) as it focuses attention on the unequal wage growth workers experienced over this period. In a simple, balanced-growth benchmark, all occupations would have experienced the same gains in log wages, so given that this is not what we observed, I am asking how much of the unequal gains over this period can be explained by the differential automation exposure of my model. Acemoglu and Restrepo (2022b) conduct a similar exercise with demographic groups and find that task displacement can explain a whopping 60% of the variation in log wage growth over this period. Therefore, my calibration is conservative in how much inequality it attributes to differential exposure to automation.

Figure 2: Observed vs. Model-Implied Occupational Elasticities of Substitution



Notes: Occupation-level elasticities of substitution come from Caunedo et al. (2023) using their 2-digit classification. Aggregation and regression weights, as well as dot scale, all make use of the weights described in Equation (19).

For my second validation exercise, I examine my calibrated model's ability to match the occupational-level elasticity of substitution estimates of Caunedo et al. (2023). As I have discussed, variation in elasticities of substitution across occupations is driven in my model by variation in comparative advantage, so this exercise assesses the ability of my calibration approach to back out reasonable elasticities of substitution not just in the aggregate, but also at the occupation level. Caunedo et al. (2023) make use of a more coarse 2-digit occupation scheme, so I aggregate my occupational elasticities up to this level using the weights described in Equation (19). I then regress their estimated elasticities on my model-implied elasticities

(also using the weights of Equation (19)) and find a coefficient of 0.96.[18] The results of this exercise are plotted in Figure 2. Thus, my model does a reasonable job of recreating occupation-level variation in the elasticity of substitution.

6 Numerical Results

In this section, I assess the quantitative importance of the threshold rule, considering its introduction in the context of both a fully optimal non-linear tax schedule as well as the US status quo tax system. I also simulate various scenarios for the introduction of AI and show that the optimal threshold rule grows alongside the productivity of capital.

6.1 Optimal Policy

First, I will consider introducing the threshold rule in the context of a optimal non-linear tax system. As shown in Proposition 3, the non-linear tax system will generally tax capital when the threshold rule is not available, and indeed, I find that a tax on capital income of 10.9% is optimal in the absence of a threshold rule.[19] Table 2 shows the change brought about by moving to the allocation with an optimal threshold rule of 10.8%. By eliminating the capital tax, the Planner achieves a slight increase in output and a consumption equivalent welfare gain of 0.05%.

Table 2: Optimal Threshold Rule

	Non-Linear Tax Function	**Status Quo Tax Function**
θ	10.8%	31.8%
$\%\Delta Y$	0.02%	-3.3%
$\%CE$	0.05%	1.4%
$\%\Delta Cov$	-5.11%	-16%

Contrast this with the change in the status quo allocation by introducing the threshold rule alone. In the absence of optimal redistribution through other policy tools, the optimal threshold rule roughly triples to 31.8%. This distortion in the task allocation generates a small reduction in output of -3.3%, but in doing so, it generates a substantial consumption equivalent welfare gain of 1.4%. The difference between the two policy experiments comes from the fact that the equity value of limiting automation depends on the level of inequality in the remainder of the economy. In particular, the non-linear tax problem allows the Planner to

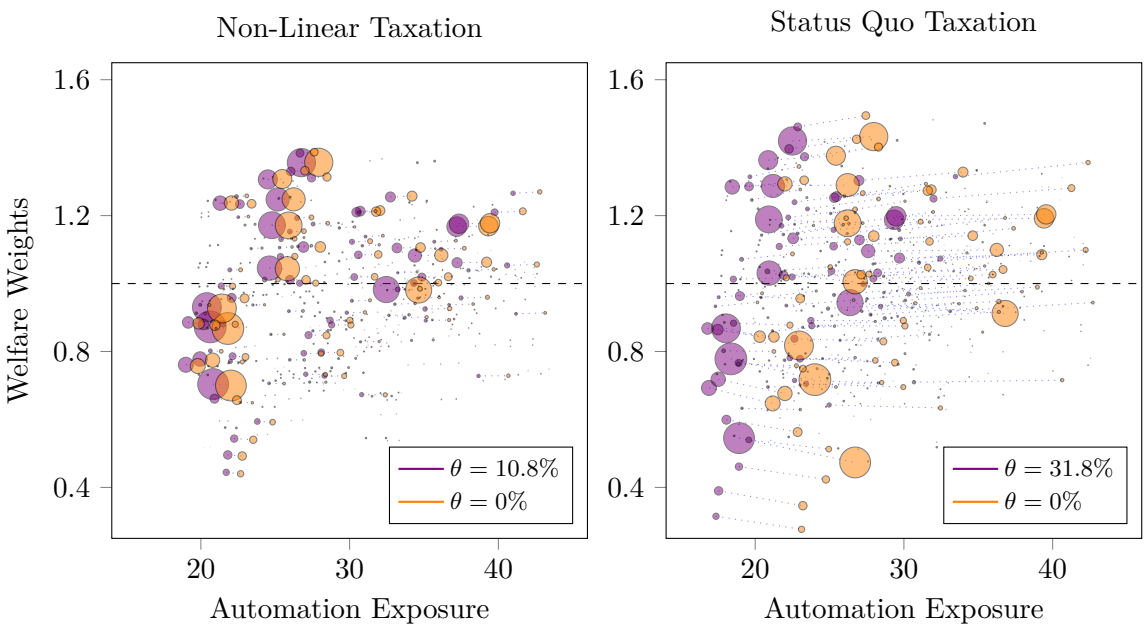
[18] This regression does not include an intercept, but when I include an intercept I get an estimated intercept and coefficient of 0.81 and 0.32, respectively

[19] I list the savings wedge as a tax on capital income $r-\delta$ for the sake of comparison with existing tax systems, but this is equivalent to a wealth tax of 0.5%.

fully tax and redistribute the initial endowment, making the ownership of capital following the first period much more equal.[20] Therefore, allowing capital to displace labor is less harmful from an equity perspective because the workers being displaced own a larger share of the capital displacing them. In contrast, the initial endowment in the status quo allocation is only subject to standard taxation, so the ownership of capital is highly concentrated. This makes a reduction in wages for workers translate into a similar-sized reduction in consumption.

The final row of Table 2 considers a measure of the change in automation-based inequality: the covariance between automation exposure and welfare weights, where consistent with Equation (13), automation exposure is given by the inverse of comparative advantage x_j/ζ_j. In both cases, the threshold rule exploits this variation to enhance equity and reduces the covariance in the process. These reductions in covariance can seen in Figure 3, where I plot welfare weights against automation exposure in the two tax settings. The orange and purple dots represent the allocations before and after the introduction of the threshold rule, respectively.

Figure 3: Equity Consequences of Automation



Notes: This graph shows the changing relationship between automation exposure and welfare weights when introducing the threshold rule for the two tax systems. The dotted lines connect an occupation's two dots across the allocations. The scale of the dots is given by the size of each occupation n. The dashed horizontal line benchmarks the average welfare weight of one.

In both cases, introducing the threshold rule reduces automation exposure by pushing down the level of automation and compresses dispersion in welfare weights by pushing them

[20]Taxing the initial endowment is generally beneficial for the Planner because it is a lump-sum tax on a sunk source of income, but it does generate the standard time inconsistency problem discussed at length in the capital taxation literature (Kydland and Prescott, 1977; Chari and Kehoe, 1990; Klein et al., 2008; Farhi et al., 2012).

towards one. The greater equity benefit of the threshold rule in the context of the status quo tax function is apparent here as well. Welfare weights are already less dispersed in the non-linear taxation case because free choice of the income tax gives the Planner much more scope for redistribution before using the threshold rule. This background redistribution justifies a smaller threshold rule, resulting in much less aggressive reductions in automation exposure in the non-linear tax case than in the status quo. Therefore, we can see why the threshold rule reduces the covariance between welfare weights and automation by less when the income tax is already optimized: there is both less variation to exploit and less need to exploit it.

6.2 Introduction of AI

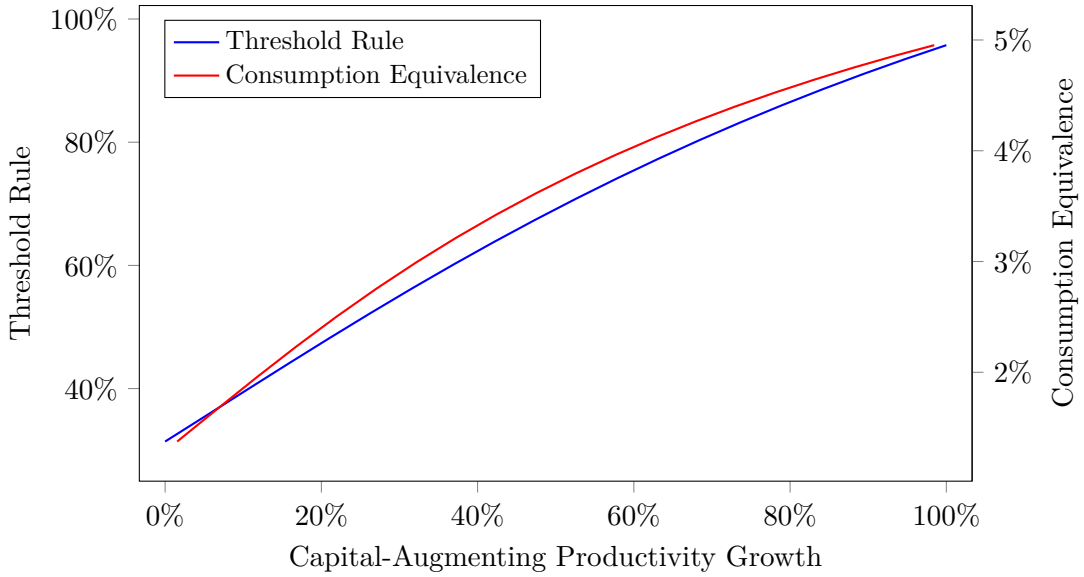
In this section, I simulate the introduction of AI and consider the optimal threshold rule when the tax function is unchanged.[21] This analysis is necessarily somewhat speculative, as the adoption of AI remains in its early stages (Acemoglu et al., 2024). In this context, empirically grounded projections – while imperfect – offer the most disciplined available approach to assessing likely impacts (Frey and Osborne, 2017; Eloundou et al., 2024). Of particular interest is the way in which occupational exposure to automation from AI way differ from past episodes of automation. To discipline AI's impact on different occupations, I make use of Webb (2020)'s exposure scores for AI technologies. Methodologically, these are measured in the same way as the exposure scores for robots and software used in the previous section, but as I show in Figure D.1, AI exposure is *positively* correlated with wages. That is, unlike previous automating technologies, which tended to displace workers with lower wages, AI has the potential to displace high wage workers by producing in cognitive tasks performed by white collar workers.

To map these AI exposure scores to occupation-level comparative advantage, I use the calibrated mapping described in Section 5 and input average exposure score across the three technologies: robots, software, and AI. I then adjust the absolute advantage and productivity parameters so that the observed equilibrium outcomes remain unchanged. I describe the details of this procedure in Appendix C.2. With occupation-level exposure to AI in hand, I simulate the introduction of AI by growing capital-augmenting productivity A_k. The growth consequences of AI are a major focus of the literature on AI's macro impacts (Aghion et al., 2019; Nordhaus, 2021; Trammell and Korinek, 2023), so to consider a variety of possible outcomes, Figure 4 shows the optimal threshold rule as a function of the growth in capital-augmenting productivity.

As one can see, automation policy should become sharply more aggressive the more we expect AI to enhance the productivity of capital. Interestingly, only changing occupation-level comparative advantage by introducing AI does little to change the optimal threshold rule. Instead, growth in the productivity of capital has a significant quantitative effect on optimal

[21] As I discuss in Appendix C.2, I do allow the tax scale parameter Ψ to change so that marginal tax rates do not mechanically grow with income.

Figure 4: Introduction of AI



Notes: This graph shows the optimal threshold rule, as well as the consumption equivalent welfare gain from adopting that threshold rule, as a function of the growth in capital-augmenting productivity. The tax system in the background is the US status quo tax function, and labor-augmenting productivity is assumed to grow at 25% the rate of capital.

policy. As capital's productivity grows to the point of doubling, the optimal threshold rule more than triples to 95.8%. Moreover, the welfare stakes of introducing optimal policy grow in tandem. The red line of Figure 4 shows the consumption equivalent welfare gain of introducing the optimal threshold rule for a given level of capital productivity growth, and this value rises to 5% as the productivity of capital doubles. Therefore, optimal policy becomes both more aggressive and more consequential as AI makes capital more productive.

Finally, it seems unlikely that AI will *only* enhance the productivity of capital, so I also consider various degrees to which labor-augmenting technology is able to rise alongside that of capital. Figure 4 assumes that all labor-augmenting technologies experience 25% of the growth achieved by capital, but in Figure D.2, I also consider 10% and 50% catchup growth for labor. Interestingly, the assumed fraction with which labor's productivity can keep up with capital has little quantitative impact on optimal policy. In fact, optimal policy is slightly *more* aggressive when labor's productivity grows at 50% of the rate for capital. This is because workers need to produce in tasks to take advantage of their enhanced productivity, so those most exposed to automation are least able to benefit from rising labor productivity. For this reason, the variance of wages is highest in the 50% case, justifying a larger threshold rule to pull back automation and share the benefit of technological change more widely. However, the consumption equivalent welfare gains from introducing the optimal threshold rule are largest for the 10% case simply because lower economic growth means greater variance in welfare weights,

increasing the welfare consequences of redistributing income. Therefore, an expectation that labor-augmenting technology can keep up with that of capital does not mean the need for policy is relaxed.

7 Conclusion

Automation changes the value of different groups' labor endowments, and a normative evaluator may find the resulting inequality unacceptable. In this paper, I have argued that the right policy response to this problem needs to consider that capital only harms labor at the extensive margin of automation. Once a task is automated, further capital deepening raises demand for labor and therefore wages. This implies that, if the Planner can introduce a wedge in the automation decision via the threshold rule, automation concerns alone do not justify capital taxation.

The dominance of the threshold rule over capital taxation is robust to a variety of extensions including intrinsic value from work, occupational mobility, and endogenous innovation. While perfect implementation of the threshold rule presents practical challenges, I have shown that it can be accomplished via a tax or subsidy on fixed costs for automation. Quantitatively, the importance of implementing the threshold rule depends on the background level of inequality in society, especially on the degree to which capital ownership is concentrated. The stakes for getting policy right are further amplified by the possibility that new AI tools could radically increase the productivity of capital. Future work could consider the role of the threshold role in a long-run transition to an economy where all tasks can be produced by capital (Ray and Mookherjee, 2022; Restrepo, 2025).

