# Should Robots Be Taxed?*

Joao Guerreiro[†], Sergio Rebelo[‡], and Pedro Teles[§]

January 2019

**Abstract**

We use a model of automation to show that with the current U.S. tax system, a fall in automation costs could lead to a massive rise in income inequality. This inequality can be reduced by making the current income-tax system more progressive and by taxing robots. But this solution involves a substantial efficiency loss. A Mirrleesian optimal income tax can reduce inequality at a smaller efficiency cost. An alternative approach is to amend the current tax system to include a lump-sum rebate. With the rebate in place, it is optimal to tax robots as long as there is partial automation.

J.E.L. Classification: H21, O33
Keywords: inequality, optimal taxation, automation, robots.

*We thank Gadi Barlevy, V.V. Chari, Bas Jacobs, and Nir Jaimovich for their comments. Teles thanks the support of FCT as well as the ADEMU project, "A Dynamic Economic and Monetary Union," funded by the European Union's Horizon 2020 Program under grant agreement N° 649396.

[†]Northwestern University.

[‡]Northwestern University, NBER and CEPR.

[§]Católica-Lisbon School of Business & Economics, Banco de Portugal and CEPR

# 1 Introduction

The American writer Kurt Vonnegut began his career in the public relations division of General Electric. One day, he saw a new milling machine operated by a punch-card computer outperform the company's best machinists. This experience inspired him to write a novel called "Player Piano." It describes a world in which school children take a test at an early age that determines their fate. Those who pass, become engineers and design robots used in production. Those who fail, have no jobs and live from government transfers. Are we converging to this dystopian world? How should public policy respond to the impact of automation on the demand for labor?

These questions have been debated ever since 19th-century textile workers in the U.K. smashed the machines that eliminated their jobs. As the pace of automation quickens and affects a wide range of economic activities, Bill Gates re-ignited this debate by proposing that robots should be taxed. Policies that address the impact of automation on the labor market have been discussed in the European Parliament and have been implemented in South Korea.

In this paper, we use a simple model of automation to compare the equilibrium that emerges under the current U.S. tax system (which we call the status quo), the first-best solution to a planner's problem without information constraints, and the second-best solutions associated with different configurations of the tax system.

Our model has two types of workers which we call routine and non-routine. Routine workers perform tasks that can be automated by using intermediate inputs that we refer to as robots.[1] We find that robot taxes are optimal as long as there is partial automation. These taxes increase the wages of routine workers, and decrease those of non-routine workers, giving the government an additional instrument to reduce income inequality. With full automation, it is not optimal to tax robots. Rou-

[1] See Acemoglu and Autor (2011) and Cortes, Jaimovich and Siu (2017) for a discussion of the impact of automation on the labor market for routine workers.

tine workers no longer work, so taxing robots distorts production decisions without reducing income inequality.[2]

We model the current U.S. tax system using the after-tax income function proposed by Feldstein (1969), Persson (1983), and Benabou (2000) and estimated by Heathcote, Storesletten and Violante (2017). With this tax system, as the cost of automation falls, the wages of non-routine workers rise while the wages of routine workers fall to make them competitive with robot use. The result is a large rise in income inequality and a substantial decline in the welfare of routine workers.

The level of social welfare obtained in the status quo is much worse than that achieved in the first-best solution to an utilitarian social planner problem without information constraints. But this first-best solution cannot be implemented when the government does not observe the worker type. The reason is that the two types of workers receive the same level of consumption but non-routine workers supply more labor than routine workers. As a result, non-routine workers have an incentive to act as routine workers and receive their bundle of consumption and hours worked.

To circumvent this problem, we solve for the optimal tax system imposing, as in Mirrlees (1971), the constraint that the government does not observe the worker type or the workers' labor input. The government can observe total income and consumption of the two types of workers, as well as the use of robots by firms. We assume that taxes on robots are linear for the reasons emphasized in Guesnerie (1995): non-linear taxes on intermediate inputs are difficult to implement in practice because they create arbitrage opportunities. A Mirrleesian optimal tax system can improve welfare relative to the status quo. In fact, it can yield a level of welfare that is close to that of

[2]These results show that the reason why it can be optimal to tax robots in our model differs from the rationale used by Bill Gates to motivate robot taxation. Gates argued that robots should be taxed to replace the tax revenue that the government collected from routine workers before their jobs were automated. In our model, when there is full automation the government collects no tax revenue from routine workers yet it is optimal not to tax robots.

the first-best allocation.

We also study the optimal policy when the tax schedule is constrained to take a simple, exogenous form. We consider the income tax schedule proposed by Heathcote, Storesletten and Violante (2017) and linear robot taxes. We compute the parameters of the income tax function and the robot tax rate that maximize social welfare. We find that income inequality can be reduced by raising marginal tax rates and taxing robots. Tax rates on robot use can be as high as 30 percent. Routine workers supply a constant number of hours over time even though their wages fall. This solution yields poor outcomes in terms of efficiency and distribution.

We consider a modification of the Heathcote, Storesletten and Violante (2017) tax schedule that allows for lump-sum transfers that ensure that all workers receive a minimum income. We find that this modification improves both efficiency and distribution relative to a tax system without transfers. Hours worked by routine and non-routine workers diverge over time. Full automation occurs in finite time, so hours worked by routine workers fall to zero. Once full automation occurs, routine workers pay no income taxes and the tax system can be designed so that the labor-supply decisions of non-routine workers are not distorted. The economy with full automation resembles the world of "Player Piano." Only non-routine workers have jobs. Routine workers live off government transfers and, despite losing their jobs, are better off than in the status quo. The fact that full automation occurs in finite time reflects the rudimentary nature of the tax system available to the government. When the government has access to a more flexible non-linear tax schedule, as in the Mirrleesian solution, full automation occurs only asymptotically.

One might expect optimal robot taxation to follow from well-known principles of optimal taxation in the public finance literature. We know from the intermediate-goods theorem of Diamond and Mirrlees (1971) that it is not optimal to distort production decisions by taxing intermediate goods. Since robots are in essence an intermediate good, taxing them should not be optimal.

The intermediate-good theorem relies on the assumption that "net trades" of different goods can be taxed at different rates. In our context, this assumption means that the government can use different tax schedules for routine and non-routine workers. We study two environments where there are limits to the government's ability to tax different workers at different rates, Mirrlees (1971)-type information constraints and a simple exogenous tax system common to both types of workers. We find that it is optimal to tax robots in both environments.

This finding seems to contradict the key result in Atkinson and Stiglitz (1976). These authors show that when the income tax system is non-linear it is not optimal to distort production decisions by taxing intermediate goods. But, as stressed by Naito (1999), Scheuer (2014), and Jacobs (2015), Atkinson and Stiglitz (1976)'s result depends critically on the assumption that workers with different productivities are perfect substitutes in production. This assumption does not hold in our model. Taxing robots can be optimal because it affects relative productivities, loosening the incentive constraint of non-routine workers.

We extend our model to allow workers to switch their occupations by paying a cost. In the first-best solution, workers who have a low cost of becoming non-routine workers do so. Those with a high cost become routine workers. In the Mirrlees solution to the model with occupational choice, it is optimal to use robot taxes to loosen the incentive constraint of non-routine workers. The planner can use the income tax schedule to redistribute income or to induce more agents to become non-routine workers. When the cost of becoming non-routine are high (low), the planner resorts more (less) to using the income tax schedule to redistribute income.

We generalize our static model to a dynamic setting in which robots are an investment good. The properties of the Mirleesian solution of the dynamic model are similar to those of the static model. It is optimal to tax robots to loosen the incentive constraint of non-routine workers. The levels of taxation are similar to those of the static model. The tax rate on robots converges to zero as the degree of automation converges to one.

The paper is organized as follows. In Section 2, we describe our static model of automation. Subsection 2.1 describes the status-quo equilibrium, i.e. the equilibrium under the current U.S. income tax system and no robot taxes. Subsection 2.2 describes the first-best solution to the problem of an utilitarian planner. In subsection 2.3, we analyze a Mirrleesian second-best solution to the planner's problem. In subsection 2.4, we study numerically the optimal tax system that emerges when income taxes are constrained to take the functional form proposed by Heathcote, Storesletten and Violante (2017) both with and without lump-sum rebates. In subsection 2.5, we compare the implications of different policies for social welfare and for the utility of different agents. Subsection 2.6 discusses the model with endogenous occupation choice. In Section 3, we analyze a dynamic model of automation. Section 4 relates our findings to classical results on production efficiency and capital taxation in the public finance literature. Section 5 concludes. To streamline the main text, we relegate the more technical proofs to the appendix.

**Related literature** ++++++++++++++++++

Missing literature

++++++++++++++++++

In section 3, we enrich the analysis of the static model to a more complex dynamic overlapping-generations model. This model features several important characteristics. First, in the dynamic model, robots are an investment good which can be accumulated over time. Second, we consider technological progress in the form of investment-specific technical change as in Greenwood, Hercowitz, and Krusell (1997). Because robots are more substitutable with routine workers than non-routine workers, the technical change is skill biased which impacts the routine-skill premium.[3] Technological advances lead to rising productivity, but increase pretax wage

[3]This is analogous to the skill-premium effects of capital-skill complementarity in Krusell, Oha-

inequality among occupations. For that reason, a good description of the long-run consequences of technological transitions requires an explicit model of skill acquisition by workers. Our model endogeneizes the choice of skills by adopting an overlapping-generations framework in which workers can choose their skills upon entering the economy. In that feature, the model is related to a large literature studying the effects of technology-specific human capital for the diffusion of technologies, e.g. Chari and Hopenhayn (1991), Caselli (1999), Adao, Beraja, and Pandalai-Nayar (2019), among others.

In this model, taxing robots affects the relative wages of routine and non-routine workers thereby affecting human capital decisions and occupational choice. In that sense, our analysis is related to Saez (2004), Scheuer (2014), Rothschild and Scheuer (2013), and Gomes, Lozachmeur and Pavan (2017). These authors characterize Mirrlees-style optimal tax plans in static models with endogenous occupation choice. Our approach is closest to Scheuer (2014) who considers a model in which an agent faces the choice of becoming a worker and an entrepreneur. Scheuer finds that, in the absense of differential taxation for these two occupations, the optimal plan may feature production distortions, much like the ones we find.

There is a large literature extending the Mirrlees (1971) approach to dynamic settings, e.g. Golosov, Kocherlakota, and Tsyvinski (2003), Werning (2007), Farhi and Werning (2013), among others.[4] Our analysis is closest to Slavík and Yazici (2014). These authors consider optimal Mirrleesian taxation in an infinite-horizon model with low- and high-skill workers and capital-skill complementarity. They find that it is optimal to tax equipment capital because it is complementary to high-skill workers and substitute to low skill. Furthermore, capital taxes are high initially and rise over time.

Our analysis departs from theirs in two main ways. First, because we focus on

nian, Rios-Rull, and Violante (2000).

[4]A review of the dynamic Mirrleesian approach can be found in Kocherlakota (2010).

the effects of automation as a technological transition, our model features technical progress over time. As automation advances, the productivity of routine workers falls. As a result, the optimal solution is such that routine workers eventually supply zero labor hours. This means that there is no role for robot taxes in affecting relative wages in the long run, so those taxes should be zero. We recover the celebrated Chamley-Judd result of zero long-run capital income taxes.[5] The reasons for positive long-run capital-income taxes in Slavík and Yazici (2014)[6] cease to be relevant once we take into account skill-biased technical progress.

Second, we also consider an overlapping–generations model with skill choice. In this model, taxation has a direct distribution effect on after-tax income and an indirect distribution effect on the choice of skills. As technology improves, the indirect redistribution channel becomes more important, i.e. the planner wants to induce a higher fraction of the population to acquire non-routine skills. Since taxes on robots are desirable only insofar as they improve the direct redistribution mechanism, the reduced relative importance of this form of redistribution also implies that taxes on robots should become zero.

**THIS USED TO BE IN THE RELATED LITERATURE AT THE END** The public-finance literature discusses several reasons why it might be optimal to tax capital, introducing intertemporal distortions.[7] First, it might be optimal to use intertemporal distortions to confiscate the initial stock of capital. Second, intertemporal distortions can be optimal when the elasticities of the marginal utility of consumption and labor are time varying. Third, intertemporal distortions can be used to provide insurance in models with idiosyncratic risk. All three reasons are absent in our model.

[5]See Chamley (1986) and Judd (1985).

[6]The same reasons for the optimality of positive long-run income taxes when there are restrictions to the taxation of different labor types are present in Jones, Manuelli, and Rossi (1997).

[7]For a recent overview of the literature on optimal capital taxation in a dynamic Ramsey setting see Chari, Nicolini and Teles (2018).

We consider Mirleesian taxes which allow for lump-sum taxation so there is no reason to confiscate the initial stock of capital. In addition, we assume that utility is separable in consumption and labor and the disutility of labor is isoelastic. Werning (2007) shows that under these conditions and with perfect substitutability of labor types, the optimal tax on capital is zero. Because our dynamic model abstracts from idiosyncratic risk, the reasons for capital taxation discussed in Golosov, Kocherlakota and Tsyvinski (2003) do not apply.

Our results are related to work by Slavík and Yazici (2014). These authors study optimal taxation in a model with two types of capital, structures and equipment, and no technical progress. They assume that equipment raises the marginal product of skilled workers relative to that of unskilled workers. In their set up, the optimal tax on equipment rises over time. In contrast, the optimal tax on robots in our model converges to zero. As discussed in Section 3, this property reflects the presence of technical progress in our model.

In sum, the classical results on production efficiency in the public finance literature depend on one of two key assumptions: (i) the government can tax differently every consumption good and labor type; or (ii) the environment is such that production distortions do not help in shaping incentives. Both assumptions fail in our model. On the one hand, the government cannot design the income tax system to independently target each type of worker. On the other hand, robots are substitutes for routine workers and complements to non-routine workers, so a tax on robots affects the ratio of the wages of these two types of workers.