I have focused on the process of automation, but the distinction between the extensive and intensive margin of economic displacement applies more generally. For example, similar reasoning could apply to trade; another topic where policy's role in alleviating disruption is highly debated (Grossman and Helpman, 1994; Autor et al., 2013). In that setting, the arguments of this paper would imply that, like capital taxation, tariffs are an excessively blunt instrument, with gains coming from more precise targeting of the extensive margin of displacement from import competition.

Finally, this paper's focus on the policy implications of distinguishing between the extensive and intensive margin of automation could have further implications for growth. The initial switch to a new technology may bring about small productivity gains and large disruptions precisely because the first switch happens when the new technology has only just surpassed the incumbent. Therefore, the eventual benefit of new automating technologies such as AI will depend critically on the degree to which further productivity improvements and capital deepening occur after the initial switch (Acemoglu, 2025; Jones and Tonetti, 2025). How the threshold rule may interfere with the process, e.g. by preventing learning-by-doing, is a

question that warrants further study.

References

Acemoglu, D. (2025). The Simple Macroeconomics of AI. *Economic Policy*, 40(121):13–58.

Acemoglu, D., Akcigit, U., Hanley, D., and Kerr, W. (2016). Transition to Clean Technology. *Journal of Political Economy*, 124(1):52–104.

Acemoglu, D., Anderson, G., Beede, D., Buffington, C., Childress, E., Dinlersoz, E., Foster, L., Goldschlag, N., Haltiwanger, J., Kroff, Z., Restrepo, P., and Zolas, N. (2024). Automation and the Workforce: A Firm-Level View from the 2019 Annual Business Survey. In *Technology, Productivity, and Economic Growth*. University of Chicago Press.

Acemoglu, D. and Autor, D. (2011). Skills, Tasks, and Technologies: Implications for Employment and Earnings. In *Handbook of Labor Economics*, volume 4, pages 1043–1171. Elsevier.

Acemoglu, D. and Lensman, T. (2024). Regulating Transformative Technologies. *American Economic Review: Insights*, 6(3):359–76.

Acemoglu, D., Manera, A., and Restrepo, P. (2020). Does the US Tax Code Favor Automation? *Brookings Papers on Economic Activity*.

Acemoglu, D. and Restrepo, P. (2018). The Race Between Man and Machine: Implications of Technology for Growth, Factor Shares, and Employment. *American Economic Review*, 108(6):1488–1542.

Acemoglu, D. and Restrepo, P. (2020). Robots and Jobs: Evidence from US Labor Markets. *Journal of Political Economy*, 128(6):2188–2244.

Acemoglu, D. and Restrepo, P. (2022a). Demographics and Automation. *The Review of Economic Studies*, 89(1):1–44.

Acemoglu, D. and Restrepo, P. (2022b). Tasks, Automation, and the Rise in U.S. Wage Inequality. *Econometrica*, 90(5):1973–2016.

Aghion, P., Antonin, C., Bunel, S., and Jaravel, X. (2022). Modern Manufacturing Capital, Labor Demand, and Product Market Dynamics: Evidence from France. *Working Paper*.

Aghion, P., Barrage, L., Donald, E., Hémous, D., and Liu, E. (2025). Transition to Green Technology along the Supply Chain. Working Paper 33934, National Bureau of Economic Research.

Aghion, P., Jones, B., and Jones, C. (2019). Artificial Intelligence and Economic Growth. In *The Economics of Artificial Intelligence: An Agenda*, pages 237–290. University of Chicago Press.

Ales, L., Kurnaz, M., and Sleet, C. (2015). Technical Change, Wage Inequality, and Taxes. *American Economic Review*, 105(10).

Atkinson, A. and Stiglitz, J. (1976). The Design of Tax Structure: Direct Versus Indirect Taxation. *Journal of Public Economics*, 6(1-2):55–75.

Atolia, M., Holland, M., and Kreamer, J. (2025). Optimal Taxes and Basic Income during an Episode of Sutomation: A Worker's Perspective. *Macroeconomic Dynamics*, 29:e28.

Autor, D., Chin, C., Salomons, A., and Seegmiller, B. (2024). New Frontiers: The Origins and Content of New Work, 1940–2018. *The Quarterly Journal of Economics*, 139(3):1399–1465.

Autor, D., Dorn, D., and Hanson, G. (2013). The China Syndrome: Local Labor Market Effects of Import Competition in the United States. *American Economic Review*, 103(6):2121–68.

Autor, D., Levy, F., and Murnane, R. (2003). The Skill Content of Recent Technological Change: An Empirical Exploration. *The Quarterly Journal of Economics*, 118(4):1279–1333.

Beraja, M. and Zorzi, N. (2024). Inefficient Automation. *The Review of Economic Studies*, 92(1):69–96.

Bond, P. and Kremens, L. (2024). Income and Inequality under Asymptotically Full Automation. *Working Paper*.

Burstein, A., Morales, E., and Vogel, J. (2019). Changes in Between-Group Inequality: Computers, Occupations, and International Trade. *American Economic Journal: Macroeconomics*, 11(2):348–400.

Caliendo, L., Dvorkin, M., and Parro, F. (2019). Trade and Labor Market Dynamics: General Equilibrium Analysis of the China Trade Shock. *Econometrica*, 87(3):741–835.

Caunedo, J., Jaume, D., and Keller, E. (2023). Occupational Exposure to Capital-Embodied Technical Change. *American Economic Review*, 113(6):1642–1685.

Chamley, C. (1986). Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives. *Econometrica*, 54(3):607–622.

Chari, V. and Kehoe, P. (1990). Sustainable Plans. *Journal of Political Economy*, 98(4):783–802.

Chetty, R., Guren, A., Manoli, D., and Weber, A. (2011). Are Micro and Macro Labor Supply Elasticities Consistent? A Review of Evidence on the Intensive and Extensive Margins. *American Economic Review: Papers and Proceedings*, 101(3):471–475.

Costinot, A. and Werning, I. (2023). Robots, Trade, and Luddism: A Sufficient Statistic Approach to Optimal Technology Regulation. *The Review of Economic Studies*, 90(5):2261–2291.

de Souza, G. and Li, H. (2023). Robots, Tools, and Jobs: Evidence from Brazilian Labor Markets. Working Paper Series WP 2023-42, Federal Reserve Bank of Chicago.

Deming, D. (2017). The Growing Importance of Social Skills in the Labor Market. *The Quarterly Journal of Economics*, 132(4):1593–1640.

Diamond, P. and Mirrlees, J. (1971a). Optimal Taxation and Public Production I: Production Efficiency. *The American Economic Review*, 61(1):8–27.

Diamond, P. and Mirrlees, J. (1971b). Optimal Taxation and Public Production II: Tax Rules. *The American Economic Review*, 61(3):261–278.

Domar, E. (1961). On the Measurement of Technological Change. *The Economic Journal*, 71(284):709–729.

Donald, E. (2023). Spillovers and the Direction of Innovation: An Application to the Clean Energy Transition. *Working Paper*.

Donald, E., Fukui, M., and Miyauchi, Y. (2025a). Optimal Dynamic Spatial Policy. Working Paper 34290, National Bureau of Economic Research.

Donald, E., Fukui, M., and Miyauchi, Y. (2025b). Unpacking Aggregate Welfare in a Spatial Economy. Working Paper 34075, National Bureau of Economic Research.

Eloundou, T., Manning, S., Mishkin, P., and Rock, D. (2024). GPTs are GPTs: Labor Market Impact Potential of LLMs. *Science*, 384(6702):1306–1308.

Farhi, E., Sleet, C., Werning, I., and Yeltekin, S. (2012). Non-Linear Capital Taxation Without Commitment. *Review of Economic Studies*, 79(4):1469–1493.

Frey, C. B. and Osborne, M. (2017). The Future of Employment: How Susceptible are Jobs to Computerisation? *Technological Forecasting and Social Change*, 114:254–280.

Golosov, M., Hassler, J., Krusell, P., and Tsyvinski, A. (2014). Optimal Taxes on Fossil Fuel in General Equilibrium. *Econometrica*, 82(1):41–88.

Golosov, M., Kocherlakota, N., and Tsyvinski, A. (2003). Optimal Indirect and Capital Taxation. *The Review of Economic Studies*, 70(3):569–587.

Grossman, G. and Helpman, E. (1994). Protection for Sale. *American Economic Review*, 84(4):833–850.

Guerreiro, J., Rebelo, S., and Teles, P. (2022). Should Robots be Taxed? *The Review of Economic Studies*, 89(1):279–311.

Guerreiro, J., Rebelo, S., and Teles, P. (2023). Regulating Artificial Intelligence. Working Paper 31921, National Bureau of Economic Research.

Heathcote, J., Storesletten, K., and Violante, G. (2017). Optimal Tax Progressivity: An Analytical Framework. *The Quarterly Journal of Economics*, 132(4):1693–1754.

Hémous, D. and Olsen, M. (2022). The Rise of the Machines: Automation, Horizontal Innovation, and Income Inequality. *American Economic Journal: Macroeconomics*, 14(1):179–223.

Hémous, D., Olsen, M., Zanella, C., and Dechezleprêtre, A. (2025). Induced Automation Innovation: Evidence from Firm-Level Patent Data. *Journal of Political Economy*, 133(6):1975–2028.

Hosseini, R. and Shourideh, A. (2020). Inequality, Redistribution, and Optimal Trade Policy: A Public Finance Approach. *Working Paper*.

Hubmer, J. and Restrepo, P. (2021). Not a Typical Firm: The Joint Dynamics of Firms, Labor Shares, and Capital–Labor Substitution. Working Paper 28579, National Bureau of Economic Research.

Hulten, C. (1978). Growth Accounting with Intermediate Inputs. *The Review of Economic Studies*, 45(3):511–518.

Humlum, A. (2021). Robot Adoption and Labor Market Dynamics. *Working Paper*.

Jones, C. (2024). The AI Dilemma: Growth versus Existential Risk. *American Economic Review: Insights*, 6(4):575–90.

Jones, C. (2025). How Much Should We Spend to Reduce A.I.'s Existential Risk? Working Paper 33602, National Bureau of Economic Research.

Jones, C. and Tonetti, C. (2025). Past Automation and Future A.I.: How Weak Links Tame the Growth Explosion. *Working Paper*.

Kaplow, L. (2006). On the Undesirability of Commodity Taxation even when Income Taxation is not Optimal. *Journal of Public Economics*, 90(6-7):1235–1250.

Karabarbounis, L. and Neiman, B. (2014). The Global Decline of the Labor Share. *The Quarterly Journal of Economics*, 129(1):61–103.

Klein, P., Krusell, P., and Rios-Rull, J.-V. (2008). Time-Consistent Public Policy. *The Review of Economic Studies*, 75(3):789–808.

Krusell, P., Ohanian, L., Ríos-Rull, J.-V., and Violante, G. (2000). Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis. *Econometrica*, 68(5):1029–1053.

Kurnaz, M., Michelini, M., Özdenören, H., and Sleet, C. (2025). Tax Design for the Long Run. *Working Paper*.

Kydland, F. and Prescott, E. (1977). Rules Rather than Discretion: The Inconsistency of Optimal Plans. *Journal of Political Economy*, 85(3):473–491.

Laroque, G. (2005). Indirect Taxation is Superfluous under Separability and Taste Homogeneity: A Simple Proof. *Economics Letters*, 87(1):141–144.

Lehr, N. and Restrepo, P. (2022). Optimal Gradualism. Working Paper 30755, National Bureau of Economic Research.

Lehr, N. and Restrepo, P. (2025). The Price of Intelligence: How Should Socially-Minded Firms Price and Deploy AI? Working Paper 34424, National Bureau of Economic Research.

Liu, E. and Ma, S. (2021). Innovation Networks and R&D Allocation. Working Paper 29607, National Bureau of Economic Research.

Mirrlees, J. (1971). An Exploration in the Theory of Optimum Income Taxation. *The Review of Economic Studies*, 38(2):175–208.

Moll, B., Rachel, L., and Restrepo, P. (2022). Uneven Growth: Automation's Impact on Income and Wealth Inequality. *Econometrica*, 90(6):2645–2683.

Naito, H. (1999). Re-Examination of Uniform Commodity Taxes under a Non-Linear Income Tax System and its Implication for Production Efficiency. *Journal of Public Economics*, 71(2):165–188.

Nordhaus, W. (2021). Are we approaching an Economic Singularity? Information Technology and the Future of Economic Growth. *American Economic Journal: Macroeconomics*, 13(1):299–332.

Oberfield, E. and Raval, D. (2021). Micro Data and Macro Technology. *Econometrica*, 89(2):703–732.

Ray, D. and Mookherjee, D. (2022). Growth, Automation, and the Long-Run Share of Labor. *Review of Economic Dynamics*, 46:1–26.

Restrepo, P. (2023). Automation: Theory, Evidence, and Outlook. *Annual Review of Economics*, 16.

Restrepo, P. (2025). We Won't be Missed: Work and Growth in the AGI World. Working Paper 34423, National Bureau of Economic Research.

Saez, E. (2001). Using Elasticities to Derive Optimal Income Tax Rates. *The Review of Economic Studies*, 68(1):205–229.

Saez, E. and Stantcheva, S. (2018). A Simpler Theory of Optimal Capital Taxation. *Journal of Public Economics*, 162:120–142.

Saez, E. and Zucman, G. (2016). Wealth Inequality in the United States since 1913: Evidence from Capitalized Income Tax Data. *The Quarterly Journal of Economics*, 131(2):519–578.

Schulz, K., Tsyvinski, A., and Werquin, N. (2023). Generalized Compensation Principle. *Theoretical Economics*, 18(4):1665–1710.

Slavík, C. and Yazici, H. (2014). Machines, Buildings, and Optimal Dynamic Taxes. *Journal of Monetary Economics*, 66:47–61.

Straub, L. and Werning, I. (2020). Positive Long-Run Capital Taxation: Chamley-Judd Revisited. *American Economic Review*, 110(1):86–119.

Thuemmel, U. (2023). Optimal Taxation of Robots. *Journal of the European Economic Association*, 21(3):1154–1190.

Trammell, P. and Korinek, A. (2023). Economic Growth under Transformative AI. Working Paper 31815, National Bureau of Economic Research.

Webb, M. (2020). The Impact of Artificial Intelligence on the Labor Market. *Working Paper*.

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

Online Appendix

Optimal Taxation with Automation:
Navigating Capital and Labor's Complicated Relationship

Eric Donald

Contents

A Theory Derivations & Proofs **1**
A.1 Derivation of Aggregate Production Function . . . 1
A.2 Aggregate Elasticity of Substitution . . . 2
A.3 Proof of Lemma 1 . . . 3
A.4 Proof of Proposition 1 . . . 4
A.5 Proof of Proposition 2 . . . 5
A.6 Proof of Corollary 1 . . . 6
A.7 Proof of Proposition 3 . . . 8
A.8 Cobb-Douglas Production . . . 9

B Extensions **9**
B.1 Infinite Horizon . . . 9
B.2 Within-Type Heterogeneity . . . 14
B.3 Intrinsic Value from Work . . . 15
B.4 Type Switching . . . 16
B.5 Endogenous Innovation . . . 18
B.6 Implementation of Threshold Rule with Fixed Costs . . . 22

C Details on Calibration & Simulation **23**
C.1 Description of Data Sources . . . 23
C.2 Calibration Moments . . . 24
C.3 Computing Threshold Rule Perturbations . . . 27
C.4 Non-Linear Tax Solution Algorithm . . . 28

D Additional Tables & Figures **29**

A Theory Derivations & Proofs

A.1 Derivation of Aggregate Production Function

The representative firm solves the profit maximization problem described in Equation (9). First, because production is perfect substitutes within tasks, the factor with the lowest marginal

cost (conditional on the threshold rule) will produce in a given task, as described by the task assignment sets of Equations (10) and (11).

The firm's optimality conditions follow

$$\Big(\frac{Y}{A_j\alpha_j(x)\ell_j(x)}\Big)^{\frac{1}{\sigma}} = \frac{w_j}{A_j\alpha_j(x)}$$
$$\Big(\frac{Y}{A_k\alpha_k(x)k(x)}\Big)^{\frac{1}{\sigma}} = \frac{r}{A_k\alpha_k(x)}.$$

Aggregating over the total supply of labor, we can derive a formula for wages:

$$L_j = \int_{\Omega_j} \frac{Y}{w_j^\sigma} A_j^{\sigma-1}\alpha_j(x)^{\sigma-1}dx$$
$$\Rightarrow w_j = \Big(\int_{\Omega_j} \alpha_j(x)^{\sigma-1}dx\Big)^{\frac{1}{\sigma}} A_j^{\frac{\sigma-1}{\sigma}} \Big(\frac{Y}{L_j}\Big)^{\frac{1}{\sigma}},$$

and analogously for the rental rate. Plugging these factor prices into the ideal price index gives us Equation (14).

A.2 Aggregate Elasticity of Substitution

This appendix derives the mapping from labor-type-level elasticities of substitution to the aggregate elasticity of substitution, as described in Equation (19). In what follows, I will only assume that the production function is constant returns. To start, we have that $S_j/S_k = (L_j/K)/(r/w_j)$. This implies that

$$\frac{\partial \ln(S_j)}{\partial \ln(r)} - \frac{\partial \ln(S_k)}{\partial \ln(r)} = (\Sigma_j - 1)\Big(1 - \frac{\partial \ln(w_j)}{\partial \ln(r)}\Big). \tag{A.1}$$

If we now consider the marginal cost function $\mathcal{C}(r, \{w_j\}) = 1$, we have

$$\sum_j S_j \frac{\partial \ln(w_j)}{\partial \ln(r)} + S_k = 0. \tag{A.2}$$

Finally, the fact that income shares must sum to one implies

$$\sum_j S_j \frac{\partial \ln(S_j)}{\partial \ln(r)} + S_k \frac{\partial \ln(S_k)}{\partial \ln(r)} = 0. \tag{A.3}$$

Combining, we have

$$\frac{\partial \ln(S_k)}{\partial \ln(r)} = 1 - \Sigma_k = 1 - \sum_j S_j\Big(1 - \frac{\partial \ln(w_j)}{\partial \ln(r)}\Big)\Sigma_j,$$

which straightforwardly implies Equation (19). Note that Equation (A.2) implies that the weights on the final sum must add up to one.

A.3 Proof of Lemma 1

The automation thresholds X are the tasks at which the marginal cost of producing with labor equals the adjusted marginal cost of producing with capital, where the adjustment comes from the threshold rule θ. These threshold conditions, given by Equation (12), allow us to define implicit functions for the automation thresholds with respect to (L, K, θ).

Formally, if we define the function $G_j = \frac{r(1+\theta)}{A_k \alpha_k(x_j)} - \frac{w_j}{A_j \alpha_j(x_j)}$, then the Implicit Function Theorem implies that

$$\left(\frac{\partial X}{\partial(L,K,\theta)}\right) = -\left(\frac{\partial G}{\partial X}\right)^{-1}\left(\frac{\partial G}{\partial(L,K,\theta)}\right),$$

evaluated along the level set where $G = 0$.

Denote by $\tilde{P}_j \equiv \frac{w_j}{A_j \alpha_j(x_j)}$ the marginal cost of producing with labor at x_j. We have that

$$\frac{\partial G_j}{\partial K} = -\frac{\tilde{P}_j}{\sigma K} \tag{A.4}$$

$$\frac{\partial G_j}{\partial \theta} = \frac{\tilde{P}_j}{1+\theta}. \tag{A.5}$$

Thus, $\frac{\partial G_j}{\partial K} \propto \frac{\partial G_j}{\partial \theta}$, but because these two columns will be multiplied by the same matrix, we also have $\frac{\partial x_j}{\partial K} \propto \frac{\partial x_j}{\partial \theta}$, as desired. Furthermore, it will be useful to derive $\frac{\partial X}{\partial(K,\theta)}$ in closed form. First, we have that

$$\frac{\partial G_j}{\partial x_j} = \tilde{P}_j\Big(\frac{1}{\sigma}\Big[\frac{\alpha_k(x_j)^{\sigma-1}}{\Lambda_k} + \frac{\alpha_j(x_j)^{\sigma-1}}{\Lambda_j}\Big] + \frac{\partial \ln\left(\alpha_j(x_j)/\alpha_k(x_j)\right)}{\partial x_j}\Big),$$

as well as

$$\frac{\partial G_i}{\partial x_j}|_{i\neq j} = \frac{\tilde{P}_i}{\sigma}\frac{\alpha_k(x_j)^{\sigma-1}}{\Lambda_k}.$$

Therefore, we can write $\frac{\partial G}{\partial X}$ as the sum of a diagonal matrix with diagonal components $\tilde{P}_j(\frac{1}{\sigma}\frac{\alpha_j(x_j)^{\sigma-1}}{\Lambda_j} + \frac{\partial \ln(\alpha_j(x_j)/\alpha_k(x_j))}{\partial x_j})$ plus the outer product of a column vector given by $\frac{\tilde{P}_i}{\sigma}$ and a row vector given by $\frac{\alpha_k(x_j)^{\sigma-1}}{\Lambda_k}$. That is, we have a matrix of the form $Z + uv'$, which means, to find its inverse, we can apply the Sherman-Morrison Formula where

$$\left(Z + uv'\right)^{-1} = Z^{-1} - \frac{Z^{-1}uv'Z^{-1}}{1 + v'Z^{-1}u}.$$

Luckily, because Z is diagonal, we have that Z^{-1} is just the multiplicative inverse along the

diagonal. Therefore, we have

$$Z^{-1}uv'Z^{-1} = \Big(\frac{u_i v_j}{Z_i Z_j}\Big)_{ij}$$
$$1 + v'Z^{-1}u = 1 + \sum_j \frac{u_j v_j}{Z_j}.$$

Then, if we define $\gamma \equiv 1/(1+\sum_j \frac{u_j v_j}{Z_j})$ and $\hat{\alpha}_j \equiv \frac{\alpha_j(x_j)^{\sigma-1}}{\Lambda_j} + \sigma \frac{\partial \ln(\alpha_j(x_j)/\alpha_k(x_j))}{\partial x_j}$, we have that

$$\Big(\frac{\partial G}{\partial X}\Big)^{-1}_{ii} = \frac{\sigma\gamma}{\tilde{P}_i \hat{\alpha}_i}\Big(1 + \sum_{j\neq i} \frac{\alpha_k(x_j)^{\sigma-1}/\Lambda_k}{\hat{\alpha}_j}\Big) \tag{A.6}$$

$$\Big(\frac{\partial G}{\partial X}\Big)^{-1}_{i\neq j} = -\frac{\sigma\gamma}{\tilde{P}_j \hat{\alpha}_j} \frac{\alpha_k(x_j)^{\sigma-1}/\Lambda_k}{\hat{\alpha}_i}. \tag{A.7}$$

Combining the two matrices gives us

$$\frac{\partial x_j}{\partial K} = \frac{\gamma \hat{\alpha}_j^{-1}}{K} \tag{A.8}$$

$$\frac{\partial x_j}{\partial \theta} = -\frac{\sigma\gamma\hat{\alpha}_j^{-1}}{1+\theta}. \tag{A.9}$$

□

A.4 Proof of Proposition 1

The Planner solves the problem laid out in Definition 2, where I will denote the multipliers on the resource constraints by μ_0 and μ_1 and those on the incentive compatibility constraints by η_{ij}.

The first-order condition with respect to the threshold rule θ satisfies

$$\mu_1 \frac{\partial Y}{\partial X}\frac{\partial X}{\partial \theta} + \sum_{i,j} \eta_{ij}\beta h'_{ij} l_j \frac{\partial (w_j/w_i)}{\partial X}\frac{\partial X}{\partial \theta} = 0, \tag{A.10}$$

where $h_{ij} \equiv h_i(w_j l_j / w_i)$ is the labor disutility of a type i household when they earn type j labor income. Solving out the equilibrium objects, we have

$$\frac{\mu_1}{\beta} Y \sum_j \Big(\frac{A_k \alpha_k(x_j)}{r}\Big)^{\sigma-1} \frac{(1+\theta)^{1-\sigma}-1}{1-\sigma}\hat{\alpha}_j^{-1} = \frac{1}{\sigma}\sum_{i,j}\eta_{ij}h'_{ij}\frac{w_j l_j}{w_i}\Big(\frac{\alpha_j(x_j)^{\sigma-1}}{\Lambda_j}\hat{\alpha}_j^{-1} - \frac{\alpha_i(x_i)^{\sigma-1}}{\Lambda_i}\hat{\alpha}_i^{-1}\Big).$$

The first-order condition with respect to capital K satisfies

$$\mu_0 = (1 + \frac{\partial Y}{\partial K} - \delta)\mu_1 + \mu_1 \frac{\partial Y}{\partial X}\frac{\partial X}{\partial K} + \sum_{i,j}\eta_{ij}\beta h'_{ij} l_j \frac{\partial (w_j/w_i)}{\partial X}\frac{\partial X}{\partial K}. \tag{A.11}$$

First, note that Equation (15) implies that capital only affects relative wages by changing the automation thresholds X. Next, note that Lemma 1 implies that the terms related to changes in the automation thresholds are second-order because they are proportional to the first-order condition for the threshold rule. Therefore, the first-order condition with respect to capital reduces to $\mu_0 = \mu_1(1 + \frac{\partial Y}{\partial K} - \delta)$.

The first-order conditions with respect to consumption $\{c_{j,0}, c_{j,1}\}$ satisfy

$$n_j \mu_0 = [\partial_j \mathcal{W} n_j + \sum_i (\eta_{ji} - \eta_{ij})] u'_{j,0} \tag{A.12}$$

$$n_j \mu_1 = [\partial_j \mathcal{W} n_j + \sum_i (\eta_{ji} - \eta_{ij})] \beta u'_{j,1}. \tag{A.13}$$

Combining with the first-order condition with respect to capital gives us

$$u'_{j,0} = (1 + \frac{\partial Y}{\partial K} - \delta)\beta u'_{j,1},$$

which is an undistorted Euler equation. Therefore, the capital tax should be zero.

□

A.5 Proof of Proposition 2

I will consider a perturbation argument of a similar form to that considered in Costinot and Werning (2023). Consider a perturbation of the threshold rule $d\theta$ as well as the resulting changes in equilibrium objects $\{dR, \{dw_j, dl_j, d\kappa_j, dx_j\}\}$. These changes reflect the total, general equilibrium effect of the perturbation of the threshold rule. To maintain government budget balance, I will adjust the tax function by a lump sum dD but otherwise keep the shape of the tax function unchanged. That is, I imagine the policy maker can freely choose the threshold rule and a uniform lump-sum tax/transfer, but otherwise I am not making any assumptions about the tax function.

At the optimum, this perturbation cannot improve welfare, so by Roy's Identity, we have

$$\sum_j \partial_j \mathcal{W} n_j \beta u'_{j,1}((1 - \tau_{l,j}) dw_j l_j + (1 - \tau_{k,j}) dR \kappa_j - dD) = 0. \tag{A.14}$$

Using Euler's Theorem, we have that

$$\begin{aligned}(1 - \delta) dK + \frac{\partial Y}{\partial X} dX + \sum_j w_j dL_j + r dK &= \sum_j \left(dw_j L_j + w_j dL_j\right) + R dK + dR K \\ \Rightarrow \frac{\partial Y}{\partial X} dX &= \sum_j dw_j L_j + dR K. \end{aligned} \tag{A.15}$$

Finally, the change in taxes must satisfy

$$dD + \sum_j n_j(\tau_{k,j}[dR\kappa_j + Rd\kappa_j] + \tau_{l,j}[dw_j l_j + w_j dl_j]) = 0. \tag{A.16}$$

Combining, we have

$$-Y\sum_j \Big(\frac{A_k\alpha_k(x_j)}{r}\Big)^{\sigma-1}\frac{(1+\theta)^{1-\sigma}-1}{1-\sigma}dx_j = \sum_j n_j(\lambda_j-1)((1-\tau_{l,j})w_j l_j d\ln(w_j) + (1-\tau_{k,j})R\kappa_j d\ln(R)) + \sum_j n_j(\tau_{l,j}w_j l_j d\ln(l_j) + \tau_{k,j}R\kappa_j d\ln(\kappa_j)),$$

which gives us Equation (27) of Proposition 2 when evaluated with covariance/expectation notation.