## 2 A simple static model

We first discuss a simple model of automation that allows us to address the optimal tax policy questions posed in the introduction. The model has two types of households who draw utility from consumption of private and public goods and disutility from labor. One household type supplies routine labor and the other non-routine labor. The consumption good is produced with non-routine labor, routine labor, and robots. Robots and routine labor are used in a continuum of tasks. They are both perfect substitutes in performing these tasks.[8]

**Households** There is a continuum of unit measure of households. A mass $\pi_n$ of households is composed of non-routine workers while $\pi_r$ households are composed of routine workers. The index $j = n, r$, denotes the non-routine and routine labor type, respectively.

An household of type $j$ derives utility from consumption, $c_j$, and from the provision of a public good, $G$. The household also derives disutility from the hours of labor it supplies, $l_j$. Households have a unit of time per period, so $l_j \leq 1$. The household's utility function is given by

$$U_j = u(c_j, l_j) + v(G). \tag{1}$$

Denote by $u_x = \partial u(c,l)/\partial x$ where $x = c, l$ and $u_{xy} = \partial^2 u(c,l)/\partial x \partial y$. We assume that $u_c > 0$, $u_l < 0$, $u_{cc}$, $u_{ll} < 0$ and that consumption and leisure are normal goods: $u_{lc}/u_l - u_{cc}/u_c \geq 0$, and $u_{ll}/u_l - u_{cl}/u_c \geq 0$, where one of these conditions is a strict inequality. Furthermore, we assume that utility satisfies the single-crossing property, which is equivalent to assuming that $u_{ll}l/u_l + 1 - u_{cl}l/u_c > 0$. Finally, we assume that $v_G > 0$, $v_{GG} < 0$ and that $u(c,l)$ satisfies standard Inada conditions.

Household $j$ chooses $c_j$ and $l_j$ to maximize utility (1), subject to the budget constraint

$$c_j \leq w_j l_j - T(w_j l_j),$$

where $w_j$ denotes the wage rate received by the household type $j$ and $T(\cdot)$ denotes the income tax schedule.

[8]See Autor, Levy and Murmane (2003) for a study of the importance of tasks performed by routine workers in different industries and a discussion of the impact of automating these tasks on the demand for routine labor.

**Robot producers** The cost of producing a robot is the same across tasks and is equal to $\phi$ units of output. Robots are produced by competitive firms. A representative firm producing robots chooses $X$ to maximize profits $p_x X - \phi X$. It follows that in equilibrium $p_x = \phi$ and profits are zero.

**Final good producers** The representative producer of final goods hires non-routine labor ($N_n$), routine labor, and buys intermediate goods which we refer to as robots. Aggregate production follows a task framework which has become standard in the automation literature (Acemoglu and Restrepo, 2018a, 2018b). There is a unit continuum of tasks that can be performed by either routine labor or robots. Services provided by these tasks are denoted by $y_i$ for each $i \in [0, 1]$. The production function is given by

$$Y = A \left[ \int_0^1 y_i^{\frac{\rho-1}{\rho}} di \right]^{\frac{\rho}{\rho-1}(1-\alpha)} N_n^{\alpha}, \quad \alpha \in (0,1), \rho \in [0,\infty). \tag{2}$$

Each task can be produced with workers, $n_i$, or with robots, $x_i$

$$y_i = \begin{cases} \kappa_i x_i, & \text{if } i \text{ is automated} \\ \ell_i n_i, & \text{if } i \text{ is not automated} \end{cases} \tag{3}$$

The parameters $\kappa_i$ and $\ell_i$ capture the efficiency of robots and routine labor, respectively, in task $i$. We assume that $\kappa_i / \ell_i$ is weakly decreasing in $i$. This implies that routine workers are relatively more efficient in higher ordered tasks. Given this assumption, firms choose to automate the first tasks in that continuum. We denote by $m$ the level of automation, i.e. the number of tasks that are produced by robots, and write the production function as:

$$Y = A \left[ \int_0^m (\kappa_i x_i)^{\frac{\rho-1}{\rho}} di + \int_m^1 (\ell_i n_i)^{\frac{\rho-1}{\rho}} di \right]^{\frac{\rho}{\rho-1}(1-\alpha)} N_n^{\alpha}. \tag{4}$$

The problem of the firm is to maximize profits, $Y - w_n N_n - w_r \int_m^1 n_i di - (1 + \tau_x)\phi \int_0^m x_i di$, where $Y$ is given by equation (4). The variable $\tau_x$ is an ad-valorem tax rate on intermediate goods.

The optimal choices of $N_n$, $x_i$ for $i \in [0, m]$, $n_i$ for $i \in (m, 1]$ require that the following first-order conditions be satisfied:

$$w_n = \frac{\alpha Y}{N_n}, \tag{5}$$

$$(1+\tau_x)\phi = (1-\alpha)Y \left( \int_0^m (\kappa_s x_s)^{\frac{\rho-1}{\rho}} ds + \int_m^1 (\ell_s n_s)^{\frac{\rho-1}{\rho}} ds \right)^{-1} (\kappa_i)^{\frac{\rho-1}{\rho}} x_i^{-\frac{1}{\rho}} \tag{6}$$

$$w_r = (1-\alpha)Y \left( \int_0^m (\kappa_s x_s)^{\frac{\rho-1}{\rho}} ds + \int_m^1 (\ell_s n_s)^{\frac{\rho-1}{\rho}} ds \right)^{-1} (\ell_i)^{\frac{\rho-1}{\rho}} n_i^{-\frac{1}{\rho}}, \text{ for }. \tag{7}$$

To simplify, we make the additional assumption that $\kappa_i = \ell_i = 1$ for all $i$. This assumption lends tractability and clarity to the exposition of our main results. Section 3 relaxes this assumption in a dynamic quantitative model.

Under this assumption, it follows that it is optimal to use the same level of routine labor, $n_i$, in the $1 - m$ tasks that have not been automated and that the optimal use of robots is also the same in the $m$ automated tasks.

The optimal level of automation is $m = 0$ if $w_r < (1 + \tau_x)p_x$. The firm chooses to fully automate ($m = 1$) and to employ no routine workers if $w_r > (1 + \tau_x)p_x$. If $w_r = (1 + \tau_x)p_x$, the firm is indifferent between any level of automation $m \in [0, 1]$. In this case, equations (6) and (7) imply that the levels of routine labor and robots are the same across tasks,

$$m x_i = X, \text{ for } i \in [0, m], \text{ and } (1-m) n_i = N_r, \text{ for } i \in (m, 1], \tag{8}$$

where $N_r$ denotes total routine hours and $X$ denotes total robots. Using the fact that $x_i = n_j$, with interior automation we obtain $m = X/(N_r + X)$, and we can write the production function as $Y = A\,(X + N_r)^{1-\alpha}\,N_n^{\alpha}$.[9] Since the technology has constant

[9]Under the assumption that $\ell_i = \kappa_i = 1$, our task-based production function becomes exactly the same aggregate production function studied by Autor, Levy, and Murnane (2003).

returns to scale, profits are zero.

**Government** The government chooses taxes and the optimal level of government spending, subject to the budget constraint

$$G \leq \pi_r T(w_r l_r) + \pi_n T(w_n l_n) + \tau_x p_x \int_0^m x_i di. \tag{9}$$

**Equilibrium** An equilibrium is a set of allocations $\{c_r, l_r, c_n, l_n, G, N_r, X, x_i, n_i, m\}$, prices $\{w_r, w_n, p_x\}$, and a tax system $\{T(\cdot), \tau_x\}$ such that: (i) given prices and taxes, allocations solve the households' problem; (ii) given prices and taxes, allocations solve the firms' problem; (iii) the government budget constraint is satisfied; and (iv) markets clear.

The market clearing conditions for routine and non-routine labor are

$$N_j = \pi_j l_j, \quad j = r, n. \tag{10}$$

The market-clearing condition for the output market is

$$\pi_r c_r + \pi_n c_n + G \leq Y - \phi \int_0^m x_i di. \tag{11}$$

**The equilibrium with interior automation** In an equilibrium with automation ($m \in (0,1)$), the wage rate of routine workers equals the cost of robot use: $w_r = (1+\tau_x)\phi$. This condition implies that the number of robots used in each automated task equals the number of routine workers used in each non-automated task

$$\frac{X}{m} = \frac{\pi_r l_r}{1-m}.$$

Combining this equation with the firm's first-order condition (6), we obtain

$$(1+\tau_x)\phi = (1-\alpha)[X + \pi_r l_r]^{-\alpha}(\pi_n l_n)^{\alpha}. \tag{12}$$

Using this first-order condition and replacing $X = m\pi_r l_r/(1-m)$, we find that, given aggregate labor supplies $\pi_r l_r$ and $\pi_n l_n$, the equilibrium level of automation satisfies

$$m = 1 - \left[\frac{(1+\tau_x)\phi}{(1-\alpha)A}\right]^{1/\alpha} \frac{\pi_r l_r}{\pi_n l_n}, \tag{13}$$

Finally, using the other two first order conditions, we find that wages of both non-routine and routine labor are independent of preferences,

$$w_n = \alpha \frac{A^{1/\alpha}(1-\alpha)^{\frac{1-\alpha}{\alpha}}}{\left[(1+\tau_x)\phi\right]^{\frac{1-\alpha}{\alpha}}}, \tag{14}$$

$$w_r = (1+\tau_x)\phi. \tag{15}$$

The wage of routine workers is determined by the after-tax cost of robots. Because of constant returns to scale, the ratio of inputs is pinned down, and so is the wage of the non-routine worker. An increase in $\tau_x$ raises the wage of routine workers and lowers the wage rate of non-routine agents. Furthermore, production net of the cost of robots is also given by a linear formula

$$Y - \phi X = \pi_n w_n l_n \frac{\tau_x + \alpha}{\alpha(1+\tau_x)} + \frac{\pi_r w_r l_r}{1+\tau_x}. \tag{16}$$

It is also useful to note that in any equilibrium the income shares of total production are given by

$$\frac{w_r \pi_r l_r}{Y} = (1-\alpha)(1-m), \text{ and } \frac{w_n \pi_n l_n}{Y} = \alpha.$$

An increase in automation reduces the income share of routine workers and does not change the share of non-routine workers. In this sense, an increase in automation leads to an increase in pre-tax income inequality.

## 2.1 The status-quo equilibrium

In this section, we describe the status-quo equilibrium, i.e. the equilibrium under the current U.S. income tax system and no taxes on robot use ($\tau_x = 0$). We model the U.S. income tax system using the functional form for after-tax income proposed by Feldstein (1969), Persson (1983), and Benabou (2000) and estimated by Heathcote, Storesletten and Violante (2017). In this specification, the income tax paid by household $j$ is given by[10]

$$T(w_j l_j) = w_j l_j - \lambda (w_j l_j)^{1-\gamma}, \tag{17}$$

where $\gamma < 1$. Using PSID data, Heathcote, Storesletten and Violante (2017) estimate that $\gamma = 0.181$, which means that income taxes are close to linear. They find that their specification fits the data with an $R^2$ of 0.91. The parameter $\lambda$ controls the level of taxation, higher values of $\lambda$ imply lower average taxes. The parameter $\gamma$ controls the progressivity of the tax code. When $\gamma$ is positive (negative), the average tax rate rises (falls) with income, so the tax system is progressive (regressive).

We assume in all our numerical work that the utility function takes the form:

$$u(c_j, l_j) + v(G) = \log(c_j) - \zeta \frac{l_j^{1+\nu}}{1+\nu} + \chi \log(G). \tag{18}$$

These preferences, which have been used by Ales, Kurnaz and Sleet (2015) and Heathcote, Storesletten and Violante (2017), have two desirable properties. First, they are consistent with balanced growth. Second, they are consistent with the empirical evidence reviewed in Chetty (2006).

For these preferences and the status-quo tax specification, both households choose to work the same number of hours, $l_j = [(1-\gamma)/\zeta]^{1/(1+\nu)}$, which only depend on the preference parameters, $\zeta$ and $\nu$, and the progressivity parameter, $\gamma$.

**Model calibration** We set $\zeta = 10.63$, so that in the status-quo equilibrium agents choose to work $1/3$ of their time endowment. We set $\nu = 4/3$, so that the Frisch

[10] Income in Heathcote, Storesletten and Violante (2017) includes other sources of income, other than labor earnings.

elasticity is equal to 0.75, which is consistent with the estimates discussed by Chetty, Guren, Manoli, and Weber (2011).

Following Heathcote et al. (2017), we choose $\chi = 0.233$ so that the optimal ratio of government to output is 18.9 percent, the same weight observed in the U.S. economy.[11] The tax on robots is zero in the status quo ($\tau_x = 0$). We assume that the level of progressivity of the tax system is $\gamma = 0.181$, the value estimated by Heathcote, et al. (2017). We adjust $\lambda$ to satisfy the government budget constraint.[12] On the production side, we normalize $A$ to one, choose $\alpha = 0.53$ and $\pi_r = 0.55$. These choices are consistent with the share of labor income received by non-routine workers and the fraction of workers that are routine estimated by Chen (2016).[13] These values of $A$, $\alpha$, $\pi_r$ and parameter choices are used in all our numerical experiments.

In our quantitative analysis, we consider a sequence of static economies where the cost of a robot falls geometrically over time, $\phi_t = \phi_0 e^{-g_\phi t}$. We set $g_\phi$, the rate of decline in the price of robots, equal to 0.0083. This value allows our model to match the decline in task content estimated by Acemoglu and Restrepo (2018b) for the period 2000-2008. We choose this period to abstract from the financial crises and focus on the period where automation takes off in the U.S. We set $\phi_0 = 0.4226$ which is the lowest value of $\phi$ consistent with no automation in the status-quo equilibrium (see equation (13)). We assume that time zero corresponds to year 2000 and label our figures accordingly.

Figure 1 describes the effect of changes in the cost of automation. As time goes by, $\phi$ falls causing the wage of routine workers to fall and that of non-routine workers to

[11]When utility takes the form (18), the optimal ratio of government spending to output is the same for all the tax systems we consider.