□

A.6 Proof of Corollary 1

First, to clearly distinguish between income and substitution effects, I will write the household budget as

$$(1-\tau_{k,j})R\kappa_j + (1-\tau_{l,j})w_j l_j - c_{j1} - D_j \geq 0,$$

where $D_j \equiv \mathcal{T}_j - \tau_{k,j}R\kappa_j - \tau_{l,j}w_j l_j$ represents inframarginal taxation that cannot be accounted for by linear factor taxes. Next, we want to consider a tax reform $d\mathcal{T}$ that replicates the effect of a threshold rule perturbation $d\theta$ on income before behavioral responses. That is,

$$\begin{aligned}&(1-\tau_{l,j})w_j l_j d_\theta \ln(w_j) + (1-\tau_{k,j})R\kappa_j d_\theta \ln(R)\\ &\quad = (1-\tau_{l,j})w_j l_j(d_{\mathcal{T}}\ln(1-\tau_{l,j}) + d_{\mathcal{T}}\ln(w_j)) + (1-\tau_{k,j})R\kappa_j(d_{\mathcal{T}}\ln(1-\tau_{k,j}) + d_{\mathcal{T}}\ln(R)) - d_{\mathcal{T}}D_j\end{aligned}$$

Note that Roy's Identity implies that enacting each of these perturbations in the opposite direction leaves household utility unchanged.[A.1]

Now imagine that we can satisfy Equation (27) of Proposition 2 with this joint perturbation when we have production efficiency. This requires

$$\mathbb{E}(\tau_{l,j}w_j l_j[d_\theta \ln(l_j) - d_{\mathcal{T}}\ln(l_j)] + \tau_{k,j}R\kappa_j[d_\theta\ln(\kappa_j) - d_{\mathcal{T}}\ln(\kappa_j)]) = 0 \tag{A.17}$$

because the production loss and covariance term are both zero by construction. By the Slutsky

[A.1] That is, we enact $d\theta$ and $-d\mathcal{T}$ concurrently. Note that the lump-sum transfer needed to balance the government budget following a threshold rule perturbation is simply wrapped into the tax reform for the sake of expositional convenience.

Identity, the labor behavioral effects for both perturbations follow

$$\begin{aligned}
d\ln(l_j) =& \frac{\partial \ln(l_j)}{\partial \ln((1-\tau_{l,j})w_j)} d\ln((1-\tau_{l,j})w_j) + \frac{\partial \ln(l_j)}{\partial \ln((1-\tau_{k,j})R)} d\ln((1-\tau_{k,j})R) + \frac{\partial \ln(l_j)}{\partial D_j} dD_j \\
=& \Big[\frac{\partial \ln(l_j^h)}{\partial \ln((1-\tau_{l,j})w_j)} + \frac{\partial \ln(l_j)}{\partial D_j}(1-\tau_{l,j})w_j l_j\Big] d\ln((1-\tau_{l,j})w_j) \\
&+ \Big[\frac{\partial \ln(l_j^h)}{\partial \ln((1-\tau_{k,j})R)} + \frac{\partial \ln(l_j)}{\partial D_j}(1-\tau_{k,j})R\kappa_j\Big] d\ln((1-\tau_{k,j})R) + \frac{\partial \ln(l_j)}{\partial D_j} dD_j,
\end{aligned}$$

where l_j^h is Hicksian labor supply. An analogous equation holds for the savings behavioral effects. Because the total change in income is the same for both perturbations, we have that income effects cancel and we are left with only substitution effects. Plugging this into Equation (A.17), we have

$$\begin{aligned}
\mathbb{E}\Big(&\left[\tau_{l,j} w_j l_j \varepsilon_{l,w}^{h,j} + \tau_{k,j} R\kappa_j \varepsilon_{\kappa,w}^{h,j}\right]\left[d_\theta \ln(w_j) - d_\mathcal{T} \ln((1-\tau_{l,j})w_j)\right] \\
&+ \left[\tau_{l,j} w_j l_j \varepsilon_{l,R}^{h,j} + \tau_{k,j} R\kappa_j \varepsilon_{\kappa,R}^{h,j}\right]\left[d_\theta \ln(R) - d_\mathcal{T} \ln((1-\tau_{k,j})R)\right]\Big) = 0, \qquad \text{(A.18)}
\end{aligned}$$

where $\varepsilon_{q,p}^{h,j}$ denotes Hicksian factor supply elasticities. This gives us Equation (28) of Corollary 1.

□

As I discuss in the main text, this condition can always be met with linear taxes that are specific to each factor à la Diamond and Mirrlees (1971a,b). To show this, I will construct the tax reform directly. Set $d_\mathcal{T} \ln(1-\tau_{l,j}) = d_\theta \ln(w_j) - d_\mathcal{T} \ln(w_j)$ and $d_\mathcal{T} \ln(1-\tau_k) = d_\theta \ln(R) - d_\mathcal{T} \ln(R)$. This guarantees satisfaction of Equation (A.18), but it remains to be shown that this tax reform fully offsets income changes from threshold rule perturbations and that it balances the government's budget. Assuming no change in lump-sum taxation, one can verify by inspection that there is no net change in the household budget. Therefore, there is neither a substitution nor income effect, so the joint perturbation generates no behavioral effects. Finally, the government budget will change according to

$$\begin{aligned}
&\sum_j \tau_{l,j} w_j L_j (d\ln(w_j) - d\ln(\tau_{l,j})) + \tau_k RK (d\ln(R) - d\ln(\tau_k)) \\
&\quad = \sum_j w_j L_j d\ln(1-\tau_{l,j}) + RK d\ln(1-\tau_k),
\end{aligned}$$

where the changes considered come from the joint perturbation. Euler's Theorem implies that these changes sum to zero, so my proposed tax reform balances the government's budget as well.

A.7 Proof of Proposition 3

The Planner again solves the problem laid out in Definition 2, except now the threshold rule is unavailable: $\theta = 0$. I will again denote the multipliers on the resource constraints by μ_0 and μ_1 and those on the incentive compatibility constraints by η_{ij}.

The first-order conditions with respect to capital K and consumption $\{c_{j,0}, c_{j,1}\}$ again satisfy Equations (A.11), (A.12), and (A.13). The difference is that the terms related to changes in the automation thresholds no longer envelope out. However, we do have that $\frac{\partial Y}{\partial x_j} = 0 \ \forall j$ because profit maximization picks automation thresholds to maximize output. Therefore, we have

$$\frac{\mu_0}{\mu_1} = (1 + \frac{\partial Y}{\partial K} - \delta) - \frac{\beta}{\mu_1}\frac{\gamma}{\sigma K}\sum_{i,j}\eta_{ij}h'_{ij}\frac{w_j l_j}{w_i}\Big(\frac{\alpha_j(x_j)^{\sigma-1}}{\Lambda_j}\hat{\alpha}_j^{-1} - \frac{\alpha_i(x_i)^{\sigma-1}}{\Lambda_i}\hat{\alpha}_i^{-1}\Big). \tag{A.19}$$

The consumption first-order conditions imply that $\frac{\mu_0}{\mu_1} = \frac{u'_{m,0}}{\beta u'_{m,1}}$, so plugging into the household Euler equation, we have

$$\tau_{k,m} = \frac{\beta}{\mu_1}\frac{\gamma}{\sigma RK}\sum_{i,j}\eta_{ij}h'_{ij}\frac{w_j l_j}{w_i}\Big(\frac{\alpha_j(x_j)^{\sigma-1}}{\Lambda_j}\hat{\alpha}_j^{-1} - \frac{\alpha_i(x_i)^{\sigma-1}}{\Lambda_i}\hat{\alpha}_i^{-1}\Big),$$

Note that the RHS does not depend on the household type m whose Euler equation we are considering, so a linear capital tax is optimal.

To derive Equation (29) of Proposition 3, we can again take a perturbation approach. Consider a uniform increase in the capital tax $d\tau_k$ as well as the resulting changes in equilibrium objects $\{dR, \{dw_j, dl_j, d\kappa_j, dx_j\}\}$, as before. Again, I will adjust the tax function by a lump-sum dD to maintain government budget balance.

At the optimum, this perturbation cannot improve welfare, so by Roy's Identity, we have

$$\sum_j \partial_j \mathcal{W} n_j \beta u'_{j,1}((1-\tau_{l,j})dw_j l_j + (1-\tau_k)dR\kappa_j - dD - d\tau_k R\kappa_j) = 0. \tag{A.20}$$

Because production is efficient with $\theta = 0$, Euler's Theorem now implies

$$\sum_j dw_j L_j + dRK = 0. \tag{A.21}$$

Finally, the change in taxes must satisfy

$$dD + \sum_j n_j(\tau_k[dR\kappa_j + Rd\kappa_j] + \tau_{l,j}[dw_j l_j + w_j dl_j] + d\tau_k R\kappa_j) = 0. \tag{A.22}$$

Combining, we have

$$\sum_j n_j(\lambda_j - 1)R\kappa_j = \sum_j n_j(\lambda_j - 1)\big((1-\tau_{l,j})w_j l_j d\ln(w_j) + (1-\tau_k)R\kappa_j d\ln(R)\big) + \sum_j n_j\big(\tau_{l,j} w_j l_j d\ln(l_j) + \tau_k R\kappa_j d\ln(\kappa_j)\big).$$

□

A.8 Cobb-Douglas Production

When the aggregation of tasks is Cobb-Douglas, we have

$$\ln(Y) = \sum_j \int_{\mathcal{X}_j} \ln(y(x))dx, \tag{A.23}$$

which requires that the total mass of tasks sums to one: $\sum_j |\mathcal{X}_j| = 1$. Plugging factor prices into the ideal price index, we have

$$Y = \bar{A}\Big(\frac{A_k K}{\Lambda_k}\Big)^{\Lambda_k} \prod_j \Big(\frac{A_j L_j}{\Lambda_j}\Big)^{\Lambda_j}, \tag{A.24}$$

where $\bar{A} \equiv \exp\big(\int_{\Omega_k} \ln(\alpha_k(x))dx + \sum_j \int_{\Omega_j} \ln(\alpha_j(x))dx\big)$. We can now consider the effect of changes in automation thresholds on output. Differentiating, we have

$$\frac{\partial Y}{\partial x_j} = Y \ln(1+\theta).$$

B Extensions

B.1 Infinite Horizon

This appendix considers an infinite horizon extension of my baseline two period model. To start, household utility is now a discounted sum of future utility flows:

$$V_j = \sum_{t \geq 0} \beta^t \big(u(c_{jt}) - h_j(l_{jt})\big). \tag{B.1}$$

Similarly, there is a resource constraint in each period that allows for consumption and savings:

$$\sum_j n_j c_{jt} + K_{t+1} = Y_t + (1-\delta)K_t. \tag{B.2}$$

For household budgets, I will make the without loss assumption that all taxes are paid in the initial period as a function of households' planned saving and labor supply decisions for all

future periods. The household budget constraint can then be written recursively as follows:

$$c_{j,0} + \kappa_{j,1} = R_0 \kappa_{j,0} + w_{j,0} l_{j,0} - \mathcal{T}(\{R_t \kappa_{jt}, w_{jt} l_{jt}\}) \tag{B.3}$$

$$c_{jt} + \kappa_{jt+1} = R_t \kappa_{jt} + w_{jt} l_{jt} \quad \forall t > 0. \tag{B.4}$$

In this setting, the tax function again implies implicit tax wedges that are now in present value form: $\tau_{k,jt} \equiv \prod_{s=1}^{t}(1-\tau_{k,js}) R_s \frac{\partial \mathcal{T}_j}{\partial (R\kappa)_t}$ and $\tau_{l,jt} \equiv \prod_{s=1}^{t}(1-\tau_{k,js}) R_s \frac{\partial \mathcal{T}_j}{\partial (wl)_t}$. Note that, relative to the baseline case, I no longer have a sunk initial endowment, with that role now being played by a sunk initial stock of capital. Therefore, the logic of an inelastic endowment is the same. The production side of the economy is the same as in the baseline model, but I will allow the productivities and task sets to change over time. That is, the comparative advantage of, and tasks available to, a given type of worker do not need to be constant over time. This allows for arbitrary (but exogenous) forms of technological change, with workers and capital producing an evolving set of tasks.

In this setting, the Planning problem can be defined as follows.

Definition B.1 (Infinite Horizon Non-Linear Tax Planning Problem)**.** *The Planner solves*

$$\max_{\{\{c_{jt}, l_{jt}\}, K_t, \theta_t\}} \mathcal{W}(\{n_j V_j\}) \;\; s.t. \;\; \text{(B.2)}, \tag{B.5}$$

$$\sum_{t \geq 0} \beta^t (u(c_{it}) - h_i(l_{it})) \geq \sum_{t \geq 0} \beta^t (u(c_{jt}) - h_i(w_{jt} l_{jt} / w_{it})) \quad \forall i, j.$$

We again have a mechanism design setup where the allocation of consumption, labor, and investment must be chosen at each point in time. The Planner can also stipulate a sequence of threshold rules that are allowed to vary over time. Lemma 1 applies period-by-period, so setting the threshold rule optimally again allows the Planner to invoke "as-if" Atkinson and Stiglitz (1976).

Proposition B.1. *When the threshold rule is set optimally, capital taxes beyond the initial period should be zero:*

$$\tau_{k,jt} = 0 \quad \forall j, t > 0. \tag{B.6}$$

Proof. The Planner solves the problem laid out in Definition B.1, where I will denote the multipliers on the resource constraints by μ_t and those on the incentive compatibility constraints by η_{ij}.

The first-order conditions with respect to the threshold rules θ_t satisfy

$$\mu_t \frac{\partial Y_t}{\partial X_t} \frac{\partial X_t}{\partial \theta_t} + \sum_{i,j} \eta_{ij} \beta^t h'_{ij,t} l_{jt} \frac{\partial (w_{jt}/w_{it})}{\partial X_t} \frac{\partial X_t}{\partial \theta_t} = 0. \tag{B.7}$$

Solving out the equilibrium objects, we have

$$\frac{\mu_t}{\beta^t} Y_t \sum_j \Big(\frac{A_{kt}\alpha_{kt}(x_{jt})}{r_t}\Big)^{\sigma-1} \frac{(1+\theta_t)^{1-\sigma}-1}{1-\sigma}\hat{\alpha}_{jt}^{-1} = \frac{1}{\sigma}\sum_{i,j}\eta_{ij}h'_{ij,t}\frac{w_{jt}l_{jt}}{w_{it}}\Big(\frac{\alpha_{jt}(x_{jt})^{\sigma-1}}{\Lambda_{jt}}\hat{\alpha}_{jt}^{-1} - \frac{\alpha_{it}(x_{it})^{\sigma-1}}{\Lambda_{it}}\hat{\alpha}_{it}^{-1}\Big).$$

As in Proposition 1, optimal choice of the threshold rule implies that the change in automation from changes in the capital stock has a second-order effect on welfare, so the first-order conditions with respect to capital K_t satisfy

$$\mu_{t-1} = (1 + \frac{\partial Y_t}{\partial K_t} - \delta)\mu_t. \tag{B.8}$$

The first-order conditions with respect to consumption c_{jt} satisfy

$$n_j\mu_t = [\partial_j \mathcal{W} n_j + \sum_i(\eta_{ji} - \eta_{ij})]\beta^t u'_{jt}, \tag{B.9}$$

which again implies an undistorted Euler equation:

$$u'_{jt-1} = (1 + \frac{\partial Y_t}{\partial K_t} - \delta)\beta u'_{jt}.$$

Therefore, capital taxes beyond the initial period should be zero.

□

As in the main text, we can then use a perturbation approach to characterize the optimal sequence of threshold rules.

Proposition B.2. *For any tax function with a linear capital tax and freely chosen lump-sum, the optimal threshold rules should be set to satisfy*

$$\begin{aligned}
-\sum_{t\geq 0}\prod_{s=1}^{t}\frac{1}{\tilde{R}_s} Y_t \sum_j \Big(\frac{A_{kt}\alpha_{kt}(x_{jt})}{r_t}\Big)^{\sigma-1}\frac{(1+\theta_t)^{1-\sigma}-1}{1-\sigma}dx_{jt} \\
= Cov(\lambda_j, \sum_{t\geq 0}\prod_{s=1}^{t}\frac{1}{\tilde{R}_s}[(1-\tau_{l,jt})w_{jt}l_{jt}d\ln(w_{jt}) + (1-\tau_{k,t})R_t\kappa_{jt}d\ln(R_t)]) \\
+ \mathbb{E}(\sum_{t\geq 0}\prod_{s=1}^{t}\frac{1}{\tilde{R}_s}[\tau_{l,jt}w_{jt}l_{jt}d\ln(l_{jt}) + \tau_{k,t}R_t\kappa_{jt}d\ln(\kappa_{jt})])
\end{aligned} \tag{B.10}$$

for perturbations generated by changes in each threshold rule.

This formula has the same interpretation as in the two period model. The left hand side represents the efficiency cost of distorting automation decisions, while the right hand side represents the equity benefit of reducing automation as well as the fiscal impact of the

perturbation. The main difference is that we now need to consider behavioral changes at all points in time. Thus, we simply consider these various effects in present value form with the tax-inclusive effective discount rate given by $\tilde{R}_t \equiv (1-\tau_{k,t})R_t$ and welfare weights given by marginal utility in the initial period $\lambda_j \equiv \partial_j \mathcal{W} u'_{j,0}/\mathbb{E}(\partial_i \mathcal{W} u'_{i,0})$.

Proof. I will again consider a perturbation of the threshold rule in a given period: $d\theta_t$. This perturbation, in principle, generates general equilibrium changes across all equilibrium objects $\{dR_t, \{dw_{jt}, dl_{jt}, d\kappa_{jt}, dx_{jt}\}\}$. I will again maintain budget balance with a lump sum adjustment dD.

At the optimum, this perturbation cannot improve welfare, so by Roy's Identity, we have

$$\sum_j \partial_j \mathcal{W} n_j u'_{j,0} \Big(\sum_{t\geq 0} \prod_{s=1}^{t} \frac{1}{\tilde{R}_s} \left[(1-\tau_{l,jt}) dw_{jt} l_{jt} + (1-\tau_{k,t}) dR_t \kappa_{jt} \right] - dD \Big) = 0. \tag{B.11}$$

Note that the discount term $\prod_{s=1}^{t} \frac{1}{\tilde{R}_s}$ includes capital taxation because the marginal value of changes in income in a given period connects to the marginal utility of consumption in the initial period through the household's Euler equation. Using Euler's Theorem, we have at each point in time that

$$\begin{aligned} (1-\delta)dK_t + \frac{\partial Y_t}{\partial X_t} dX_t + \sum_j w_{jt} dL_{jt} + r_t dK_t &= \sum_j \left(dw_{jt} L_{jt} + w_{jt} dL_{jt} \right) + R_t dK_t + dR_t K_t \\ \Rightarrow \frac{\partial Y_t}{\partial X_t} dX_t &= \sum_j dw_{jt} L_{jt} + dR_t K_t. \end{aligned} \tag{B.12}$$

Finally, the change in taxes must satisfy

$$dD + \sum_j n_j \sum_{t\geq 0} \prod_{s=1}^{t} \frac{1}{\tilde{R}_s} \left(\tau_{k,t} [dR_t \kappa_{jt} + R_t d\kappa_{jt}] + \tau_{l,jt} [dw_{jt} l_{jt} + w_{jt} dl_{jt}] \right) = 0. \tag{B.13}$$

Combining, we have

$$\begin{aligned} &- \sum_{t\geq 0} \prod_{s=1}^{t} \frac{1}{\tilde{R}_s} Y_t \sum_j \Big(\frac{A_{kt} \alpha_{kt}(x_{jt})}{r_t} \Big)^{\sigma-1} \frac{(1+\theta_t)^{1-\sigma} - 1}{1-\sigma} dx_{jt} \\ &\qquad = \sum_j n_j (\lambda_j - 1) \sum_{t\geq 0} \prod_{s=1}^{t} \frac{1}{\tilde{R}_s} \left[(1-\tau_{l,jt}) w_{jt} l_{jt} d\ln(w_{jt}) + (1-\tau_{k,t}) R_t \kappa_{jt} d\ln(R_t) \right] \\ &\qquad + \sum_j n_j \sum_{t\geq 0} \prod_{s=1}^{t} \frac{1}{\tilde{R}_s} \left[\tau_{l,jt} w_{jt} L_{jt} d\ln(L_{jt}) + \tau_{k,t} R_t \kappa_{jt} d\ln(\kappa_{jt}) \right], \end{aligned}$$

which gives us Equation (B.10) of Proposition B.2 when evaluated with covariance/expectation notation.

□

We can again ask what happens in the absence of the threshold rule. As in the baseline model, the dependence of relative wages on capital accumulation opens up the possibility of capital taxation, and we can use a perturbation argument to solve for the optimal capital tax formula.

Proposition B.3. *When the threshold rule is unavailable, capital taxes should be linear and set in each period to satisfy*

$$Cov(\lambda_j, \prod_{s=1}^{t} \frac{1}{\tilde{R}_s} R_t \kappa_{jt}) = Cov(\lambda_j, \sum_{t\geq 0} \prod_{s=1}^{t} \frac{1}{\tilde{R}_s} ((1-\tau_{l,jt}) w_{jt} l_{jt} d\ln(w_{jt}) + (1-\tau_{k,t}) R_t \kappa_{jt} d\ln(R_t))) + \mathbb{E}(\sum_{t\geq 0} \prod_{s=1}^{t} \frac{1}{\tilde{R}_s} (\tau_{l,jt} w_{jt} l_{jt} d\ln(l_{jt}) + \tau_{k,t} R_t \kappa_{jt} d\ln(\kappa_{jt}))). \quad \text{(B.14)}$$

for perturbations generated by changes in each capital tax.

Now we have capital taxes specific to each period, but they remain linear and solve a formula with the same interpretation as before. Like in the threshold rule perturbation formula, the only substantive difference comes from the need to consider changes to equilibrium outcomes at every point in time and take the present value of this sequence of changes. Therefore, all of the conclusions of the two period model carry over to an infinite horizon setting with minimal adjustment.

Proof. The Planner again solves the problem laid out in Definition B.1, except now the threshold rule is unavailable: $\theta_t = 0$. I will again denote the multipliers on the resource constraints by μ_t and those on the incentive compatibility constraints by η_{ij}.

The first-order conditions with respect to consumption c_{jt} again satisfy Equation (B.9). For capital, changes in the automation thresholds no longer envelope out, but we again have $\frac{\partial Y_t}{\partial x_{jt}} = 0$ because profit maximization picks automation thresholds to maximize output. Therefore, we have

$$\frac{\mu_{t-1}}{\mu_t} = (1 + \frac{\partial Y_t}{\partial K_t} - \delta) - \frac{\beta^t}{\mu_t} \frac{\gamma_t}{\sigma K_t} \sum_{i,j} \eta_{ij} h'_{ij,t} \frac{w_{jt} l_{jt}}{w_{it}} \Big(\frac{\alpha_{jt}(x_{jt})^{\sigma-1}}{\Lambda_{jt}} \hat{\alpha}_{jt}^{-1} - \frac{\alpha_{it}(x_{it})^{\sigma-1}}{\Lambda_{it}} \hat{\alpha}_{it}^{-1} \Big). \quad \text{(B.15)}$$

Again, combining with the consumption first-order conditions and household Euler equation, we have

$$\tau_{k,mt} = \frac{\beta^t}{\mu_t} \frac{\gamma_t}{\sigma R_t K_t} \sum_{i,j} \eta_{ij} h'_{ij,t} \frac{w_{jt} l_{jt}}{w_{it}} \Big(\frac{\alpha_{jt}(x_{jt})^{\sigma-1}}{\Lambda_{jt}} \hat{\alpha}_{jt}^{-1} - \frac{\alpha_{it}(x_{it})^{\sigma-1}}{\Lambda_{it}} \hat{\alpha}_{it}^{-1} \Big),$$

Note that the RHS does not depend on the household type m whose Euler equation we are considering, so a linear capital tax is optimal.

To derive Equation (B.14) of Proposition B.3, I will again consider a perturbation $d\tau_{k,t}$ of the capital tax at time t. As before, I will consider the resulting perturbations to equilibrium objects $\{dR_t, \{dw_{jt}, dl_{jt}, d\kappa_{jt}, dx_{jt}\}\}$ as well as a lump-sum adjustment dD to maintain budget balance.

At the optimum, this perturbation cannot improve welfare, so by Roy's Identity, we have

$$\sum_j \partial_j \mathcal{W} n_j u'_{j,0} \Big(\sum_{t\geq 0} \prod_{s=1}^{t} \frac{1}{\tilde{R}_s} [(1-\tau_{l,jt}) dw_{jt} l_{jt} + (1-\tau_{k,t}) dR_t \kappa_{jt}] - dD - \prod_{s=1}^{t} \frac{1}{\tilde{R}_s} d\tau_{k,t} R_t \kappa_{jt} \Big) = 0. \tag{B.16}$$

Changes in automation thresholds have a second-order impact on output when the threshold rule is zero, so Euler's Theorem implies

$$\sum_j dw_{jt} L_{jt} + dR_t K_t = 0. \tag{B.17}$$

Finally, the change in taxes must satisfy

$$dD + \sum_j n_j \sum_{t\geq 0} \prod_{s=1}^{t} \frac{1}{\tilde{R}_s} (\tau_{k,t} [dR_t \kappa_{jt} + R_t d\kappa_{jt}] + \tau_{l,jt} [dw_{jt} l_{jt} + w_{jt} dl_{jt}]) + \prod_{s=1}^{t} \frac{1}{\tilde{R}_s} d\tau_{k,t} R_t K_t = 0. \tag{B.18}$$

Combining, we have

$$\begin{aligned} 0 = & \sum_j n_j (\lambda_j - 1) \Big[\sum_{t\geq 0} \prod_{s=1}^{t} \frac{1}{\tilde{R}_s} ((1-\tau_{l,jt}) w_{jt} l_{jt} d\ln(w_{jt}) + (1-\tau_{k,t}) R_t \kappa_{jt} d\ln(R_t)) - \prod_{s=1}^{t} \frac{1}{\tilde{R}_s} R_t \kappa_{jt} \Big] \\ & + \sum_j n_j \sum_{t\geq 0} \prod_{s=1}^{t} \frac{1}{\tilde{R}_s} [\tau_{l,jt} w_{jt} l_{jt} d\ln(l_{jt}) + \tau_{k,t} R_t \kappa_{jt} d\ln(\kappa_{jt})], \end{aligned}$$

which gives us the desired result.

□

B.2 Within-Type Heterogeneity

This appendix considers a version of the model where there is heterogeneity within types of workers as well as across. Namely, I will assume that workers also have an exogenous endowment of efficiency units of labor which scales their labor effort and is private information. This can be thought of as a within-type skill level; i.e. a unit of effort from one person can mean more productive output than another. Denote this skill endowment by z, so households are indexed by identifier (j, z). For a given type, z will follow the distribution $M_j(z)$, so total labor supply will follow $L_j \equiv n_j \int z l_{jz} dM_j(z)$. I will also denote the PDF by $m_j(z)$.

Household utility will again be given by Equation (1), but now the household budget

constraint follows

$$Rc_{jz,0} + c_{jz,1} = R\omega_{jz}\bar{Y} + w_j z l_{jz} - \mathcal{T}(R\kappa_{jz}, w_j z l_{jz}). \tag{B.19}$$

Production does not change, so the Planner now solves the following problem.

Definition B.2 (Planning Problem with Within-Type Heterogeneity)**.** *The Planner solves*

$$\max_{\{c_{jz,0}, c_{jz,1}, l_{jz}\}, K, \theta} \mathcal{W}(\{n_j m_j(z) V_{jz}\}) \;\; s.t. \;\; (7), (8), \tag{B.20}$$
$$u(c_{iz,0}) + \beta(u(c_{iz,1}) - h_i(l_{iz})) \geq u(c_{jz',0}) + \beta(u(c_{jz',1}) - h_i(w_j z' l_{jz'}/w_i z)) \quad \forall i, j, z, z'.$$

The only difference from the baseline problem is that incentive compatibility must now consider both dimensions of private information. This does not substantively change the arguments of this paper as relative wages still enter the problem in an analogous manner as before. Therefore, Lemma 1 again implies "as-if" Atkinson and Stiglitz (1976), where capital taxes should be set to zero.

As for the perturbation arguments of Propositions 2 and 3, the only relevant difference is that the covariance terms now need to integrate over both dimensions of heterogeneity. With this adjustment, the same formulas hold.

B.3 Intrinsic Value from Work

In this appendix, I consider the possibility that workers get direct utility from their marginal product. This is one way of capturing the idea that people derive value from their work, and by focusing on marginal products, I am assuming that the value people derive tracks in some way the usefulness of their work in production. The main modification this requires is that household preference now follow

$$V_j = u(c_{j,0}) + \beta(u(c_{j,1}) - h_j(l_j, w_j)). \tag{B.21}$$

With this setup, the equilibrium condition for the labor market becomes a fixed point problem that satisfies

$$\frac{\partial_l h_j(l_j, w_j)}{u'(c_{j,1})} = (1 - \tau_{l,j}) w_j, \tag{B.22}$$

where the household still takes wages as given. Thus, we have a Planning problem given by the following.