[12]An alternative approach would have been to keep the tax schedule constant and adjust the level of government spending to balance the government budget. However, this approach would make it more difficult to compare the solutions for the different tax systems.

[13]The equilibrium is independent of the value of $\rho$, the parameter that controls the elasticity of substitution between different tasks. The reason for this result is that all the factors (non-routine workers and/or robots) used in equilibrium to perform these tasks have the same marginal cost.

rise. Since the utility function is logarithmic and wages are the only income source, hours worked remain constant for both routine and non-routine workers. This property reflects the offsetting nature of income and substitution effects. Given that as $\phi$ falls, wages of routine workers fall and their hours worked remain constant, their income, consumption, and utility fall. In contrast, non-routine workers benefit from rising income, consumption and utility.

As $\phi$ falls, the parameter that controls the level of taxation, $\lambda$, rises, which implies a decline in the overall level of taxation. This decline reflects the increasing share of tax revenue paid by non-routine workers pay and the fact that, as $\phi$ falls, their income rises faster than output.

In sum, our analysis suggests that the current U.S. tax system will lead to massive income and welfare inequality in response to a fall in the costs of automation.

## 2.2 The first-best allocation

It is useful to start our analysis by considering the first-best benchmark. This is the allocation that maximizes a weighted average of the utility of agents, subject only to technological constraints. Implicitly we are assuming that there are lump-sum taxes/transfers targeted to the different agents, that can be used to implement the allocation. The solution to this planning problem will be useful to serve as a comparison to the solutions with more restricted tax systems.

The welfare weights are $\omega_r$ and $\omega_n$ to routine and non-routine agents, respectively. These weights are normalized so that $\pi_r\omega_r + \pi_n\omega_n = 1$. The planner's problem is to choose $\{c_r, l_r, c_n, l_n, G, m, \{x_i, n_i\}\}$ to maximize social welfare,

$$\mathcal{W} \equiv \pi_r\omega_r \left[u(c_r, l_r) + v(G)\right] + \pi_n\omega_n \left[u(c_n, l_n) + v(G)\right]. \tag{19}$$

subject only to resource feasibility.

The first-best allocation features production efficiency. The robots' marginal productivity equals marginal cost, $\phi$, so the robot tax is zero. The first-best allocation is on the production possibilities frontier, which implies that, with advances in automation, pre-tax income inequality rises. This is still consistent with optimality because, with unrestricted taxes/transfers, it is possible to compensate for these effects. Once we introduce restrictions to the tax system, pre-tax wage inequality may become relevant for redistribution and production efficiency may cease to be optimal.

In our quantitative exercises, we assume equal welfare weights, $\omega_r = \omega_n = 1$. Figure 2 illustrates the properties of the first-best solution. Panel A shows that full automation occurs only asymptotically. The real wage rate for both types of workers are the same as in the status-quo equilibrium, meaning that pre-tax income inequality is rising.[14] The first-best allocation gives the same consumption to both types of workers and requires larger labor supply from the more productive non-routine workers.[15] The consumption and utility of both types of worker rise as $\phi$ falls. Figure 2 also shows that implementing the first-best solution requires large transfers from non-routine to routine workers.

The utility of routine workers exceeds that of non-routine workers. Clearly, the first-best solution cannot be implemented if the planner cannot discriminate between household types. In this solution, non-routine households would have an incentive to act as routine to benefit from a more generous consumption and leisure bundle.

## 2.3 Mirrleesian optimal taxation

In this section, we characterize the optimal non-linear income tax when the planner observes a worker's total income but does not observe the worker's type or labor

[14]The reason for this property is as follows. Equations (14) and (15) imply that wages depend on technological parameters ($\alpha$ and $A$), the cost of automation, and the value of $\tau_x$. Since $\tau_x = 0$ in the status quo and there is production efficiency in the first-best allocation, the wages are the same in both allocations.

[15]One interpretation of the social welfare function with equal weights is as the ex-ante expected utility of an agent who faces uncertainty about their labor market skills. The unrestricted optimal insurance with separable utility features equal consumption in all states.

supply, as in the canonical Mirrlees (1971) problem. For the reasons emphasized in Guesnerie (1995), we assume that robot taxes are linear.

In the analytical description of the optimal policy, we focus attention on plans with interior automation, $m > 0$.[16] We also assume that $\phi \leq \alpha^{\alpha}(1-\alpha)^{1-\alpha}A$, so that if $\tau_x \leq 0$ non-routine workers earn a higher wage ($w_n \geq w_r$) in an equilibrium with automation (see equations (14) and (15)).

The Mirrleesian planning problem is to choose the allocations $\{c_j, l_j\}_{j=r,n}$, $G$, and the robot tax $\tau_x$ to maximize social welfare, in equation (19), subject to the resource constraint

$$\pi_r c_r + \pi_n c_n + G \leq \pi_n w_n l_n \frac{\tau_x + \alpha}{\alpha(1+\tau_x)} + \frac{\pi_r w_r l_r}{1+\tau_x}. \tag{20}$$

and two incentive constraints (IC)

$$u(c_n, l_n) \geq u\left(c_r, \frac{w_r}{w_n} l_r\right), \tag{21}$$

$$u(c_r, l_r) \geq u\left(c_n, \frac{w_n}{w_r} l_n\right), \tag{22}$$

The wages of the two types of workers are given by equations (14) and (15). As is standard in the literature, the conditions (20), (21), and (22) are necessary and sufficient to describe a competitive equilibrium.[17]

In the Mirrlees (1971) model, the productivities of different agents are exogenous. In our model, these productivities are endogenous[18] and depend on $\tau_x$. This property is central to the question we are interested in studying: is it optimal to distort production decisions by taxing the use of robots to redistribute income from non-routine to routine workers to increase social welfare?

[16]We do not discuss the case where $m = 0$ because in this case the results in Stiglitz (1982) apply to our model.

[17]To save on space, We discuss necessity and sufficiency of these equations for a competitive equilibrium in the appendix.

[18]In this characteristic, our work relates to a large literature on Mirrleesian taxation with general equilibrium effects on prices/wages: Stiglitz (1982), Naito (1999), Rothschild and Scheuer (2013), Scheuer (2014), Ales, Kurnaz, and Sleet (2015), Sachs, Tsyvinski, and Werquin (2019), among others.

The tax on intermediate goods provides the government with an additional instrument relative to the Mirrlees (1971) setting. The planner can use this instrument to affect the income of the two types of workers but its use distorts production.

To bring the analysis closer to a canonical Mirrleesian approach, we maximize the planner's objective in two steps. First, we set $\tau_x$ to a given level and solve for the optimal allocations. Second, we find the optimal level of $\tau_x$. We define $W(\tau_x)$ as the maximum level of social welfare, (19), subject to the incentive constraints, (21) and (22), and the resource constraint, (20) for a given value of $\tau_x$. An optimal choice of $\tau_x$ requires that $W'(\tau_x) = 0$.

The expression for net output in the right-hand side of equation (20) can be written as

$$\frac{\tau_x + \alpha}{\alpha(1+\tau_x)^{1/\alpha}} \frac{\alpha A^{1/\alpha}(1-\alpha)^{\frac{1-\alpha}{\alpha}}}{\phi^{\frac{1-\alpha}{\alpha}}} \pi_n l_n + \phi \pi_r l_r.$$

The term $(\tau_x + \alpha) / \left[\alpha (1+\tau_x)^{1/\alpha}\right]$ is equal to one for $\tau_x = 0$ and strictly less than one for $\tau_x \neq 0$. This term is a measure of the production inefficiency created by the tax on robots. With automation is incomplete, the planner is willing to pay a resource cost, in terms of this production inefficiency, in order to loosen the incentive constraints that are also functions of the robot tax.

**Proposition 1.** *Suppose the optimal allocation is such that the non-routine workers' incentive constraint binds and the incentive constraint for routine workers does not bind. Then, if automation is incomplete ($m < 1$ and $l_r > 0$), robot taxes are strictly positive ($\tau_x > 0$). The optimal tax on robots satisfies*

$$\frac{\tau_x}{1+\tau_x} = \frac{\alpha}{1-\alpha} \frac{\pi_r \phi l_r}{\pi_n w_n l_n} \left[1 + \frac{\omega_r u_l\left(c_r, l_r\right)}{\mu \phi}\right], \tag{23}$$

*where $\mu$ denotes the multiplier on the resource constraint (20).*

This proposition is proved in the appendix. To see the intuition for this result, suppose that $\tau_x < 0$. A marginal increase in $\tau_x$ has two benefits. First, it strictly increases output and hence the amount of goods available for consumption. Second, it reduces the relative wage $w_n/w_r$ and makes the non-routine worker less inclined to mimic the routine workers. This property can be easily seen from the incentive constraint of the non-routine worker: $u(c_n, l_n) \geq u\left(c_r, \frac{w_n}{w_r} l_n\right)$.[19]

Consider instead $\tau_x = 0$. Since a zero tax on robots maximizes output, for fixed labor supplies, a marginal increase in that tax produces only second-order output losses. On the other hand, increasing $\tau_x$ generates a first-order gain from loosening the informational restriction. Therefore, starting from $\tau_x = 0$, the planner can always improve welfare with a marginal increase in $\tau_x$.

Robot taxes are optimal only when automation is incomplete ($m < 1$), so that routine workers are employed in production ($l_r > 0$). When full automation is optimal ($m = 1$, $l_r = 0$) there is no informational gains from taxing robots. Since the robot tax distorts production and does not help loosen the incentive constraint of the non-routine agent, the optimal value of $\tau_x$ is zero. We prove this result in the appendix.

We now turn to the study of the optimal wedges. The optimality conditions imply the following marginal rates of substitution:

$$\frac{u_l(c_n, l_n)}{u_c(c_n, l_n)} = w_n \frac{\tau_x + \alpha}{\alpha(1 + \tau_x)},$$
$$\frac{u_l(c_r, l_r) l_r}{u_c(c_r, l_r)} = \frac{\omega_r - \eta_n \frac{u_c(c_r, w_r l_r / w_n)}{u_c(c_r, l_r)}}{\omega_r - \eta_n \frac{u_l(c_r, w_r l_r / w_n) \frac{1}{w_n}}{u_l(c_r, l_r) \frac{1}{w_r}}} \frac{w_r l_r}{1 + \tau_x},$$

where $\eta_n \pi_r$ denotes the Lagrange multiplier of the incentive constraint of the non-routine worker.

One property of the original Mirrlees (1971) model is that the labor-supply decision of the high-ability agent should not be distorted. However, Stiglitz (1982)

[19]We characterize optimal allocations in which the incentive constraint of the non-routine worker binds, and the incentive constraint of the routine worker is slack. This pattern holds in all our numerical exercises.

showed that with endogenous productivities, there high-ability agents should instead be subsidized on the margin.

In our model, non-routine workers are subsidized at the margin when automation is incomplete. This subsidy corrects for the difference between the productivity as perceived by the firm ($w_n$) and the marginal increase in the resources available to the planner from a marginal increase in $\pi_n l_n$, which is equal to $w_n \left(\tau_x + \alpha\right) / \alpha(1 + \tau_x)$, where $\left(\tau_x + \alpha\right) / \alpha(1 + \tau_x) > 1$.[20]

Routine workers are taxed at the margin when automation is incomplete for two reasons. First, this tax corrects the distortion created by robot taxes, which make the wages of routine workers higher than the marginal increase in the resources available to the planner from a marginal increase in $\pi_r l_r$. Second, taxing routine workers makes it less appealing for non-routine workers to mimic routine workers and loosens the IC of non-routine workers.

Figure 3 illustrates the properties of the equilibrium associated with Mirrleesian optimal taxation. The process of automation begins later in the Mirrleesian solution than in the first best. This property reflects the presence of robot taxes in the Mirrleesian solution. These taxes increase the wages of routine workers and decrease the wages of non-routine workers. This wage compression loosens the incentive constraint of non-routine workers, which allows the government to redistribute more income from non-routine to routine workers.

The path for the tax rate on robots has a hump shape. The economy starts with inequality in wages that makes redistribution desirable. Since initially the cost of distorting automation is relatively small, the planner chooses a level of robot taxes which halts the process of automation. As the costs of automation fall, robot taxes increase to prevent automation from occurring. After this initial period, robot taxes fall. As robots become cheaper, it is inefficient to use routine workers so their labor

[20]Note that, in our model, the general-equilibrium effects emphasized by Stiglitz (1982) are reduced in our model to the impact of the robot tax on pre-tax wages.

supply falls. This decline in routine hours makes robot taxes less useful as a tool for income redistribution. In the limit, routine hours converge to zero and so do robot taxes.

Consumption of non-routine workers is higher than that of routine workers. Since non-routine workers work harder than routine workers, the former need to receive higher consumption to satisfy their incentive constraint. Both types of workers see their consumption rise as $\phi$ approaches zero. This outcome is achieved through large transfers to the routine workers.

In the limit, routine households stop working and live off government transfers. Those transfers are generous enough that the utilities of the two worker types are equalized. The reason for this equalization is that, once routine workers supply zero hours, there is no difference between the non-routine worker pretending to be routine and that of the routine worker.

## 2.4 Ramsey optimal taxation

In this section, we compare the Mirrleesian allocation with the solution to a Ramsey (1927)-style optimal taxation problem in which the tax schedule is restricted to take a parametric functional form. The particular functional form we consider is the one that has been widely used to describe the U.S. income tax system by Heathcote, Storesletten, and Violante (2017), among many others.

Our goal is to assess which tax-system characteristics play a crucial role in approximating well the Mirrleesian solution. In that objective, our analysis is related to Heathcote and Tsujiyama (2019). These authors ask whether the Mirrleesian solution is best approximated either by linear taxation with a lump-sum transfer or by a tax system that is progressive but does not feature a lump-sum component. They find that, in their model, the second form of tax system performs better. Contrary to their findings, the present section shows that during skill-biased technological transitions a lump-sum transfer is an important instrument of the tax system. These transfers can be interpreted as a form of universal basic income.