Definition B.3 (Planning Problem with Value from Work)**.** *The Planner solves*

$$\max_{\{c_{j,0},c_{j,1},l_j\},K,\theta} \mathcal{W}(\{n_j V_j\}) \;\; s.t. \;\; (7),(8), \tag{B.23}$$
$$u(c_{i,0}) + \beta(u(c_{i,1}) - h_i(l_i, w_i)) \geq u(c_{j,0}) + \beta(u(c_{j,1}) - h_i(w_j l_j / w_i, w_i)) \quad \forall i, j.$$

Taking the first-order condition with respect to capital, we have

$$\mu_0 = (1 + \frac{\partial Y}{\partial K} - \delta)\mu_1 + \underbrace{\frac{\partial \mathcal{L}}{\partial X}\frac{\partial X}{\partial K}}_{=0} + \sum_i \left[\partial_i \mathcal{W} n_i \frac{\partial V_i}{\partial w_i} + \sum_j \beta \eta_{ij}(\partial_w h_{ij} - \partial_w h_i)\right] \frac{w_i}{\sigma} \frac{\partial \ln(Y)}{\partial K}, \tag{B.24}$$

where $\frac{\partial \mathcal{L}}{\partial X}$ are all of the effects coming from changes in automation thresholds. This term again sums to zero because of the envelope condition from optimal choice of the threshold rule. The remaining additional terms represent the direct increase in wages from capital deepening, which both directly affects welfare through utility and changes incentive compatibility by potentially changing the relative utility that a type receives by truthfully reporting.

Therefore, we no longer necessarily have a zero capital tax, but deviations from zero will not be due to concerns about displacement from automation. Instead, they will be due to the interaction between capital deepening and the intrinsic value from work. Moreover, in the case where the functions $h_j(.)$ are additively separable in wages and labor, we will have $\partial_w h_{ij} - \partial_w h_i = 0$. Therefore, capital deepening will not affect incentive compatibility; it will only improve welfare directly by increasing marginal products. This means that the Planner should *subsidize* capital as a way of raising the value people derive from their work.

B.4 Type Switching

In the baseline model, I assume that type assignment is fixed. This appendix relaxes that assumption by considering a dynamic, discrete choice framework like that used in the spatial literature (Caliendo et al., 2019; Donald et al., 2025a). I will assume that a worker of a given type receives an idiosyncratic vector of preference shocks at the conclusion of the initial period, then decides their type for the following period to maximize their continuation utility level. Specifically, workers of type i with shocks ϱ_{ij} solve the following problem

$$\max_{\{c_{ij,0},c_{ij,1},l_{ij},\hat{j}\}} u(c_{i\hat{j},0}) + \beta(u(c_{i\hat{j},1}) - h_{\hat{j}}(l_{i\hat{j}}) + \varrho_{i\hat{j}}) \;\; s.t. \tag{B.25}$$
$$R c_{ij,0} + c_{ij,1} = R\omega_i \bar{Y} + w_j l_{ij} - \mathcal{T}(R\kappa_{ij}, w_j l_{ij}) \;\; \forall j.$$

That is, households make a contingent plan for consumption, savings, and labor supply for every choice of type, then pick the type $\hat{j}$ that maximizes their utility level conditional on their preference shock. Note that because the preference shocks are additive, they do not affect the

contingent allocative choices for a given potential type choice. I make no assumptions about the distribution of the preference shocks, and I will denote the fraction of initially type i agents who choose j in the second period by ϕ_{ij}. Furthermore, I will assume that the preference shocks are private information.

To map from this specification of discrete choice to a Planning problem, I will make use of the representation result of Hofbauer and Sandholm (2002) and Donald et al. (2025b). They show that discrete choice problems of this type can be represented as a maximization problem over choice probabilities with an appropriately defined cost function. That is, the expected utility of type i agents– where the expectation is over preference shocks – in the initial period is given by

$$\begin{aligned} \max_{\{c_{ij,0},c_{ij,1},l_{ij},\phi_{ij}\}} \sum_j \phi_{ij}U_{ij} - \Xi_i(\{\phi_{ij}\}) \;\; s.t. \\ \sum_j \phi_{ij} = 1 \\ Rc_{ij,0} + c_{ij,1} = R\omega_i\bar{Y} + w_j l_{ij} - \mathcal{T}(R\kappa_{ij}, w_j l_{ij}) \;\; \forall j, \end{aligned} \tag{B.26}$$

where $U_{ij} \equiv u(c_{ij,0}) + \beta\big(u(c_{ij,1}) - h_j(l_{ij})\big)$ is the utility before preference shocks that one achieves with a given type pair. That is, the above formulation of atomistic agents with preference shocks is equivalent to a representative agent with a choice cost function that embeds the utility consequences from preference shocks of different choice probabilities. Using this representative formulation, we can define the Planning problem as follows.

Definition B.4 (Planning Problem with Type Switching)**.** *The Planner solves*

$$\begin{aligned} \max_{\{c_{ij,0},c_{ij,1},l_{ij}\},K,\theta} \mathcal{W}\big(\{n_j[\sum_j \phi_{ij}U_{ij} - \Xi_i(\{\phi_{ij}\})]\}\big) \;\; s.t. \;\; (7),(8), \\ u(c_{ij,0}) + \beta\big(u(c_{ij,1}) - h_j(l_{ij})\big) \geq u(c_{i'j',0}) + \beta\big(u(c_{i'j',1}) - h_j(w_{j'}l_{i'j'}/w_j)\big) \quad \forall i,j,i',j'. \end{aligned} \tag{B.27}$$

Several features of this Planning problem warrant discussion. First, following Donald et al. (2025b), I assume that welfare is utilitarian over preference shocks for a given initial type. This can either be because welfare is taken from an ex-ante perspective before the shocks are realized, or, as shown by Eden (2020) and Donald et al. (2025b), because welfare satisfies a set of axioms over the aggregation of heterogeneous preferences. Second, choice probabilities are given by an implicit function that solves

$$\begin{aligned} \phi_i(\{U_{ij}\}) = \operatorname*{argmax}_{\{\phi_{ij}\}} \sum_j \phi_{ij}U_{ij} - \Xi_i(\{\phi_{ij}\}) \;\; s.t. \\ \sum_j \phi_{ij} = 1, \end{aligned} \tag{B.28}$$

This is a form of incentive compatibility in that choice probabilities have to respect private optimization as described by (B.26). Third and finally, the additive nature of preference shocks implies that they can be ignored in the incentive compatibility constraints. All people moving from i to j will be assigned the same bundle, and the incentive to truthfully reveal one's type pair does not depend on preference shocks.[B.1]

When workers can switch their type, an optimal threshold rule still implies that we should not tax capital.

Proposition B.4. *When the threshold rule is set optimally, capital taxes should be zero:*

$$\tau_{k,ij} = 0 \quad \forall i, j. \tag{B.29}$$

Proof. The first-order condition with respect to capital follows

$$\mu_0 = (1 + \frac{\partial Y}{\partial K} - \delta)\mu_1 + \underbrace{\frac{\partial \mathcal{L}}{\partial X}\frac{\partial X}{\partial K}}_{=0}, \tag{B.30}$$

where we again have that all of the effects of changes in automation thresholds sum to zero when the threshold rule is chosen optimally.

More important then are the first-order conditions with respect to consumption $\{c_{ij,0}, c_{ij,1}\}$. These follow

$$n_i \phi_{ij} \mu_0 = [\partial_i \mathcal{W} n_i \phi_{ij} + \sum_{i',j'} (\eta_{ij,i'j'} - \eta_{i'j',ij}) + \sum_{j'} \frac{\partial \mathcal{L}}{\partial \phi_{ij'}} \frac{\partial \phi_{ij'}}{\partial U_{ij}}] u'_{ij,0} \tag{B.31}$$

$$n_i \phi_{ij} \mu_1 = [\partial_i \mathcal{W} n_i \phi_{ij} + \sum_{i',j'} (\eta_{ij,i'j'} - \eta_{i'j',ij}) + \sum_{j'} \frac{\partial \mathcal{L}}{\partial \phi_{ij'}} \frac{\partial \phi_{ij'}}{\partial U_{ij}}] \beta u'_{ij,1}, \tag{B.32}$$

where $\frac{\partial \mathcal{L}}{\partial \phi_{ij}}$ are all of the effects of changes in choice probabilities. Therefore, we again have that households should follow an undistorted Euler equation, implying zero capital taxation. □

B.5 Endogenous Innovation

In this appendix, I consider a version of the model where the direction of innovation is endogenous, using a modified version of the model in Donald (2023). In particular, I will focus on labor- and capital-augmenting technology, rather than innovation that affects task-specific productivities or the set of tasks produced in the economy. To start, I will assume that, rather

[B.1] One may be tempted to think that the Planner could somehow improve welfare by conditioning allocations on preference shocks, but as shown by Donald et al. (2025a), this will not be incentive compatible. If different bundles are offered to the same type pair as a function of preference shocks, people will always report that their preference shock is the one with the utility maximizing bundle. Therefore, the fact that preference shocks are private information implies no differences in bundles for a given type pair.

than supplying factors directly to tasks, each factor is used in an earlier step in the production process which produces factor services. That is, for each type of labor we have

$$\begin{aligned}\int_0^1 \bar{\ell}_j(\bar{x})d\bar{x} &= L_j \\ \ln(\mathcal{Y}_j) &= \int_0^1 \ln(a_l(\bar{x})\bar{\ell}_j(\bar{x}))d\bar{x} \\ \int_{\mathcal{X}_j} \ell_j(x)dx &= \mathcal{Y}_j,\end{aligned}$$

and likewise for capital. That is, a factor is used to competitively produce a unit interval of intermediates, which aggregate Cobb-Douglas up to a factor service. The total supply of this factor service is then spread across production tasks. The productivity schedule for the intermediates $\{a_k(\bar{x}), a_l(\bar{x})\}$ defines labor- and capital-augmenting technology. Note that this recovers our baseline formulation (with some relabeling and competitive pricing) if we define $A_k \equiv \exp\left(\int_0^1 \ln(a_k(\bar{x}))d\bar{x}\right)$ and $A_l \equiv \exp\left(\int_0^1 \ln(a_l(\bar{x}))d\bar{x}\right)$.[B.2]

I will allow the two schedules of factor-augmenting technologies to evolve endogenously with step-ladder innovation, where the step size is given by υ. Intermediate producers have market power, but because the unrestricted markup in a Cobb-Douglas demand system is infinite, I will assume that intermediate producers limit price one step down the productivity ladder to a competitive fringe. This implies that profits for each factor-augmenting technology are proportional to that factor's income share; a standard market size result. That is, $\Pi_k \propto S_k Y$ and $\Pi_l \propto (1 - S_k)Y$.

Denote by $\mathcal{Z}_k$ and $\mathcal{Z}_l$ the mass of intermediates that receive innovations. Thus, the two technologies will evolve according to

$$\begin{aligned}A_k &= \upsilon^{\mathcal{Z}_k} A_{k,0} \\ A_l &= \upsilon^{\mathcal{Z}_l} A_{l,0},\end{aligned}$$

where $A_{k,0}$ and $A_{l,0}$ are exogenous initial states for technology. Innovations are produced with scientists according to the idea production functions $\mathcal{Z}_k = s_k^\varsigma$ and $\mathcal{Z}_k = s_k^\varsigma$, where ς is a measure of diminishing returns in innovation. The total supply of scientists is fixed at $\mathcal{S}$, so the focus here is on the composition off research effort, not the overall level. A competitive research firm solves

$$\max_{s_k, s_l} \xi_k \mathcal{Z}_k \Pi_k + \xi_l \mathcal{Z}_l \Pi_l - w_s(s_k + s_l), \tag{B.33}$$

where ξ_k and ξ_l are innovation subsidies, and w_s is the wage for scientists. By no arbitrage, the price of an innovation is the profit it can generate. This gives us the following equilibrium

[B.2] The relabeling is just a matter of interpreting task-level factor supply in efficiency units and absorbing differences in factor-augmenting technology across labor types in their task-level productivity schedule.

condition for innovation:

$$\left(\frac{s_l}{s_k}\right)^{1-\varsigma} = \left(\frac{\xi_l}{\xi_k}\right)\left(\frac{\Pi_l}{\Pi_k}\right). \tag{B.34}$$

Thus, the direction of innovation is driven by market size and innovation subsidies. We can now consider a Planning problem in this endogenous innovation context.

Definition B.5 (Planning Problem with Innovation)**.** *The Planner solves*

$$\begin{aligned}
&\max_{\{c_{j,0},c_{j,1},l_j\},s_k,s_l,A_k,A_l,K,\theta} \mathcal{W}(\{n_j V_j\}) \;\; s.t. \;\; (7),(8), \\
&u(c_{i,0}) + \beta\big(u(c_{i,1}) - h_i(l_i)\big) \geq u(c_{j,0}) + \beta\big(u(c_{j,1}) - h_i(w_j l_j / w_i)\big) \quad \forall i,j \\
&\ln(A_k) = \ln(\upsilon)\mathcal{Z}_k + \ln(A_{k0}) \\
&\ln(A_l) = \ln(\upsilon)\mathcal{Z}_l + \ln(A_{l0}) \\
&s_k + s_l = \mathcal{S}.
\end{aligned} \tag{B.35}$$

This formulation assumes that the Planner can freely subsidize markups and innovation as well as tax and redistribute all profits and innovation income (subject to private information about type). Note that I also abstract from knowledge spillovers as a justification for intervention. See Liu and Ma (2021) and Donald (2023) for optimal innovation policy in the presence of a spillover network. While these are stark assumptions, they retain focus on the relevant question for this paper: do concerns about displacement from automation warrant intervention in the direction of innovation? To answer this question, we will first derive another Lemma.

Lemma B.1. *The equilibrium change in automation thresholds from changing factor-augmenting technologies is proportional to that from changing the threshold rule. That is,*

$$\frac{\partial x_j}{\partial A_l} \propto \frac{\partial x_j}{\partial A_k} \propto \frac{\partial x_j}{\partial \theta}. \tag{B.36}$$

Proof. Defining G as in the proof of Lemma 1 and again invoking the Implicit Function Theorem, we have

$$\begin{aligned}
\frac{\partial G_j}{\partial A_k} &= -\frac{\tilde{P}_j}{\sigma A_k} \\
\frac{\partial G_j}{\partial A_l} &= \frac{\tilde{P}_j}{\sigma A_l}.
\end{aligned}$$

As before, this implies $\frac{\partial x_j}{\partial A_l} \propto \frac{\partial x_j}{\partial A_k} \propto \frac{\partial x_j}{\partial \theta}$ as desired. □

To consider the desirability of intervening in the direction of innovation, we will start by

taking the first-order conditions with respect to scientists, which follow

$$\epsilon_k \ln(\upsilon)\varsigma s_k^{\varsigma-1} = \varpi$$
$$\epsilon_l \ln(\upsilon)\varsigma s_l^{\varsigma-1} = \varpi,$$

where ϵ_k and ϵ_l are the Lagrange multipliers on the laws of motion for technology, and ϖ is the Lagrange multiplier on the supply of scientists. This implies that the Planner can implement their desired innovation allocation by setting innovation subsidies according to $\frac{\xi_l}{\xi_k} = \frac{\epsilon_l/\Pi_l}{\epsilon_k/\Pi_k}$. Therefore, the Planner will want to intervene whenever this ratio differs from one because that would mean private profit signals do not reflect the true social benefit of innovation.