The tax/transfer system is assumed to take the form described in equation (17), extended to also include a lump-sum transfer. This function has three parameters, $\gamma$ and $\lambda$, and $\Omega$. When $\gamma$ is zero, the tax system is linear with a rate $\lambda$. This parameter captures the level of taxes. The parameter $\gamma$ captures progressivity in the tax system. Finally, $\Omega$ denotes the lump-sum transfer. If the household has income $y$, their after-tax income is thus $\lambda y^{1-\gamma} + \Omega$.[21][22] In our numerical exercises, we will consider both the case when transfers are restricted to be zero, $\Omega = 0$, and when they can be freely chosen by the government. Because these forms of taxation have been extensively studied by Heathcote, Storesletten, and Violante (2018), we call the case with $\Omega = 0$ the HSV-type taxes case. We refer to the case $\Omega \neq 0$ as the HSV-type taxes with lump-sum transfers case.

We characterize the competitive equilibrium for this economy in the Appendix. Using these equations, we can write the ratio of the consumption of routine and non-routine workers as:

$$\frac{c_r - \Omega}{c_n - \Omega} = \left[ \frac{(1-\alpha)(1-m)}{\alpha} \frac{\pi_n}{\pi_r} \right]^{1-\gamma}. \tag{24}$$

This condition results from the fact that the government must set the same income tax schedule for both routine and non-routine agents.

Suppose first that $\Omega = 0$. In this case, equation (24) shows that there are two ways to make the ratio $c_r/c_n$ closer to one. One way is to raise $\tau_x$ which leads to a fall in the level of automation, $m$. The other way is to make $\gamma$ closer to one, i.e. make the tax

[21]This function has been widely used to study the U.S. income tax system and has recently been estimated by Heathcote, Storesletten and Violante (2017), for the case $\Omega = 0$.

[22]When $\Omega \neq 0$ progressivity the tax system may still be progressive even if $\gamma = 0$, as long as the transfer is positive. Overall progressivity of the tax system should be analyzed from average taxes which are given by

$$\frac{T(y)}{y} = 1 - \lambda y^{-\gamma} - \frac{\Omega}{y}.$$

The average tax is increasing as long as $\gamma \geq 0$ and $\Omega \geq 0$.

system more progressive. Both approaches have drawbacks. Taxing robots distorts production. Increasing progressivity reduces incentives to work. If the government can also use transfers, $\Omega$, then it is clear that higher positive transfers can help in bringing the ratio of consumption closer to one, holding everything else constant. Again, this benefit comes at the cost of desincentivizing labor supply.

Since we are interested in studying optimal taxation in an economy with automation, we focus on equilibria where $m > 0$. In this case, equation (24) can be written as:

$$c_r - \Omega = (c_n - \Omega)\left[\frac{(1-\alpha)}{\alpha}\frac{\pi_n}{\pi_r}\left(\frac{\phi(1+\tau_x)}{(1-\alpha)A}\right)^{1/\alpha}\frac{\pi_r l_r}{\pi_n l_n}\right]^{1-\gamma}, \tag{25}$$

The planning problem is to choose allocations $\{c_n, c_r, l_r, l_n, G\}$ and the tax parameters $\{\tau_x, \gamma, \Omega\}$, subject to the following conditions

$$u_c(c_j, l_j)\left(c_j - \Omega\right) + \frac{u_l(c_j, l_j) l_j}{1-\gamma} = 0,\ j = r, n, \tag{26}$$

$$u\left(c_j, l_j\right) \geq u\left(\Omega, 0\right)\ \text{if}\ \Omega \geq 0,\ j = r, n\ , \tag{27}$$

and the resource constraint with interior automation, (20).

Conditions (26) are obtained from the budget constraints for each household type, combining the first-order conditions to replace prices and taxes. They are the usual Ramsey implementability constraints. The second condition (25) imposes that the tax system is the same for both household types. With positive lump-sum transfers and regressivity, $\gamma < 0$, the solution to the households problem may not be interior, meaning that the household may choose to work zero hours and set consumption equal to the transfer. The conditions (27) impose that the household's allocation does not yield lower utility than that corner solution. We show in the appendix that these conditions are necessary and sufficient for a competitive equilibrium in the quantities $\{c_n, c_r, l_r, l_n, G\}$ and the tax parameters $\{\tau_x, \gamma, \Omega\}$.

The optimal tax on robots satisfies

$$\frac{\tau_x}{1+\tau_x} = \frac{\alpha}{1-\alpha}\frac{\eta\,(1-\gamma)}{\mu w_n \pi_n l_n}\left(\frac{c_r-\Omega}{c_r}\right), \tag{28}$$

where $\eta/c_r$ is the multiplier for the no-discrimination constraint (25) and $\mu$ is the multiplier for the resource constraint.[23] This expression also holds for the case in which the lump-sum transfers is constrained, just by setting $\Omega = 0$.

Since the marginal utility of public expenditures is always positive, the marginal value of resources to the planner, given by the multiplier $\mu$, is strictly positive. The multiplier $\eta$ captures the marginal value of redistributing income to routine households, which is limited by the assumptions on the income-tax function. If $\eta > 0$, the marginal value of additional redistribution of income towards routine workers is positive and robot taxes are strictly positive as long as $c_r > \Omega$, i.e. as long as routine workers also have some labor income.

In the case in which transfers are constrained to be zero, then robot taxes must always be strictly positive. The intuition for this result is that since the government has to use the same income tax function for both types of workers, taxing robots helps redistribute income by increasing the pre-income tax wage of routine workers and lowering that of non-routine workers.

When transfers can be freely chosen, the optimal plan may feature $c_r = \Omega$, in which case the routine worker supplies zero labor hours. In this limiting case, there is no role for redistributing by affecting the relative wages, and robot taxes should be set to zero.

Figure 4 shows the optimal policy when transfers are constrained to be zero, $\Omega = 0$. It shows that the form of the tax function constrains heavily the outcomes that can be achieved. As discussed above, the planner can redistribute income by taxing robots or by increasing progressivity. Taxing robots distorts production and

[23]See the appendix for the derivation.

higher progressivity reduces incentives to work. Since initially the cost of distorting automation is relatively small, the planner chooses a level of robot taxes consistent with no automation. Robot taxes are heavily used, reaching values as high as $\tau_x = 0.33$. As the costs of automation decline, the progressivity of the income tax rises. But there is still a large divergence in wage rates, consumption and utility across the two types of workers.

Figure 5 illustrates the properties of this allocation when transfers can be freely chosen. Workers have two sources of income: wages and transfers. For this reason, income and substitution effects of changes in wages are no longer offsetting. As a consequence, the two types of workers supply a different number of hours and their hours vary with $\phi$.

The optimal solution with lump-sum transfers features full automation even when robots are still relatively expensive. In contrast, full automation occurs only asymptotically with Mirrleesian taxes and optimal simple taxes. Once full automation occurs, routine households have no labor income. Since only non-routine workers pay income taxes, the planner designs the tax system to avoid distorting their marginal labor-supply decisions. This result is achieved by increasing the regressivity of the tax system (i.e. lowering $\gamma < 0$) so that, given the level of taxation implied by $\lambda$, the marginal income tax rate for non-routine workers is zero. In this way, the taxation of non-routine households is effectively equivalent to lump-sum taxes.[24] The level of transfers is chosen so that the non-routine worker is indifferent between the interior solution, with positive labor, and the corner solution with zero labor and consumption equal to transfers. For a higher level of transfers, both agents would supply zero labor.[25]

[24]The same logic implies that in a representative-agent economy it is possible to use the Heathcote et al. (2018) tax function to obtain the same allocation as with lump-sum taxes. This allocation is achieved by choosing a regressive tax system such that the marginal tax rate is zero. Government expenditures are financed with the revenue raised by the infra-marginal tax rates.

[25]In the appendix, we show a numerical example of the individual agent's problem with one such solution.

The availability of lump-sum transfers is essential for full automation to occur in finite time. Without lump-sum transfers, a routine worker who drops out of the labor force has zero consumption. For this reason, routine workers never drop out of the labor force.

One surprising result is that full automation occurs with this restricted tax system and not in the Mirrleesian solution which also allows for lump-sum transfers. The intuition for this result is as follows. Because preferences are separable and have constant Frisch elasticity, the marginal disutility of labor converges to zero as labor hours approach zero. For this reason, a Mirrleesian planner finds it optimal to have routine workers supply positive labor, even if productivity is very low. With a general tax function it is possible to implement very different marginal distortions for both types of worker. As we have seen before, the Mirrleesian solution features a negative marginal distortion for non-routine workers, and a positive marginal distortion for routine workers. This solution cannot be obtained with the simple tax function. The restrictions on the marginal distortions associated with this function are such that the planner prefers a corner solution for the labor supply of routine workers. By excluding routine workers from the labor force, the planner can design the simple tax system to target only non-routine workers, reducing their marginal tax rate to zero so that in effect they are taxed in a lump-sum fashion.

In such an equilibrium, income is redistributed through a large lump-sum transfer. This transfer can be interpreted as a minimum income that is guaranteed to all agents in the economy. When automation is incomplete, robot taxes are used as an additional source of redistribution and $\tau_x$ can be as high as 37 percent. Complete automation occurs shortly after 2050, once the cost of robots drops below $\phi = 0.27$.

## 2.5 Comparing different policies

In this section, we compare the first-best allocation with the allocations associated with different policies in terms of social welfare and the utility of routine and non-routine workers. In the figures discussed below, we use the labels FB, SQ, OT, HSV and HSV-L to refer to the first-best, status-quo, Mirrleesian optimal taxes, HSV-type taxes, and HSV-type taxes with lump-sum transfers, respectively.

Figure 6 shows the welfare of the utilitarian social planner under the different policies, between the years 2000 and 2150. Recall that the level of $\phi$ in the year 2000 was chosen as the lowest value for which there is no automation in the status quo. Social welfare rises as the costs of automation fall both for the first best and for all the policies we consider. We see that the Mirrlees allocation is relatively close in terms of welfare to the first-best allocation. The solution with simple taxation and lump-sum transfers ranks next in terms of welfare, followed by the solution with simple taxes without rebates. The status quo is by far the worst allocation.

A fall in the cost of automation can have very different consequences for routine and non-routine workers. To illustrate this property, we measure the utility of the two types of workers relative to the status-quo equilibrium in year 2000. We call this allocation the no-automation benchmark. Panel A (B) of Figure 7 shows how much routine (non-routine) workers would have to be compensated in the no-automation benchmark to be as well off as in the policy under consideration, for the different years. The measure is computed as a percentage of consumption.

Panel A of Figure 7 shows that the utility of routine workers in the first-best allocation improves as $\phi$ falls over time. In contrast, in the status quo, routine workers become increasingly worse as automation becomes more pronounced. With Mirrleesian optimal taxation, routine workers are always made better off. With simple income taxes, routine workers are not made better off until after 2150. We can see that including a universal form of income is a simple way of recovering gains for routine workers. Indeed, shortly after 2050 the routine worker is almost as well off in this solution as in the solution with Mirrleesian taxes.

Panel B of Figure 7 shows that non-routine workers prefer the no-automation benchmark to the first best while automation costs are relatively high (almost until

2100). This preference reflects the large transfers that non-routine workers make to routine workers in the first best. Once automation costs fall by 62 percent relative to their 2000 value, which happens in 2116, non-routine workers prefer the first best to the no-automation benchmark. The reason is that the wage of non-routine workers is high enough to compensate for the transfers they make to routine workers. Starting in 2015, non-routine workers prefer the status quo to all other allocations. This preference results from a combination of high wages and relatively low taxes in the status quo.

Routine workers always rank the first-best allocation first, Mirrleesian optimal taxation second, HSV-L third, HSV fourth, and the status quo last. The utility for the allocation with simple taxes and lump-sum transfers approaches that under Mirrleesian taxes as the cost of automation falls. In contrast, non-routine workers rank the status quo first and the first best last. Mirrleesian optimal taxation, and HSV and HSV-L transfers rank in between the two extremes.

# 3 A dynamic model

In this section, we enrich the analysis of the static model to a more complex dynamic overlapping-generations model. This model features several important characteristics. First, in the dynamic model, robots are an investment good which can be accumulated over time. Second, we consider technological progress in the form of investment-specific technical change as in Greenwood, Hercowitz, and Krusell (1997). Because robots are more substitutable with routine workers than non-routine workers, the technical change is skill biased which impacts the routine-skill premium.[26] Technological advances lead to rising productivity, but increase pretax wage inequality among occupations. For that reason, a good description of the long-

[26] This is analogous to the skill-premium effects of capital-skill complementarity in Krusell, Ohanian, Rios-Rull, and Violante (2000).

run consequences of technological transitions requires an explicit model of skill acquisition by workers. Our model endogeneizes the choice of skills by adopting an overlapping-generations framework in which workers can choose their skills upon entering the economy.

**Workers and preferences** Time is discrete and infinite, $t = 1, 2, ...$ Each worker is characterized by a duplet $(\theta, t) \in \Theta \times \mathbb{N}_0$, where $t$ denotes the time at which the household enters the economy, and $\theta \in \Theta \subset \mathbb{R}$ denotes the household's cost of skill acquisition. There is a continuum of unit measure of workers born in each period. We assume that $\theta$ is distributed according to an absolutely continuous distribution function $\Lambda$, with p.d.f. $\lambda$. Each worker lives for two periods. For the sake of exposition, we say that workers born in period $t$ are "young" in period $t$, and "old" in period $t+1$.

When young, workers choose their skills, work, consume, and save. When old, workers only have a consumption decision. Agent's who were born in $t = 0$ enter the economy as old.

We denote the consumption of the young household at time $t$ by $c_t^y(\theta)$, and consumption of the old by $c_t^o(\theta)$. Labor supply by the young household is denoted by $l_t(\theta)$. Households can choose whether to learn non-routine occupations, we denote this decision by $s_t(\theta)$. If $s_t(\theta) = 1$ then the household has chosen to learn non-routine worker skills, while $s_t(\theta) = 0$ denotes that the worker is a routine worker.

The household's choice of skills determines their occupation, either routine or non-routine. We denote by $\Theta_r$ and $\Theta_n$ the subsets of $\Theta$ which become routine and non-routine, respectively, i.e.

$$\Theta_{n,t} \equiv \{\theta : s_t(\theta) = 1\} \tag{29}$$

and $\Theta_{r,t} \equiv \Theta - \Theta_{n,t}$.

The utility function of household $(\theta, t)$ is given by

$$U_t(\theta) = u\left(c_t^y(\theta)\right) - \psi\left(l_t(\theta)\right) + v(G_t) + \beta[u(c_{t+1}^o(\theta)) + v(G_{t+1})] - \theta s_t(\theta) \qquad (30)$$

where $G_t$ denotes public goods at time. We assume that the utility function is separable in consumption and labor[27] and verifies the standard assumptions of monotonicity and concavity. Old workers in period 1 are all alike, they derive utility from consumption, $c_1^0$, and public goods, $G_1$, and their utility function is

$$U_0 = \beta[u(c_1^o) + v(G_1)].$$

The utility representation above has the following interpretation: households have heterogeneous costs with respect to the acquisition of routine and non-routine skills. A worker with a positive $\theta$ faces a positive cost of acquiring non-routine skills, which means that all else equal they would prefer to acquire routine skills. If $\theta$ is negative, then all else equal the worker would prefer to acquire non-routine skills.

**Firm's, technology, and equilibrium** At time $t$, the cost of producing a robot is equal to $\phi_t$ units of output. Robots are produced by competitive firms. A representative firm producing robots chooses $i_t$ to maximize profits $p_{x,t}i_t - \phi_t i_t$. It follows that in equilibrium $p_{x,t} = \phi_t$ and profits are zero. The stock of robots evolves according to:

$$X_{t+1} = (1-\delta)X_t + i_t/\phi_t,$$

where $\delta$ denotes the depreciation rate of robots.

Following the literature on automation, we adopt a task-based framework to model the production of final goods.[28] The production of final goods combines

[27] As is well known, without weak separability between consumption and leisure, the uniform taxation result in Atkinson and Stiglitz (1976) fails. In a dynamic setting, this failure would mean that the optimal plan features intertemporal distortions, i.e. it is optimal to tax consumption in different periods at different rates. This reason to tax capital is orthogonal to the one we focus on and, for that reason, we assume separability.

[28] See Acemoglu and Restrepo (2018a, 2018b).

non-routine labor, $N_{n,t}$, with a continuum of routine tasks $y_{i,t}$ for $i \in [0,1]$, in the following way

$$Y_t = A\left[\int_0^1 y_{i,t}^{\frac{\rho-1}{\rho}}\,di\right]^{\frac{\rho}{\rho-1}(1-\alpha)} N_{n,t}^{\alpha}, \tag{31}$$

where $A$ denotes total factor productivity and $\alpha$ is the share of non-routine workers in production. The assumption that the elasticity between total routine services and non-routine labor is unitary is important because it ensures the existence of a balanced growth path which is reached asymptotically.[29]

As in the static model, each task $i$ can be done by workers $n_i$, or it can be automated and performed by robots.

$$y_{i,t} = \begin{cases} \kappa_i x_{i,t}, & \text{if } i \text{ is automated} \\ \ell_i n_{i,t}, & \text{if } i \text{ is not automated} \end{cases} \tag{32}$$

The parameters $\kappa_i$ and $\ell_i$ capture the efficiency of robots and routine labor, respectively, in task $i$. We assume that $\kappa_i/\ell_i$ is weakly decreasing in $i$.

Firm's maximize after-tax dividends, which are given by

$$Y_t + p_{x,t}(1-\delta)X_t - R_t X_t - w_{r,t}N_{r,t} - w_{n,t}N_{n,t} - \tau_t^p \pi_t \tag{33}$$

where $\pi_t = Y_t - p_{x,t}\delta X_t - w_{r,t}N_{r,t} - w_{n,t}N_{n,t}$. As in the static model, $X_t = \int_0^1 x_{i,t}$ and $N_{r,t} = \int_0^1 n_{i,t}$.

The optimal allocation of robots and routine workers over the different tasks follows the same principles of our static model, i.e. in period $t$, the firm uses robots in the first $m_t$ and routine workers in the final $1-m_t$ tasks. The optimal allocation of routine workers and robots to each of those tasks have the same first-order condi-

[29] It is well known that our results hold for more general production structures. In the main text, we restrict the analysis to parametric assumptions which ensure the existence of an asymptotic balanced growth path, and keep our exposition close to the literature on automation. The results could be easily extended.

tions as in the static model. These conditions imply:

$$x_{i,t} = \frac{\kappa_i^{\rho-1}}{\int_0^{m_t} \kappa_j^{\rho-1} dj} X_t, \ i \in [0, m_t] \quad \text{and} \quad n_{i,t} = \frac{\ell_i^{\rho-1}}{\int_{m_t}^1 \ell_j^{\rho-1} dj} N_{r,t}, \ i \in (m_t, 1].$$

As in Acemoglu and Restrepo (2018a,b), we can replace these expressions in the production function to obtain:

$$Y_t = A \left[ \left( \int_0^{m_t} \kappa_i^{\rho-1} di \right)^{\frac{1}{\rho}} X_t^{\frac{\rho-1}{\rho}} + \left( \int_{m_t}^1 \ell_i^{\rho-1} di \right)^{\frac{1}{\rho}} N_{r,t}^{\frac{\rho-1}{\rho}} \right]^{\frac{\rho}{\rho-1}(1-\alpha)} N_{n,t}^{\alpha} \tag{34}$$

The optimization with respect to $m_t$ implies the following relation:

$$\frac{X_t}{N_{r,t}} = \frac{\int_0^{m_t} \kappa_i^{\rho-1} di}{\int_{m_t}^1 \ell_i^{\rho-1} di} \frac{\ell_{m_t}^{\rho}}{\kappa_{m_t}^{\rho}}.$$

At this level of generality, the above expression cannot be solved in closed form. However, under a suitable assumption this expression can be solved and yield an aggregate production function which renders our analysis tractable.

**Assumption 1.** *Assume that* $\kappa_i = \left[1 - \frac{\rho-1}{\varepsilon-1}\right] i^{-\frac{1}{\varepsilon-1}}$ *and* $\ell_i = \left[1 - \frac{\rho-1}{\varepsilon-1}\right] (1-i)^{-\frac{1}{\varepsilon-1}}$*, where* $\varepsilon > \rho$*.*

Under this assumption, the solution with respect to $m_t$ is just

$$m_t = \frac{X_t^{\frac{\varepsilon-1}{\varepsilon}}}{X_t^{\frac{\varepsilon-1}{\varepsilon}} + N_{r,t}^{\frac{\varepsilon-1}{\varepsilon}}}, \tag{35}$$

and the routine services aggregator becomes just a CES aggregator of total robots and routine work, where $\varepsilon > 0$ denotes the elasticity of substitution between robots and routine workers:[30]

$$Y_t = A \left[ X_t^{\frac{\varepsilon-1}{\varepsilon}} + N_{r,t}^{\frac{\varepsilon-1}{\varepsilon}} \right]^{\frac{\varepsilon-1}{\varepsilon}(1-\alpha)} N_{n,t}^{\alpha}. \tag{36}$$

[30]This derivation is similar to the one in Chen (2017).

For short we write this production function as $F(X_t, N_{r,t}, N_{n,t})$. Defining the derivatives $F_{x,t} \equiv dF(X_t, N_{r,t}, N_{n,t})/dX_t$, and $F_{j,t} = \partial F(N_{r,t}, N_{n,t}, X_t)/\partial N_{j,t}$ for $j = r, n$, we have that the first order conditions with respect to $X_t$, $N_{r,t}$, and $N_{n,t}$ imply that

$$\begin{aligned} w_{j,t} &= F_{j,t}, \quad \text{for } i = r, n, \\ R_t &= p_{x,t} + (1 - \tau_t^p)(F_{x,t} - p_{x,t}\delta). \end{aligned}$$

As discussed in Chari, Nicolini, and Teles (2018), with the decentralization we adopt, the linear tax on profits acts like a distortion on the optimality conditions for robots in the same way as was originally assumed by Judd (1985) and Chamley (1986).

As in the previous section, the important characteristic is that robots share a higher degree of complementarity with non-routine work than with routine work. For that, we assume that $\varepsilon > 1$ such that routine workers and robots are substitutes. Note that this implies that

$$\mathcal{E}_t \equiv \frac{d \log (F_{r,t}/F_{n,t})}{d \log X_t} = -\frac{\varepsilon - 1}{\varepsilon} m_t \leq 0.$$

The resource constraint in period $t$ can be written as

$$\sum_{a=y,o} \int_\theta c_t^a(\theta)\lambda(\theta)d\theta + G_t + \phi_t \left[X_{t+1} - (1-\delta) X_t\right] \leq F\left(X_t, N_{r,t}, N_{n,t}\right), \tag{37}$$

where $N_{j,t} \equiv \int_{\Theta_{j,t}} l_t(\theta)\lambda(\theta)d\theta$. The budget constraint of the government is implies by Walras' law.

**Definition 1.** *An equilibrium is a set of individual allocations* $\{(c_t^y(\theta), c_t^o(\theta), l_t(\theta), s_t(\theta), x_t(\theta))_{\theta\in\Theta}\}_{t\geq 1}$*, aggregates* $\{G_t, N_{r,t}, N_{n,t}, X_t, m_t\}$*, prices* $\{w_{r,t}, w_{n,t}, p_{x,t}\}_{t\geq 1}$*, and a tax system* $\{T_t^y(\cdot), T_t^o(\cdot), \tau_t^a(\cdot), \tau_t^p\}_{t\geq 1}$ *such that: (i) given prices and taxes, allocations solve the households' problem; (ii) given prices and taxes, allocations solve the firms' problem; (iii) the government budget constraint is satisfied; and (iv) markets clear*

## 3.1 First-best allocation

We assume that the planner assigns Pareto weights $\beta^t \omega_t(\theta)$ to agents of type $(\theta, t)$, and we normalize the weights, $\omega_t(\theta)$, such that

$$(1-\beta)\sum_{t=0}^{\infty}\int_{\Theta}\beta^t \omega_t(\theta)\lambda(\theta)d\theta = 1.$$

The planner's objective function is

$$\mathcal{W} \equiv \sum_{t=0}^{\infty}\int_{\Theta}\beta^t \omega_t(\theta)U_t(\theta)\lambda(\theta)d\theta. \tag{38}$$

We assume that the weights converge in the long run, i.e. for all $\theta$, $\omega_t(\theta) \to \omega(\theta) \geq 0$ as $t \to \infty$.

The first-best allocation maximizes this welfare function subject to the resource constraints (37). The solution to this problem implies the following efficiency conditions, which equate marginal rates of substitution to marginal rates of transformation

$$\frac{\psi'(l_t(\theta))}{u'(c_t^y(\theta))} = F_{s_t(\theta)}(t),$$

and

$$\frac{u'(c_t^y(\theta))}{\beta u'(c_{t+1}^o(\theta))} = \frac{F_x(t+1) + \phi_{t+1}(1-\delta)}{\phi_t},$$

for all $\theta$.

Assume that the Pareto weights are such that $\omega_t(\theta)\theta$ is weakly growing in $\theta$, then the first-best level of skill acquisition is determined by a threshold rule. This means that there exists $\theta_t^* \in \Theta$ such that for smaller enough $\theta$, $\theta < \theta_t^*$, we have that $s_t(\theta_i = n$ and all large enough $\theta > \theta_t^*$ become routine workers, $s_t(\theta) = r$. The optimal threshold is determined by

$$\omega_t(\theta_t^*)\theta_t^* = \frac{F_n(t) - F_r(t)}{v'(G_t)} l_t(\theta_t^*).$$

## 3.2 Mirrleesian taxation

In this section, we characterize the optimal non-linear income tax when the planner observes a worker's total income but does not observe the worker's type, skill choice, or labor supply. The planner can discriminate across generations.

In the appendix, we show that the necessary and sufficient conditions for implementation are the same as those that characterize a direct implementation mechanism where households declare their type $\theta$ and get assigned an allocation. Income and consumption are observable, but the government cannot observe labor, wage or skill choice. Given this informational asymmetry, the constraints that guarantee truth-telling are as follows.

The first condition is the same incentive constraint on the choice of hours worked that we have used in the static model:

$$u\left(c_t^y(\theta)\right) - \psi\left(l_t(\theta)\right) + \beta u\left(c_{t+1}^o(\theta)\right) \geq u\left(c_t^y(\theta')\right) - \psi\left(\frac{F_{s_t(\theta'),t}}{F_{s_t(\theta),t}} l_t(\theta')\right) + \beta u\left(c_{t+1}^o(\theta')\right), \tag{39}$$

for all $\theta, \theta' \in \Theta$. This *labor-supply incentive constraint* guarantees that the household chooses the assigned allocation, fixing their occupation choice.

The second condition is the incentive constraint for the choice of occupation of an individual of type $\theta$:

$$\begin{aligned} u\left(c_t^y(\theta)\right) - \psi\left(l_t(\theta)\right) + \beta u\left(c_{t+1}^o(\theta)\right) - s_t(\theta)\theta \\ \geq u\left(c_t^y(\theta')\right) - \psi\left(l_t(\theta')\right) + \beta u\left(c_{t+1}^o(\theta')\right) - s_t(\theta')\theta,. \end{aligned} \tag{40}$$

for all $\theta, \theta' \in \Theta$ and $t = 1, 2, 3...$ This *occupation-choice incentive constraint* ensures that the household chooses the assigned occupation.[31] The planning problem is to maximize (38) subject to these incentive constraints and the resource constraints (37).

[31]These constraints do not explicitly take into account the possibility that agent $\theta$ might choose an allocation $(c_t^y(\theta'), c_t^o(\theta'), y_t(\theta'))$ at a different occupational choice than $s_t(\theta')$. However, those additional constraints are redundant.