Taking the first-order conditions with respect to technology, we have

$$\epsilon_k = \mu_1 S_k + A_k \underbrace{\frac{\partial \mathcal{L}}{\partial X}\frac{\partial X}{\partial A_k}}_{=0}$$
$$\epsilon_l = \mu_1(1 - S_k) + A_l \underbrace{\frac{\partial \mathcal{L}}{\partial X}\frac{\partial X}{\partial A_l}}_{=0}.$$

That is, changing technology has two effects. First, from Hulten's Theorem, it expands output according to its income share (Hulten, 1978). Second, it change automation thresholds, but because of Lemma B.1, this effect envelopes out when the threshold rule is chosen optimally.

Therefore, the social value of innovation is proportional to factor shares, implying that market size effects generate an efficient allocation of research effort. This means that, with the threshold rule set optimally, the Planner does not need to engage in addition intervention in the direction of innovation.

We can ask an analogous question about the capital tax: in the absence of the threshold rule, does selecting an optimal capital tax imply that the equilibrium direction of innovation is efficient? The answer is no. In fact, the opposite is the case. To see why, we can can see from examining the proofs of Lemmas 1 and B.1 that

$$\frac{\partial x_j}{\partial A_k} = \frac{K}{A_k}\frac{\partial x_j}{\partial K}$$
$$\frac{\partial x_j}{\partial A_l} = -\frac{K}{A_l}\frac{\partial x_j}{\partial K}.$$

Following the same reasoning as in the proof of Proposition 3, we have that

$$\tau_k = -\frac{1}{\mu_1 R}\frac{\partial \mathcal{L}}{\partial X}\frac{\partial X}{\partial K}.$$

This implies that the first-order conditions for technology become

$$\epsilon_k = \mu_1 S_k - \mu_1 RK\tau_k$$
$$\epsilon_l = \mu_1(1 - S_k) + \mu_1 RK\tau_k,$$

which implies that relative innovation subsidies follow

$$\frac{\xi_l}{\xi_k} = \frac{1 + \frac{S_k}{1-S_k} Y\tau_k}{1 - Y\tau_k}. \tag{B.37}$$

This implies that the optimal degree of intervention in the direction of innovation *increases* with the capital tax; quite the opposite of an optimal capital tax allowing for laissez-faire innovation. Namely, because changes in automation from changing technology are also proportional to those from changing capital, we have that influencing innovation is another avenue through which the Planner can address the displacement effect of automation. And the more it is desirable to tax capital to reduce automation, the more it will be desirable to push the direction of innovation in favor of labor-augmenting technology.

B.6 Implementation of Threshold Rule with Fixed Costs

In this appendix, I show how the threshold rule can be implemented via a tax (or subsidy) on fixed cost for automation. In particular, I will assume that automating up to the threshold task requires the payment of convex fixed costs $\mathcal{F}_j(x)Y$. Justifying payment of these fixed costs requires market power, so I will assume a markup structure given by the following

$$\upsilon(x) = \max\big[\frac{w_j}{r}\frac{A_k\alpha_k(x)}{A_j\alpha_j(x)}, \overline{\upsilon}\big]. \tag{B.38}$$

That is, those who pay the fixed cost to automate a task are granted an exclusive right to produce that task with capital. They then limit price to labor or charge an exogenously given maximum markup.

For policy, I will consider a tax or subsidy that scales the fixed costs by ξ_j. In this setting, the marginal task will satisfy

$$(\upsilon(x_j) - 1)\frac{r}{A_k\alpha_k(x_j)} y(x_j) = \xi_j \mathcal{F}_j'(x_j). \tag{B.39}$$

In words, the threshold of automation will be the point where the profit from automating is equal to the marginal increase in fixed costs.

The question then is whether the Planner can select ξ_j so as to replicate a threshold rule of a given size θ. The answer is yes, and the formula for doing so will depend on whether the

upper bound for the markup is binding. If the upper bound is binding, ξ_j must follow

$$\xi_j = \frac{\overline{\upsilon} - 1}{\overline{\upsilon}^{\sigma}} \Big(\frac{w_j}{A_j \alpha_j(x_j)} \Big)^{1-\sigma} \frac{1}{(1+\theta)^{1-\sigma} \mathcal{F}_j'(x_j)}.$$

If the upper bound is not binding, this implies a tight connection between the markup and the threshold rule: $\upsilon(x_j) = (1+\theta)$. In that case, ξ_j must follow

$$\xi_j = \Big(\frac{r}{A_k \alpha_k(x_j)} \Big)^{1-\sigma} \frac{\theta}{(1+\theta)^{\sigma} \mathcal{F}_j'(x_j)}.$$

In either case, the Planner can implement a given threshold rule by modifying the fixed cost schedule for automation.

C Details on Calibration & Simulation

C.1 Description of Data Sources

This appendix provides greater detail on the datasets that I use for my quantitative analysis. To derive occupation-level statistics, I use household-level data from both the 2016 ACS and 1980 Census. Nominal wages in the 1980 Census are converted to 2016 dollars using the CPI. The occupational scheme in both of these datasets is occ2010, which I map to Deming (2017)'s occ1990dd scheme using a crosswalk from Webb (2020). I only include those aged 18 to 65 who worked last year and drop those who reported no earnings. Taking the intersection of availability across the two datasets, I end up with $J = 300$ occupations.

My calculation of $\{n_j, l_j, \tilde{w}_j\}$ is as follows.[C.1] I will index people by ι, and with a slight abuse of notation, I will say a person is in occupation j whenever $\iota \in j$. Denote person weight by $perwt(\iota)$, number of weeks worked last year by $wkswork(\iota)$, usual hours worked in a week by $uhrswork(\iota)$, and pretax wage and salary income by $incwage(\iota)$. Number of weeks worked is reported in intervals (e.g. 27 to 39 weeks), so I follow Acemoglu and Autor (2011) and assign a person the midpoint of their reported interval. For the number of people in each occupation, I compute

$$n_j = \frac{\sum_{\iota \in j} perwt(\iota)}{\sum_i \sum_{\iota \in i} perwt(\iota)}. \tag{C.1}$$

This is also the proportion of the population in a given occupation. For labor supply, I compute

$$l_j = \frac{\sum_{\iota \in j} perwt(\iota) \cdot wkswork(\iota) \cdot uhrswork(\iota)}{\sum_{\iota \in j} perwt(\iota)}. \tag{C.2}$$

[C.1] I use tilde notation to distinguish between observed wages $\tilde{w}_j$ in 2016 dollars and model unit wages w_j. These will only differ by a factor of proportionality.

This gives us a measure of the average hours worked in an occupation. For wages, I compute

$$\tilde{w}_j = \frac{\sum_{\iota \in j} perwt(\iota) \frac{incwage(\iota)}{wkswork(\iota) \cdot uhrswork(\iota)}}{\sum_{\iota \in j} perwt(\iota)}. \tag{C.3}$$

This gives us a measure of the average hourly wage in an occupation. Finally, I compute labor income shares as

$$S_j = (1 - S_k) \frac{\tilde{w}_j n_j l_j}{\sum_i \tilde{w}_i n_i l_i}. \tag{C.4}$$

To compute labor income and wealth shares by decile from the 2016 SCF, I make use of the variables weight, net worth, and wage income. I then separately compute decile cutoffs for both labor income and wealth using the weight variable to define the distribution. Finally, I compute each decile's share of total labor income or wealth. The share of wealth owned by the top decile is 77.1%, which is very similar to the 77.2% reported by Saez and Zucman (2016) for the year 2012.

C.2 Calibration Moments

This appendix describes my calibration procedure in greater detail. I will start be describing the breakdown of income in the initial period. Throughout, I normalize production in the second period to $Y = 100$. For the initial endowment, I set

$$\bar{Y} = Y_0 + (1 - \delta) K_0 = \frac{1 + (1 - \delta)\mathcal{K}}{1 + g} Y. \tag{C.5}$$

That is, the initial endowment is equal to initial output and capital. Division by $1 + g$, where I set $g = 0.02$, reflects TFP growth across the two periods. I assume that income shares are the same in each period. Wealth shares follow Equation (33), so ownership of the initial endowment follows

$$n_j \omega_j \bar{Y} = \frac{S_j + S_j^{\kappa} \left[S_k + (1 - \delta)\mathcal{K} \right]}{1 + g} Y. \tag{C.6}$$

Next, I directly observe labor supply and income shares, so I can set factor prices as $w_j = S_j Y / L_j$ and $r = S_k / \mathcal{K}$. Note that the status quo tax function described in Equation (34) applies to households' budgets in both periods, with income allocated across capital- and labor-based sources in the initial period according to the above discussion. The scale parameter of the labor tax Ψ is the solution to

$$\mathcal{G} = (1 - S_k) + \tau_k (r - \delta)\mathcal{K} - \Psi \frac{\sum_j n_j (w_j l_j)^{1 - \psi}}{Y}, \tag{C.7}$$

where $\mathcal{G} = 0.3$ is the share of output raised as tax revenue.[C.2]

[C.2] I modify Ψ in the initial period to maintain the same government share of income before income growth

I set consumption $\{c_{j,0}, c_{j,1}\}$ and the household discount factor β using a modified version of my two period model, where the second period is interpreted as a permanent steady-state. Specifically, households solve the problem

$$\begin{aligned} \max_{c_{j,0},c_{j,1},l_j,\kappa_j} u(c_{j,0}) + \beta \sum_{t\geq 0} \beta^t \big[u((1+g)^t c_{j,1}) - h_j(l_j)\big] \;\; s.t. \\ c_{j,0} + \kappa_j = \omega_j \bar{Y} \\ c_{j,1} + (1+g)\kappa_j = \Psi(w_j l_j)^{1-\psi} + D + \big(1 + (1-\tau_k)(r-\delta)\big)\kappa_j. \end{aligned} \tag{C.8}$$

That is, it is as if households are committing to a consumption, savings, and labor supply profile that can be maintained indefinitely in a steady-state with income growth g. This is isomorphic to a two period problem with effective discount factor $\beta/(1-\beta(1+g)^{1-\vartheta})$ and depreciation rate $1+\delta+g$. Consumption profiles and the discount factor then follow from satisfaction of the household budget constraints (20) and Euler equations (22) when the the capital-output ratio is constant in both the initial period and stead-state. Finally, I set the scale parameters of labor disutility φ_j to exactly match the labor supply condition (5). Thus, $\varphi_j l_j^{\psi+1/\varepsilon} c_{j,1}^{\vartheta} = \Psi(1-\psi) w_j^{1-\psi}$.

The remaining parameters to calibrate are those describing the productivity schedules of the economy: $\{\Gamma, A_k, \{\nu_j, A_j\}\}$. From the parametric assumptions for task-level productivity described in Equations (35) and (36), we have that

$$\Lambda_j = \frac{\overline{x}^{\zeta_j \nu_j (\sigma-1)+1} - x_j^{\zeta_j \nu_j (\sigma-1)+1}}{\zeta_j \nu_j (\sigma-1)+1} \tag{C.9}$$

$$\Lambda_k = \sum_j \frac{x_j^{\zeta_j (\nu_j - 1)(\sigma-1)+1}}{\zeta_j (\nu_j - 1)(\sigma-1)+1}, \tag{C.10}$$

and for labor-augmenting productivity we have

$$A_j = \frac{w_j}{x_j^{\zeta_j}} \frac{A_k}{r}. \tag{C.11}$$

Next, matching the capital share with my assumed capital-output ratio implies

$$A_k = r^{\frac{\sigma}{\sigma-1}} \Big(\frac{\mathcal{K}}{\Lambda_k}\Big)^{\frac{1}{\sigma-1}}. \tag{C.12}$$

To match labor income shares, absolute advantage ν_j must satisfy

$$\Lambda_j = \Lambda_k x_j^{\zeta_j(\sigma-1)} \frac{S_j}{S_k}, \tag{C.13}$$

across periods. That is, I set $\Psi_0 = \Psi(1+g)^{-\psi}$. Without this adjustment, pure income growth would scale up the share of income collected as tax revenue.

which required plugging in Equations (C.11) and (C.12). Note that Equation (C.13) only depends on data as well as absolute and comparative advantage parameters. Therefore, for a given guess of Γ, we can solve out $\{A_k, \{\nu_j, A_j\}\}$. This gives us a one-dimensional root-solving problem as the outer loop, where Γ must match the aggregate elasticity of substitution Σ_k according to Equation (19).

To unpack Equation (19) in the context of my parametric task model, we have that

$$\frac{\partial \ln(w_j)}{\partial \ln(r)} = -\frac{1}{\sigma}\frac{\alpha_j(x_j)^{\sigma-1}x_j}{\zeta_j \Lambda_j}\Big(\frac{\partial \ln(w_j)}{\partial \ln(r)} - 1\Big) + \frac{1}{\sigma}\Sigma_k + \frac{1}{\sigma}\Sigma_j\Big(\frac{\partial \ln(w_j)}{\partial \ln(r)} - 1\Big).$$

Plugging in Equation (18) and solving gives us

$$\frac{\partial \ln(w_j)}{\partial \ln(r)} = 1 + \frac{\sigma - \Sigma_k}{\alpha_k(x_j)^{\sigma-1}x_j/\zeta_j\Lambda_k}.$$

Therefore, Γ is chosen to satisfy

$$\Sigma_k = \sum_j \Big(\frac{S_j(\Sigma_k - \sigma)}{\alpha_k(x_j)^{\sigma-1}x_j/\zeta_j\Lambda_k}\Big)\Sigma_j. \tag{C.14}$$

For the introduction of AI in Section 6.2, I use the same calibrated value of Γ but change the exposure score χ_j to be the average across robots, software, and AI. Given the resulting array of comparative advantage ζ_j, I solve out $\{A_k, \{\nu_j, A_j\}\}$ using the same procedure as above. This keeps all equilibrium outcomes the same; the only thing that now differs is the counterfactual change in automation thresholds in response to price, productivity, or policy changes.

For Figure 4, I vary capital-augmenting productivity growth g_k, so that capital's productivity equals $A_k' = (1+g_k)A_k$. I also allow labor to achieve catchup growth at a fixed fraction of the growth of capital. Finally, I adjust the scale of the tax function according to $\Psi' = \Psi(1+g_y)^{\psi}$, where $g_y = S_k g_k + (1-S_k)g_l$ is the first-order change in output. This adjustment prevents the progressivity of the status quo tax function from mechanically increasing tax rates in response to income growth.