We now state two results which allow us to simplify our analysis. These same results appear in Scheuer (2014). The proofs can be found in the appendix.

**Lemma 1.** *An allocation satisfies the extensive margin incentive constraints if and only if for all t there exists* $U_{r,t}, U_{r,t} \in \mathbb{R}$ *and* $\theta^* = U_{n,t} - U_{r,t}$ *such that:*

1. *If* $\theta < \theta^*_t$*, then* $s_t(\theta) = 1$ *and* $U_t(\theta) = U_{n,t} - \theta$*;*

2. *If* $\theta > \theta^*_t$*, then* $s_t(\theta) = 0$ *and* $U_t(\theta) = U_{r,t}$*.*

This lemma allows us to simplify the incentive constraints. For an allocation to be incentive compatible, all workers that make the same skill choice should have the same utility gross of skill acquisition costs. As a result, these incentive constraints can be simplified to a cut-off rule. This rule implies that all workers with relatively low $\theta$ acquire non-routine skills, while those with high $\theta$ acquire non-routine skills.

We can do one further simplification. As the next lemma shows all workers that have the same skill choice should have the same allocation in terms of consumption and labor.

**Lemma 2.** *In the optimum, if* $s_t(\theta) = s_t(\theta')$ *then* $c^y_t(\theta) = c^y_t(\theta')$*,* $c^o_{t+1}(\theta) = c^o_{t+1}(\theta')$*, and* $l_t(\theta) = l_t(\theta')$*.*

As a result, in the optimum we need only find the allocations for agents that acquire routine skills and for the agents that acquire non-routine skills. We can find the optimum by maximizing welfare

$$\omega_0 U_0 + \sum_{t=1}^{\infty} \beta^t \left\{ \sum_{j=n,r} \Lambda_{j,t} \omega_{j,t} U_{j,t} - \int_{-\infty}^{\theta^*_t} \theta \lambda(\theta) \omega(\theta) d\theta \right\} \tag{41}$$

where $U_{j,t} \equiv u(c^y_{j,t}) - \psi(l_{j,t}) + v(G_t) + \beta(u(c^o_{j,t+1}) + v(G_{t+1}))$, $\Lambda_{n,t} \equiv \int_{-\infty}^{\theta^*_t} \lambda(\theta) d\theta$, $\Lambda_{r,t} \equiv \int_{\theta^*_t}^{\infty} \lambda(\theta) d\theta$, and $\omega_{j,t} \equiv \int_{\theta \in \Theta_{j,t}} \frac{\lambda(\theta)}{\Lambda_{j,t}} \omega(\theta) d\theta$. Subject to the constraint on $\theta^*_t$

$$\theta^*_t = U_{n,t} - U_{r,t}, \tag{42}$$

two intensive margin incentive compatibility constraints

$$u(c^y_{n,t}) + \beta u(c^y_{n,t+1}) - \psi(l_{n,t}) \geq u(c^y_{r,t}) + \beta u(c^y_{r,t+1}) - \psi\left(\frac{F_{r,t}}{F_{n,t}} l_{r,t}\right), \tag{43}$$

$$u(c^y_{r,t}) + \beta u(c^y_{r,t+1}) - \psi(l_{r,t}) \geq u(c^y_{n,t}) + \beta u(c^y_{n,t+1}) - \psi\left(\frac{F_{n,t}}{F_{r,t}} l_{n,t}\right) \tag{44}$$

and the resource constraint

$$\sum_{j=r,n} \Lambda_{j,t} c^y_{j,t} + \sum_{j=r,n} \Lambda_{j,t-1} c^o_{j,t} + G_t + \phi_t \left[X_{t+1} - (1-\delta) X_t\right] = F\left(X_t, \Lambda_{r,t} l_{r,t}, \Lambda_{n,t} l_{n,t}\right) \tag{45}$$

for $t \geq 2$, and the resource constraint at $t = 1$ which is similar but there is only one $c^o_1$.

The next proposition states results analogous to those we obtained for the static model: as long as automation is incomplete and the intensive margin incentive constraint of non-routine binds, (43), there are positive wedges in the accumulation of robots.

**Proposition 2.** *In the optimal plan, the consumption-intertemporal wedge is the same for all workers,*

$$\frac{u_c(c^y_{r,t})}{\beta u_c(c^o_{r,t+1})} = \frac{u_c(c^y_{n,t})}{\beta u_c(c^o_{n,t+1})}.$$

*Suppose furthermore that the incentive compatibility of non-routine workers at time $t+1$,* (43)*, is binding, and that of routine workers at time $t+1$,* (44)*, is not. Then, as long as routine labor hours are strictly positive ($l_{r,t+1} > 0$), the intertemporal wedge is strictly positive,*

$$\frac{u_c(c^y_{r,t})}{\beta u_c(c^o_{r,t+1})} = \frac{u_c(c^y_{n,t})}{\beta u_c(c^o_{n,t+1})} < \frac{F_{x,t+1} + (1-\delta)\phi_{t+1}}{\phi_t}$$

.

As in the static model, it may be optimal to tax robots when the hours supplied by routine workers are positive. Since robots are a form of capital, their use is taxed by creating a positive intertemporal wedge which distorts the accumulation of capital goods. Notice also that in this economy the intertemporal marginal rates of substitution of the two agents are equated. This is a consequence of the assumption that consumption and labor are separable in the utility function. The corollary shows that there is an implementation of this optimum which does not require taxes on asset returns. The tax on profits is sufficient to achieve the distortions on the intertemporal margins.

Unlike the static model, because this dynamic model features endogenous skill acquisition it is no longer clear that the intensive margin incentive constraint of non-routine workers should bind even if the government wants to redistribute income. That is because in this model, taxation affects the composition of the labor force.

The government faces two options to achieve better outcoems. The first is to redistribute income from non-routine to routine agents. By doing so the government redistributes after-tax income. We call this the *direct redistribution mechanism*, which is the driving force of our results in the static model. In a model with endogenous skill choice this comes at the cost of desincentivizing workers from acquiring non-routine skills. The second way is through an *indirect redistribution mechanism*. In this solution, the planner provides little redistribution to routine workers in order to incentive agents to acquire non-routine skills, and consequently the intensive margin incentive constraint no longer binds. Because taxes on robots are only desirable insofar as they help provide incentives on this intensive margin, in this solution there is no longer a need to tax robots and distort production. Which of these solutions turns out to be optimal is a quantitative question.

**Asymptotic balanced growth** We assume that the cost of robots declines geometrically over time, $\phi_t = \phi e^{-g_\phi t}$, which means that the dynamic problem features growth arising from investment-specific technical change as in Greenwood, Hercowitz and Krusell (1997). We assume that technological progress is exogenous and not affected by taxation.

Because our model features economic growth, we require preferences to be consistent with steady state balanced growth. This amounts to assuming that $u(\cdot)$ and $v(\cdot)$ are logarithmic functions.

**Assumption 2** (BG Preferences)**.** *The utility function takes the form* $u(c) = \log(c)$ *and* $v(G) = \chi \log(G)$*, with* $\chi > 0$*.*

These preferences have been used in different public finance applications, especially the ones featuring technical change, e.g. Ales, Kurnaz, and Sleet (2015). Recall that these preferences are also compatible the empirical evidence reviewed in Chetty (2006).

In the appendix we show how the variables in our model can be normalized to remove trends. We call this the *normalized economy*. We say that the economy is in a *steady state growth path* if the allocations of the normalized economy are constant over time, i.e. if the allocations of the economy with the normalized variables is in steady state.

In general, in dynamic optimal taxation problems long-run/steady-state allocations are a function of initial conditions. This is the case, for instance, of optimal taxation of Chamley (1986), Judd (1985), Slavík and Yazici (2014), among others. Our next proposition shows that this is not the case in our model. Because our model features overlapping generations, the inter-period link is broken, in the sense that the very long-run allocations do not affect the provision of incentives in the short run. As a consequence, if the optimal plan converges to an interior steady state, it converges to a unique one.[32]

**Lemma 3.** *Generically, for all initial conditions, there exists a unique steady state growth path of the planning problem with interior automation.*

[32] A similar argument for an Aiyagari economy in Acikgoz, Hagedorn, Holter, and Wang (2018).

This result provides a great deal of tractability to our quantitative implementation. We can compute the steady state of the normalized economy separately from the transition to that steady state.

*NOTE:* The reason why this is generically is because there are two possible steady states: one in which the intensive marign incentive compatibility constraint of non-routine workers binds and then $\theta^*_\infty = 0$, and one in which it does not and $\theta^*_\infty > 0$. Which of these is the solution depends on the distribution of $\theta$. There may be distributions for which both are a solution but those are not generic.

**Proposition 3.** *Suppose that the optimal plan is such that the allocations converge to the steady state growth path with interior automation. Then, the intertemporal wedge should converge to zero.*

**Corollary 1.** *Suppose that the optimal plan is such that the allocations converge to the interior steady state growth path. Then, the optimal plan can be decentralized with zero taxes on savings, $\tau^a_t(y) = 0$, for all $t$ and $y$, and taxes on profits which converge to zero $\tau^p_t \to 0$ as $t \to \infty$.*

Proposition 3 states that the tax on robots should asymptotically converge to zero. This result is reminiscent of the celebrated Chamley-Judd result, and stands in sharp contrast with the findings of Jones, Manuelli, and Rossi (1997) and Slavík and Yazici (2014). The reason why those authors find a role for steady-state taxes on capital is because they ignore technical progress. The proposition shows that, once skill-biased technical change is incorporated into the analysis, the steady-state supply of labor by routine agents should be zero. Because relative wages stop being relevant, there is no advantage to using production distortions in the provision of incentives. As a result, the optimal tax on robots should be zero.

### 3.3 Implementation

### 3.4 Quantitative analysis

In this section, we solve the planning problem and quantify the effects of advances in automation for the optimal taxation of robots and income.

**The status-quo economy** We calibrate a status-quo economy to match salient features on the U.S. economy. We assume that a time period corresponds to 30 years, and set $\beta = 0.3083$, which are corresponds to a 4% yearly real interest rate. We use assume that $\psi(l) = \zeta l^{1+\nu}/(1+\nu)$, and set $\nu = 1/0.75$ and set $\zeta$ such that the average labor average labor supply in the original steady state is $1/3$.

We assume that the cross-sectional distribution of $\theta$, $\lambda(\theta)$, is a Logistic distribution with parameters $\mu \in \mathbb{R}$ and $\sigma \in \mathbb{R}_{++}$:

$$\lambda(\theta) = \frac{\exp\{-(x-\mu)/\sigma\}}{\sigma\left(1+\exp\{-(x-\mu)/\sigma\}\right)^2}.$$

This is equivalent to a model in which we instead assume that the agent has a occupation-specific cost to acquire skills, and these costs are distributed according to a Gumbel distribution. This parametrization has been widely used in the applied literature on occupation choice, among other models with binary choices, e.g. Johnson and Keana (2013), and Roys and Taber (2017).

We model the taxation of labor earnings as in Heathcote, Storesletten, and Violante (2017):

$$T_t^y(y_t(\theta)) = y_t(\theta) - \lambda_t y_t(\theta)^{1-\gamma_t}$$

and we model the social security system as in Guvenen and Smith (2014)

$$T_{t+1}^o(y_t(\theta)) = -Y_t \times \begin{cases} 0.9\frac{y_t(\theta)}{Y_t}, & \text{if } \frac{y_t(\theta)}{Y_t} \leq 0.3 \\ 0.27 + 0.32\left(\frac{y_t(\theta)}{Y_t} - 0.3\right), & \text{if } 0.3 < \frac{y_t(\theta)}{Y_t} \leq 2 \\ 0.81 + 0.15\left(\frac{y_t(\theta)}{Y_t} - 2\right), & \text{if } 2 < \frac{y_t(\theta)}{Y_t} \leq 4.1 \\ 1.13, & \text{if } 4.1 < \frac{y_t(\theta)}{Y_t} \end{cases}$$

where $Y_t \equiv \int_\Theta \lambda(\theta) y_t(\theta) d\theta$. Or we could take a simplified version where we just take

$$T^o_{t+1}(y_t(\theta)) = -Y_t \left[ 0.27 + 0.32 \left( \frac{y_t(\theta)}{Y_t} - 0.3 \right) \right].$$

We normalize $A$ to one, and set $\delta = 0.9575$ which corresponds to a 10% yearly depreciation of robots.

**Mirrleesian optimal taxation** Figure 10 displays the optimal policy solution. The properties of this solution are similar to those of the static model. The maximum value of the optimal robot tax is 14 percent, which is higher than the maximum value attained in the static model (9 percent). The tax rate on robots converges to zero as the degree of automation, defined as $m_t = X_t / (X_t + \pi_r l_{r,t})$, converges to one.

Technical progress induces a fall over time in the relative productivity of routine workers. For this reason, it is optimal for the number of routine hours of work to decline over time. As routine hours fall, there is less incentive for the planner to tax robots to distort the ratio of wages and loosen the incentive constraint of non-routine workers. As routine hours converge to zero, the optimal robot tax converges to zero. This mechanism is also present in our static model.

# 4 Relation to the public finance literature

In this section we discuss how our results relate to classical results on production efficiency and taxation of capital in the public finance literature.

**Relating our results to Diamond and Mirrlees (1971)** Our results stand in sharp contrast to the celebrated Diamond and Mirrlees (1971) result that an optimal tax system should ensure efficiency in production and therefore leave intermediate goods untaxed. In our framework, this property would imply that the tax on robots should be zero.