I validate my calibration procedure by comparing my model's implications to the empirical results of Acemoglu and Restrepo (2022). They estimate the role of automation in explaining the changes in the US wage structure from 1980 to 2016 using an empirical analog of the total differential of log wages:

$$d\ln(w_j) = \frac{1}{\sigma}d\ln(\Lambda_j) + \frac{\sigma-1}{\sigma}d\ln(A_j) + \frac{1}{\sigma}d\ln(Y) - \frac{1}{\sigma}d\ln(L_j)$$

In particular, they focus on the term $d\ln(\Lambda_j)$ by measuring the loss of routine tasks for 500 demographic groups over this period. They then regress a long difference in log wages on their

measure of task displacement to determine the role of automation in explaining the heterogeneous wage gains in this period. Acemoglu and Restrepo (2022) weight their regressions by total hours worked, so I make use of labor supply weights $L_j / \sum_i L_i$ in all of the comparisons discussed below.

I will make use of this evidence to validate my calibration strategy in two ways. First, to justify setting automation thresholds at 20% of the total task interval – i.e. $x_j = \varkappa \overline{x}$, where $\varkappa = 0.2$ – I compare their point estimate of the impact of task displacement on log wages with the analog in my model. To be more specific, I take $-d\ln(\overline{x} - x_j)$ as the model analog of their measure of task displacement, and I interpret their point estimate as describing the average log wage reduction from this loss in tasks. Unpacking the $d\ln(\Lambda_j)$ term in my model implies that their point estimate maps to the average value of $-\alpha_j(x_j)^{\sigma-1}(\overline{x} - x_j)/(\sigma\Lambda_j)$. The estimates from their baseline specification range from -1.3 to -1.9, but they also find estimates around -3 in several of their robustness exercises. Therefore, my model-implied estimate of -3.02 is high but still within the range of empirically validated estimates.

Second, Acemoglu and Restrepo (2022) emphasize that their measure of task displacement can explain between 50% to 70% of the unequal wage gains over this period. To compare my model with this finding, I regress observed changes in log wages $\Delta\ln(\tilde{w}_j) \equiv \ln(\tilde{w}_{j,16}) - \ln(\tilde{w}_{j,80})$ on the component of $d\ln(\Lambda_j)$ that is driven by reductions in the user cost of capital. This corresponds to Acemoglu and Restrepo (2022)'s preferred measure of task displacement, which predicts industry-level reductions in the labor share from several proxies of the adoption of automation technologies. Formally, I take the R-squared from the regression

$$\Delta\ln(\tilde{w}_j) = \Upsilon_0 + \Upsilon_1 \frac{\alpha_j(x_j)^{\sigma-1}x_j}{\zeta_j\Lambda_j} + \epsilon_j, \tag{C.15}$$

where the multiplicative component of task displacement, such as the decline in the user cost of capital $(d\ln(r) - d\ln(A_k))$, is absorbed by the coefficient and so not relevant in explaining variation.[C.3] I estimate an R-squared of 12.6%, implying that my calibration strategy takes a conservative stance on the role of automation in explaining wage inequality.

C.3 Computing Threshold Rule Perturbations

Implementing Proposition 2 quantitatively requires me to compute the changes in equilibrium objects from perturbations in the threshold rule $d\theta$. To accomplish this, I will consider a function $H(\{c_{j0}, c_{j1}, l_j\}; \theta)$ that achieves a root when the equilibrium conditions of Definition 1 are satisfied. By the Implicit Function Theorem, we have that

[C.3] Equation (C.15) only makes use of first-order changes, but this is the empirical strategy taken by Acemoglu and Restrepo (2022) as well. Thus, second-order effects will show up in the residual in both cases.

$$\left(\frac{\partial[c_{j0}, c_{j1}, l_j]}{\partial\theta}\right) = -\left(\frac{\partial H}{\partial[c_{j0}, c_{j1}, l_j]}\right)^{-1}\left(\frac{\partial H}{\partial\theta}\right). \tag{C.16}$$

From this, we have changes in the capital supply as well: $dK = \sum_j n_j d\kappa_j = -\sum_j n_j dc_{j0}$. Then for factor prices and output we have

$$d\ln(w_j) = -\frac{1}{\sigma}\frac{\alpha_j(x_j)^{\sigma-1}}{\Lambda_j}dx_j + \frac{1}{\sigma}d\ln(Y) - \frac{1}{\sigma}d\ln(L_j) \tag{C.17}$$

$$d\ln(r) = \frac{1}{\sigma}\sum_i \frac{\alpha_k(x_i)^{\sigma-1}}{\Lambda_k}dx_i + \frac{1}{\sigma}d\ln(Y) - \frac{1}{\sigma}d\ln(K) \tag{C.18}$$

$$d\ln(Y) = \sum_j \left(\frac{\partial\ln(Y)}{\partial x_j}dx_j + S_j d\ln(L_j)\right) + S_k d\ln(K). \tag{C.19}$$

Changes in the automation thresholds dx_j can be computed in the manner described in Appendix A.3.

C.4 Non-Linear Tax Solution Algorithm

To solve numerically for the Mirrlees allocations, I run the following algorithm. Holding factor prices fixed, the inner problem solves for an incentive compatible allocation. Holding consumption and labor fixed, the outer problem updates the threshold rule and automation thresholds. I find a fixed point with the following iterative procedure:

1. For a given array of factor prices, pick consumption and labor to solve

$$\max_{\{c_{j,0}, c_{j,1}, l_j\}} \mathcal{W}(\{n_j V_j\}) - \Delta\sum_{i,j}\min\left[u(c_{i,0}) + \beta(u(c_{i,1}) - h_i(l_i)) - u(c_{j,0}) - \beta(u(c_{j,1}) - h_i(w_j l_j/w_i)), 0\right]^2$$
$$s.t.\ \ (7), (8).$$

   That is, rather than imposing the IC constraints directly, I impose them as a penalty with cost parameter Δ. This cost parameter is scaled up until the worst IC violation is no more than 1/1000 of welfare. Note that with fixed factor prices, I assume a linear production technology in the resource constraint.

2. Holding automation thresholds fixed, update factor prices with the new factor supplies of capital and labor.

3. Holding fixed factor prices, the allocation of consumption and labor, and automation thresholds, update the threshold rule to satisfy the perturbation formula of Proposition 2. This requires computing the second derivative of the tax function, so I approximate the global shape of the tax function with a fourth degree polynomial. Furthermore, I assume

that initial endowments are fully taxed and redistributed, so all initial endowments are equal to $\bar{Y}$.

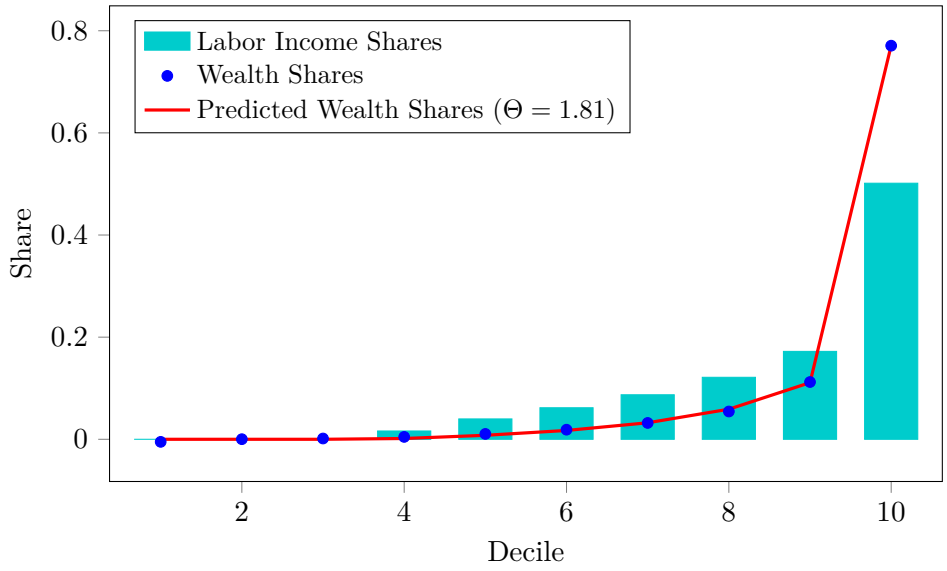
4. Update automation thresholds using factor prices and the threshold rule.

5. Check if the new threshold rule and automation thresholds are equal to their previous value. If so, this fixed point is a solution to the Mirrlees problem. If not, update factor prices using the new automation thresholds and return to step 1.

This algorithm can also find the allocation when the threshold rule is unavailable by simply skipping step 3 and always setting $\theta = 0$.

D Additional Tables & Figures

Figure D.1: Shares by Decile



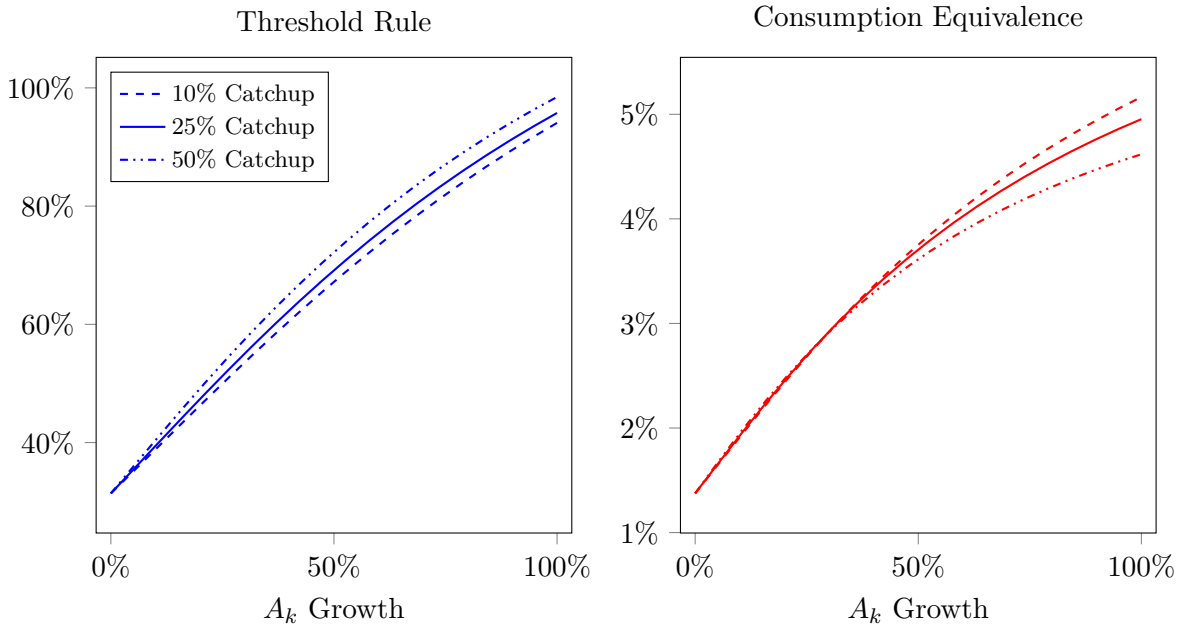
Notes: Match of model predictions to the data for wealth shares by decile. Model predictions come from Equation (33). Data on labor income and wealth shares by decile come from the 2016 SCF.

Table D.1: Correlation of Wages and Exposure Measures

	Wages	Software Exposure	Robot Exposure	AI Exposure
Wages	1.0			
Software Exposure	-0.06	1.0		
Robot Exposure	-0.36	0.63	1.0	
AI Exposure	0.26	0.65	0.23	1.0

Notes: Exposure scores are percentiles taken from Webb (2020), and wages are from the 2016 ACS. Both variables are at the occupation level.

Figure D.2: AI Introduction with Variable Labor Catchup



Notes: This graph reproduces Figure 4 with alternative assumptions about the catchup growth fraction for labor-augmenting productivity.

References

Acemoglu, D. and Autor, D. (2011). Skills, Tasks, and Technologies: Implications for Employment and Earnings. In *Handbook of Labor Economics*, volume 4, pages 1043–1171. Elsevier.

Acemoglu, D. and Restrepo, P. (2022). Tasks, Automation, and the Rise in U.S. Wage Inequality. *Econometrica*, 90(5):1973–2016.

Atkinson, A. and Stiglitz, J. (1976). The Design of Tax Structure: Direct Versus Indirect Taxation. *Journal of Public Economics*, 6(1-2):55–75.

Caliendo, L., Dvorkin, M., and Parro, F. (2019). Trade and Labor Market Dynamics: General Equilibrium Analysis of the China Trade Shock. *Econometrica*, 87(3):741–835.

Costinot, A. and Werning, I. (2023). Robots, Trade, and Luddism: A Sufficient Statistic Approach to Optimal Technology Regulation. *The Review of Economic Studies*, 90(5):2261–2291.

Deming, D. (2017). The Growing Importance of Social Skills in the Labor Market. *The Quarterly Journal of Economics*, 132(4):1593–1640.

Diamond, P. and Mirrlees, J. (1971a). Optimal Taxation and Public Production I: Production Efficiency. *The American Economic Review*, 61(1):8–27.

Diamond, P. and Mirrlees, J. (1971b). Optimal Taxation and Public Production II: Tax Rules. *The American Economic Review*, 61(3):261–278.

Donald, E. (2023). Spillovers and the Direction of Innovation: An Application to the Clean Energy Transition. *Working Paper*.

Donald, E., Fukui, M., and Miyauchi, Y. (2025a). Optimal Dynamic Spatial Policy. Working Paper 34290, National Bureau of Economic Research.

Donald, E., Fukui, M., and Miyauchi, Y. (2025b). Unpacking Aggregate Welfare in a Spatial Economy. Working Paper 34075, National Bureau of Economic Research.

Eden, M. (2020). Aggregating Welfare Gains. Discussion Paper DP14783, Centre for Economic Policy Research.

Hofbauer, J. and Sandholm, W. (2002). On the Global Convergence of Stochastic Fictitious Play. *Econometrica*, 70(6):2265–2294.

Hulten, C. (1978). Growth Accounting with Intermediate Inputs. *The Review of Economic Studies*, 45(3):511–518.

Liu, E. and Ma, S. (2021). Innovation Networks and R&D Allocation. Working Paper 29607, National Bureau of Economic Research.

Saez, E. and Zucman, G. (2016). Wealth Inequality in the United States since 1913: Evidence from Capitalized Income Tax Data. *The Quarterly Journal of Economics*, 131(2):519–578.

Webb, M. (2020). The Impact of Artificial Intelligence on the Labor Market. *Working Paper*.