At the heart of the failure of the Diamond and Mirrlees (1971) intermediate-good theorem in our model is the fact that the government cannot discriminate between the two types of workers. If tax functions could be worker specific, production efficiency would be recovered in our model. To see this result, consider type-specific tax functions of the form used by Heathcote et al. (2018) with different tax levels, $\lambda_r$ and $\lambda_n$, but with the same progressivity parameter

$$T_i\left(w_i l_i\right) = w_i l_i - \lambda_i \left(w_i l_i\right)^{1-\gamma}.$$

In this case, household optimality requires

$$\frac{-u_l(c_j, l_j) l_j}{u_c(c_j, l_j)} = \lambda_j \left(1-\gamma\right) \left(w_j l_j\right)^{1-\gamma}, \quad \text{and} \quad c_j = \lambda_j \left(w_j l_j\right)^{1-\gamma}.$$

Given that the planner can choose $\lambda_r$ and $\lambda_n$ to target each marginal rate of substitution independently, the only constraints faced by the planner are the resource constraint (20) which can be written as

$$\pi_r c_r + \pi_n c_n + G \leq \frac{\tau_x + \alpha}{\alpha(1+\tau_x)^{1/\alpha}} \frac{\alpha A^{1/\alpha}(1-\alpha)^{\frac{1-\alpha}{\alpha}}}{\phi^{\frac{1-\alpha}{\alpha}}} \pi_n l_n + \phi \pi_r l_r,$$

and the implementability conditions

$$u_c(c_j, l_j) c_j + \frac{u_l(c_j, l_j) l_j}{1-\gamma} = 0, \text{ for } j = r, n.$$

These three conditions are necessary and sufficient for an equilibrium. Recall that the term $\left(\tau_x + \alpha\right) / \alpha(1+\tau_x)^{1/\alpha} \leq 1$, and is strictly less than one if $\tau_x \neq 0$.

The robot tax only affects directly the resource constraint and not the implementability conditions. Since the robot tax does not interfere with incentives, it is chosen to maximize output for given levels of hours worked. This objective is achieved by not distorting production, setting $\tau_x = 0$.

When the tax system requires that all workers face the same income-tax function ($\lambda_r = \lambda_n$), the planner must satisfy the following additional implementability constraint

$$\frac{c_r}{c_n} = \left(\frac{w_r l_r}{w_n l_n}\right)^{1-\gamma}. \tag{46}$$

The value of $\tau_x$ no longer appears only in the resource constraint; it also appears in equation (46) because the wage ratio is a function of $\tau_x$. To relax restriction (46), it might be optimal to choose values of $\tau_x$ that are different from zero. This result depends crucially on the fact that different labor types interact differently with the intermediate good, which means that distorting the use of intermediate goods affects in different ways the wage rates of routine and non-routine workers. If the production function was weakly separable in labor types and intermediate inputs, the wage ratio would be independent of the usage of intermediate inputs and production efficiency would be optimal. In our model, robots are substitutes of routine workers and complements of non-routine workers. A tax on robots decreases the wage rate of non-routine workers and increases the wage rate of routine workers. This property implies that it can be optimal to use robot taxes.

**Relating our results to Atkinson and Stiglitz (1976)** Our result that in the Mirrleesian optimal taxation problem production efficiency is not optimal stands in contrast with the well-known result in Atkinson and Stiglitz (1976) that, for preferences that are separable in commodities and leisure, uniform commodity taxation is optimal. Since uniform taxation can be interpreted as production efficiency, their result seems to contradict ours.

The difference between our results and those of Atkinson and Stiglitz (1976) stem from the determinants of worker productivity. In Atkinson and Stiglitz (1976), workers' productivities are exogenous. In our setup workers' productivity are endogenous so, it may be optimal to deviate from production efficiency to induce changes in those productivities. In particular, by taxing robots the Mirrleesian planner is able to change pre-tax wages through general-equilibrium effects, relaxing the incentive constraint, and improving welfare.[33]

Naito (1999) shows that uniform taxation may not be optimal in an economy in which the intensity of high- and low-skilled workers in production varies across goods. This form of production non-separability implies that commodities interact differently with different agent types and, as a result, it might be optimal to deviate from uniform commodity taxation.[34] Similarly, in our model the assumption that production is not separable in the use of robots and the two labor types is key to generate deviations from production efficiency.

The intuition for the importance of general-equilibrium effects of taxation on wages and prices is the same we emphasized in our discussion of proposition 1. Because the planner does not know the type of the agent and only observes income, it is restricted to use incentive-compatible tax systems. Since different types interact differently with the intermediate good, distorting production decisions may help in the screening process. To see this property, it is useful to write the incentive constraint as: $u(c_i, l_i) \geq u\left(c_j, w_j l_j / w_i\right)$. Crucially, this incentive constraint involves the wage ratio. Whenever the taxation of intermediate goods affects this ratio, production efficiency may no longer be optimal. When intermediate goods are not separable in production from the two labor types, taxing intermediate goods affects the wage ratio and it might be optimal to distort production.

The importance of general-equilibrium effects of taxes on wages in shaping the optimal tax policy was originally emphasized by Stiglitz (1982) and Stern (1982) in a Mirrlees (1971) environment. Mirrlees assumes that production is linear in labor, so taxation does not affect wages through general-equilibrium effects. Stiglitz (1982) and Stern (1982) show that when production is not linear in labor, the optimal tax

[33] An important assumption is that when workers imitate others, they retain their productivity. See Scheuer and Werning (2016) for a discussion.

[34] Jacobs (2015) shows that production efficiency is generally not optimal in a model where commodity prices are exogenous but wages are not. In his model, goods are produced with commodities and labor according to production functions that are worker specific. Taxation of commodities has a differential impact on the marginal productivities and wages of the different workers.

schedule is more regressive than in the Mirrlees model and the top marginal income tax is negative instead of zero.[35] The reason for this result is that it is optimal to encourage high-skilled workers to exert more effort so as to reduce their relative wages, making their incentive constraint easier to satisfy.

# 5 Conclusions

Our analysis suggests that without changes to the current U.S. tax system, a sizable fall in the costs of automation would lead to a massive rise in income inequality. Even though routine workers keep their jobs, their wages fall to make them competitive with the possibility of automating production.

Income inequality can be reduced by raising the marginal tax rates paid by high-income individuals and by taxing robots to raise the wages of routine workers. But this solution involves a substantial efficiency loss. A Mirrleesian optimal income tax can reduce inequality at a smaller efficiency cost than the variants of the U.S. tax system discussed above, coming close to the levels of social welfare obtained in the first-best allocation.

An alternative, less ambitious, approach is to amend the tax system to include a transfer that is independent of income. The desirability of this type of universal basic income system has been debated since Thomas More proposed it in his 1516 book, Utopia. With this transfer in place, it is optimal in our model to tax robots for values of the automation cost that lead to partial automation. For values of the automation cost that lead to full automation, it is not optimal to tax robots. Routine workers lose their jobs and live off government transfers, just like in Kurt Vonnegut's "Player Piano."

[35] Rothschild and Scheuer (2013) generalize the results of Stern (1982) and Stiglitz (1982) to an environment in which occupational choice is endogenous and there is a continuous distribution of agent types.

# 6 References


Acemoglu, Daron and David Autor. "Skills, tasks and technologies: implications for employment and earnings." in *Handbook of Labor Economics*, vol. 4B, edited by Orley Ashenfelter and David Card, chap. 12. Elsevier, (2011): 1043-1171.

Acemoglu, Daron and Pascual Restrepo. "Robots and Jobs: evidence from US labor markets." mimeo, Massachusetts Institute of Technology (2018a).

Acemoglu, Daron and Pascual Restrepo. "Automation and New Tasks: The Implications of the Task Content of Technology for Labor Demand." mimeo, Massachusetts Institute of Technology (2018b).

Acikgoz, Omer T., Marcus Hagedorn, Hans A. Holter, and Yikai Want. "The Optimum Quantity of Capital and Debt." working paper (2018).

Adao, Rodrigo, Martin Beraja, Nitya Pandalai-Nayar. "Technological Transitions with Skill Heterogeneity Across Generations." working paper (2019).

Ales, Lawrence, Musab Kurnaz, and Christopher Sleet. "Technical change, wage inequality, and taxes." *American Economic Review*, 105(10) (2015): 3061-3101.

Atkinson, Anthony Barnes, and Joseph E. Stiglitz. "The design of tax structure: direct versus indirect taxation." *Journal of Public Economics* 6.1-2 (1976): 55-75.

Autor, David H., David Dorn, and Gordon H. Hanson. "The China Syndrome: Local Labor Market Effects of Import Competition in the United States" *The American Economic Review*, Vol. 103, No. 6 (2013), 2121-2168.

Autor, David H., Frank Levy, and Richard J. Murnane. "The skill content of recent technological change: An Empirical Exploration," *The Quarterly Journal of Economics*, Vol. 118, No. 4 (2003), pp. 1279-1333.

Caselli, Francesco. "Technological Revolutions." *The American Economic Review*, Vol. 89, No. 1 (1999), pp. 78-102.

Chamley, Christophe. "Optimal taxation of capital income in general equilibrium with infinite lives." *Econometrica* (1986): 607-622.

Chari, V. V., and Hugo Hopenhayn. "Vintage Human Capital, Growth, and the Diffusion of New Technology." *Journal of Political Economy*, Vol. 99, No. 6 (1991), pp. 1142-1165.

Chari, V. V., Juan Pablo Nicolini, and Pedro Teles. "Optimal capital taxation revisited," manuscript, Federal Reserve Bank of Minneapolis, (2018).

Chen, Yiqun G. "The Rise of the Machines and the U.S. Labor Market," manuscript, (2017)

Cortes, Guido Matias, Nir Jaimovich, and Henry E. Siu. "Disappearing routine jobs: who, how, and why? " *Journal of Monetary Economics*, Volume 91, (2017): 69-87.

Diamond, Peter A., and James A. Mirrlees. "Optimal taxation and public production I: production efficiency." *The American Economic Review* 61.1 (1971): 8-27.

Farhi, Emmanuel, and Iván Werning. "Insurance and Taxation over the Life Cycle." *The Review of Economic Studies*, (2013) 810(2), 596-635.

Golosov, Mikhail, Narayana Kocherlakota, and Aleh Tsyvinski. "Optimal Indirect and Capital Taxation." *Review of Economic Studies,* Vol. 70 (2003), 569-587.

Gomes, Renato, Jean-Marie Lozachmeur, and Alessandro Pavan. "Differential taxation and occupational choice." *The Review of Economic Studies* (2017): rdx022.

Greenwood, Jeremy, Zvi Hercowitz, and Per Krusell. "Long-run implications of investment-specific technological change" *American Economic Review* 87.3 (1997): 342-362.

Guesnerie, Roger. *A contribution to the pure theory of taxation*. Cambridge-UK: Cambridge University Press, 1995.

Guvenen, Fatih, and Anthony A. Smith. "Inferring labor income risk and partial insurance from economic choices." *Econometrica* 82, no. 6 (2014): 2085-2129.

Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L. Violante. "Optimal tax progressivity: An analytical framework." *The Quarterly Journal of Economics* 132.4 (2017): 1693-1754.

Heathcote, Jonathan, Hitoshi Tsujiyama. 2019. "Optimal Income Taxation: Mirrlees meets Ramsey." working paper.

Jacobs, Bas "Optimal inefficient production," manuscript, Erasmus University, Rotterdam (2015).

Johnson, Matthew, and Michael P. Keane. "A dynamic equilibrium model of the US wage structure, 1968?1996." *Journal of Labor Economics* 31, no. 1 (2013): 1-49.

Jones, Larry E., Rodolfo E. Manuelli, and Peter E. Rossi. "On the optimal taxation of capital income." *Journal of Economic Theory* 73, No.1 (1997): 93-117.

Judd, Kenneth L. "Redistributive taxation in a simple perfect foresight model." *Journal of public Economics* 28, no. 1 (1985): 59-83.

Kocherlakota, Narayana R. 2005. "Zero Expected Wealth Taxes: A Mirrlees Approach to Dynamic Optimal Taxation," *Econometrica*, 73(5), 1587-1621.

Kocherlakota, Narayana R. *The New Dynamic Public Finance.* (2010) Princeton, NJ: Princeton University Press.

Krusell, Per, Lee E. Ohanian, Jose-Victor Rios-Rull, Giovanni L. Violante. "Capital-skill complementarity and inequality: A macroeconomic analysis." *Econometrica* 68.5 (2000): 1029-1053.

Mirrlees, James A. "An exploration in the theory of optimum income taxation." *The Review of Economic Studies* 38.2 (1971): 175-208.

Naito, Hisahiro. "Re-examination of uniform commodity taxes under a non-linear income tax system and its implications for production efficiency." *Journal of Public Economics* 71 (1999): 165-188.

Ramsey, Frank P. "A contribution to the theory of taxation." *The Economic Journal* 37.145 (1927): 47-61.

Rothschild, Casey and Florian Scheuer. "Redistributive taxation in the Roy model," *The Quarterly Journal of Economics*, Volume 128, Issue 2, (2013): 623–668.

Roys, Nicolas, and Christopher Taber. "Skills Prices, Occupations and Changes in the Wage Structure for Low Skilled Men." working paper (2017).

Sachs, Dominik, Aleh Tsyvinsky, and Nicolas Werquin. 2019. "Nonlinear Tax Incidence and Optimal Taxation in General Equilibrium." NBER Working Paper No. 22646.

Saez, Emmanuel. "Direct or indirect tax instruments for redistribution: short-run versus long-run." *Journal of Public Economics* 88 (2004): 503-518.

Scheuer, Florian. "Entrepreneurial taxation with endogenous entry." *American Economic Journal: Economic Policy* 6:2 (2014): 126-163.

Scheuer, Florian and Iván Werning "Mirrlees meets Diamond-Mirrlees," Working paper No. w22076 (2016), National Bureau of Economic Research.

Slavík, Ctirad, and Hakki Yazici. "Machines, buildings, and optimal dynamic taxes." *Journal of Monetary Economics* 66 (2014): 47-61.

Stern, Nicholas. "Optimum taxation with errors in administration." *Journal of Public Economics*, 17:2 (1982): 181 - 211.

Stiglitz, Joseph. "Self-selection and Pareto efficient taxation." *Journal of Public Economics*, 17:2 (1982): 213-240.

Werning, Iván. "Optimal Fiscal Policy with Redistribution." *The Quarterly Journal of Economics*, Vol. 122, No.3 (2007), pp. 925-967.

# A Appendix

## A.1 The first-best allocation

We define the first-best allocation in this economy as the solution to an utilitarian welfare function, absent informational constraints. This absence implies that the planner can perfectly discriminate among agents and enforce any allocation. The optimal plan solves the following problem

$$W = \max \omega_r \pi_r \left[ u(c_r, l_r) + v(G) \right] + \omega_n \pi_n \left[ u(c_n, l_n) + v(G) \right].$$

$$\pi_r c_r + \pi_n c_n + G \leq A \left[ \int_0^m x_i^\rho di + \int_m^1 n_i^\rho di \right]^{\frac{1-\alpha}{\rho}} (\pi_n l_n)^\alpha - \phi \int_0^m x_i di, \quad [\mu],$$

$$\int_m^1 n_i di = \pi_r l_r, \quad [\eta].$$

The first-order conditions with respect to $n_i$ and $x_i$ are

$$\mu(1-\alpha) A \left[ \int_0^m x_i^\rho di + \int_m^1 n_i^\rho di \right]^{\frac{1-\alpha}{\rho}-1} (\pi_n l_n)^\alpha n_i^{\rho-1} = \eta, \quad \forall i \in (m, 1]$$

$$(1-\alpha) A \left[ \int_0^m x_i^\rho di + \int_m^1 n_i^\rho di \right]^{\frac{1-\alpha}{\rho}-1} (\pi_n l_n)^\alpha x_i^{\rho-1} = \phi, \quad \forall i \in [0, m].$$

The first equation implies that the marginal productivity of routine labor should be constant across the activities that use routine labor. This property means that $(1-m)\, n_i = \pi_r l_r$ for $i \in (m, 1]$ and $n_i = 0$, otherwise. The same property applies to robots used in the activities where they are used, $x_i = x$ for $i \in [0, m]$ and $x_i = 0$, otherwise.

To characterize the optimal allocations we replace $n_i$ and $x_i$ in the planner's problem, which can be rewritten as

$$W = \max \omega_r \pi_r \left[ u(c_r, l_r) + v(G) \right] + \omega_n \pi_n \left[ u(c_n, l_n) + v(G) \right].$$

$$\pi_r c_r + \pi_n c_n + G \leq A \left[ m x^\rho + (1-m) \left( \frac{\pi_r l_r}{1-m} \right)^\rho \right]^{\frac{1-\alpha}{\rho}} (\pi_n l_n)^\alpha - \phi m x, \quad [\mu].$$

The first-order conditions with respect to $x$ and $m$ are, respectively,

$$(1-\alpha) A \left[ m x^\rho + (1-m) \left( \frac{\pi_r l_r}{1-m} \right)^\rho \right]^{\frac{1-\alpha}{\rho}-1} N_n^\alpha x^{\rho-1} = \phi,$$

$$\frac{1-\alpha}{\rho} A \left[ m x^\rho + (1-m) \left( \frac{\pi_r l_r}{1-m} \right)^\rho \right]^{\frac{1-\alpha}{\rho}-1} N_n^\alpha \left[ x^\rho - (1-\rho) \left( \frac{\pi_r l_r}{1-m} \right)^\rho \right] = \phi x.$$

The ratio of these two equations implies that if automation is positive, $m \in (0,1)$, then $x = \pi_r l_r/(1-m)$. Using this condition, we obtain

$$W = \max \omega_r \pi_r \left[u(c_r, l_r) + v(G)\right] + \omega_n \pi_n \left[u(c_n, l_n) + v(G)\right].$$

$$\pi_r c_r + \pi_n c_n + G \leq A \left(\frac{\pi_r l_r}{1-m}\right)^{1-\alpha} (\pi_n l_n)^{\alpha} - \phi m \frac{\pi_r l_r}{1-m}, \quad [\mu].$$

The first-order condition with respect to the level of automation implies that

$$(1-\alpha) A \frac{1}{(1-m)^{2-\alpha}} (\pi_r l_r)^{1-\alpha} (\pi_n l_n)^{\alpha} - \phi \frac{\pi_r l_r}{(1-m)^2} = 0 \Leftrightarrow m = 1 - \left[\frac{\phi}{A(1-\alpha)}\right]^{1/\alpha} \frac{\pi_r l_r}{\pi_n l_n},$$

provided that $m$ is interior. Then,

$$m = \max \left\{1 - \left[\frac{\phi}{A(1-\alpha)}\right]^{1/\alpha} \frac{N_r}{N_n}, 0\right\}.$$

Furthermore, the first-order conditions with respect to $c_r$, $c_n$, $l_r$, $l_n$, and $G$ are

$$\omega_r u_c(c_r, l_r) = \mu,$$

$$\omega_n u_c(c_n, l_n) = \mu,$$

$$\omega_r u_l(c_r, l_r) \geq \frac{\mu}{\pi_r l_r}(1-\alpha)(1-m)Y,$$

$$\omega_n u_l(c_n, l_n) = \mu \frac{\alpha Y}{\pi_n l_n},$$

$$g'(G) = \mu.$$

The first-order condition with respect to $N_r$ is presented with inequality, because the constraint $N_r \geq 0$ may bind when automation costs are low. The combination of the first two equations implies that

$$\omega_r u_c(c_r, l_r) = \omega_n u_c(c_n, l_n).$$

The optimal marginal rates of substitution are given by the combination of the marginal utility of consumption and leisure for each individual

$$\frac{u_l(c_r, l_r)}{u_c(c_r, l_r)} \geq (1-\alpha)(1-m)\frac{Y}{\pi_r l_r},$$

$$\frac{u_l(c_n, l_n)}{u_c(c_n, l_n)} = \alpha \frac{Y}{\pi_n l_n}.$$

Finally, from the first-order conditions for $G$ and $c_r$ it follows that

$$g'(G) = \omega_r u'(c_r). \tag{47}$$

## A.2 Necessity and sufficiency in the static model

Household optimality implies that the utility associated with the bundle of consumption and income assigned to agent $j$, $\{c_j, l_j\}$, must be at least as high as the utility associated with any other bundle $\{c, l\}$ that satisfies the budget constraint $c \leq w_j l - T(w_j l)$, implying that $u(c_j, l_j) \geq u(c, l)$. In particular, routine workers must prefer their bundle, $\{c_r, l_r\}$, to the bundle that they would get if they pretended to be non-routine workers while keeping the routine wage, $\{c_n, w_n l_n / w_r\}$. Similarly, non-routine workers must prefer their bundle, $\{c_n, l_n\}$, to the bundle they would get if they pretended to be routine workers, $\{c_r, w_r l_r / w_n\}$. These requirements correspond to the two IC constraints, (21), and (22), so these conditions are necessary.

We show in the Appendix that equation (20) is necessary by combining the first-order conditions to the firms' problems with the resource constraint, (11). In addition, we show that conditions (20), (21), and (22), are also sufficient. To see that equations (21) and (22) summarize the household problem, note that it is possible to choose a tax function such that agents prefer the bundle $\{c_j, l_j\}$ to any other bundle. For example, the government could choose a tax function that sets the agent's after-tax income to zero for any choice of $w_j l$ different from $w_j l_j$, $j = r, n$. These results are summarized in the following proposition.

**Lemma 4.** *Equations* (20), (22) *and* (21) *characterize the set of implementable allocations. These conditions are necessary and sufficient for a competitive equilibrium.*

In an equilibrium, robot producers set the price of robots equal to their marginal cost

$$p_i = \phi. \tag{48}$$

Optimality for final goods producers implies that

$$x_i = \begin{cases} \frac{\pi_r l_r}{1-m}, & i \in [0, m], \\ 0, & \text{otherwise} \end{cases} \tag{49}$$

$$n_i = \begin{cases} \frac{\pi_r l_r}{1-m}, & i \in (m, 1], \\ 0, & \text{otherwise} \end{cases} \tag{50}$$

$$m = \max\left\{1 - \left[\frac{(1+\tau_x)\phi}{(1-\alpha)A}\right]^{1/\alpha} \frac{\pi_r l_r}{\pi_n l_n}, 0\right\}, \tag{51}$$

$$Y = A\left[\int_0^m x_i^\rho di + \int_m^1 n_i^\rho di\right]^{\frac{1-\alpha}{\rho}} (\pi_n l_n)^\alpha, \tag{52}$$

$$w_r = (1-\alpha)(1-m)\frac{Y}{\pi_r l_r}, \tag{53}$$

$$w_n = \alpha \frac{Y}{\pi_n l_n}. \tag{54}$$

The resource constraint is

$$\pi_r c_r + \pi_n c_n + G \leq Y - \int_0^m \phi x_i, \tag{55}$$

We can let equation (48) define the price of robots, equation (49) define $x_i$, equations (50), (51) and (52) determine $n_i$, $m$, and $Y$, respectively. Assuming that $m$ is interior, the wage equations (53) and (54) can be written as (14) and (15). These equations can be used to solve for the equilibrium wage rates. Combining the results above, we can write the resource constraint as

$$\pi_r c_r + \pi_n c_n + G \leq \alpha \frac{A^{1/\alpha}(1-\alpha)^{\frac{1-\alpha}{\alpha}}}{[(1+\tau_x)\phi]^{\frac{1-\alpha}{\alpha}}} \frac{\tau_x + \alpha}{\alpha(1+\tau_x)} \pi_n l_n + \phi \pi_r l_r.$$

Replacing the wage rates we can write

$$\pi_r c_r + \pi_n c_n + G \leq \pi_n w_n l_n \frac{\tau_x + \alpha}{\alpha(1+\tau_x)} + \frac{\pi_r w_r l_r}{1+\tau_x}. \tag{56}$$

This derivation makes it clear that the resource constraint (56) summarize the equilibrium conditions of the production side of the economy.

Household optimality requires that

$$u\left(c_j, l_j\right) \geq u\left(c, l\right), \quad \forall (c,l) : c \leq w_j l - T(w_j l).$$

The following incentive constraint are necessary constraints

$$u\left(c_n, l_n\right) \geq u\left(c_r, \frac{w_r}{w_n} l_r\right)$$
$$u\left(c_r, l_r\right) \geq u\left(c_n, \frac{w_n}{w_r} l_n\right).$$

These are also sufficient conditions, because the planner can set the tax schedule $T(\cdot)$ such that for all $Y \notin \{Y_n, Y_r\}$ the allocation is worse for both agents than their respective allocation. This is done by setting

$$T(y) = y - \max\left\{c | u(c_i, l_i) \geq u\left(c, \frac{y}{w_i}\right), \text{for } i = r, n\right\}.$$

Since the government can choose an arbitrary tax function, it is only bound by the incentive constraints which characterize the informational problem. This property means that the income tax function that is assumed here to implement the optimal allocation is without loss of generality. Any other implementation would at least have to satisfy the same two incentive constraints.

## A.3 Proof of proposition 1

The allocations solve the original optimization problem, or equivalently they solve

$$W(\tau_x) = \max \pi_r \omega_r u(c_r, l_r) + \pi_n \omega_n u(c_n, l_n) + v(G)$$

subject to

$$\begin{aligned}
&[\eta_r \pi_r] \;\; u(c_r, l_r) \geq u\left(c_n, \frac{w_n}{w_r} l_n\right), \\
&[\eta_n \pi_n] \;\; u(c_n, l_n) \geq u\left(c_r, \frac{w_r}{w_n} l_r\right), \\
&[\mu] \;\; \pi_r c_r + \pi_n c_n + G \leq \pi_n w_n l_n \frac{\tau_x + \alpha}{\alpha(1+\tau_x)} + \pi_r \frac{w_r l_r}{1+\tau_x}.
\end{aligned}$$

Assume that the routine IC constraint does not bind, then $\eta_r = 0$. The envelope condition is

$$W'(\tau_x) = -\eta_n \pi_n u_l\left(c_r, \frac{w_r}{w_r} l_r\right) \frac{d\log(w_r/w_n)}{d\log(1+\tau_x)} \frac{1}{1+\tau_x} \frac{w_r l_r}{w_n} + \mu \left[ \begin{array}{c} \pi_n w_n l_n \frac{\tau_x+\alpha}{\alpha(1+\tau_x)^2} \left[\frac{d\log w_n}{d\log(1+\tau_x)} + \frac{1-\alpha}{\tau_x+\alpha}\right] \\ +\pi_r \frac{w_r l_r}{(1+\tau_x)^2} \left[\frac{d\log w_r}{d\log(1+\tau_x)} - 1\right] \end{array} \right].$$

Using the wages we have that

$$\begin{aligned}
w_r &= \phi(1+\tau_x) \Rightarrow \frac{d\log w_r}{d\log(1+\tau_x)} = 1, \\
w_n &= \alpha \frac{A^{1/\alpha}(1-\alpha)^{\frac{1-\alpha}{\alpha}}}{[(1+\tau_x)\phi]^{\frac{1-\alpha}{\alpha}}} \Rightarrow \frac{d\log w_n}{d\log(1+\tau_x)} = -\frac{1-\alpha}{\alpha}, \\
\frac{w_r}{w_n} &= \frac{[(1+\tau_x)\phi]^{\frac{1}{\alpha}}}{\alpha A^{1/\alpha}(1-\alpha)^{\frac{1-\alpha}{\alpha}}} \Rightarrow \frac{d\log w_r/w_n}{d\log(1+\tau_x)} = \frac{1}{\alpha}.
\end{aligned}$$

Plugging these into the envelope condition we obtain

$$\begin{aligned}
W'(\tau_x) &= -\eta_n \pi_n u_l\left(c_r, \frac{w_r}{w_n} l_r\right) \frac{1}{\alpha(1+\tau_x)} \frac{w_r l_r}{w_n} + \mu \pi_n w_n l_n \frac{\tau_x+\alpha}{\alpha(1+\tau_x)^2}\left[-\frac{1-\alpha}{\alpha} + \frac{1-\alpha}{\tau_x+\alpha}\right] \\
&= \frac{1}{\alpha(1+\tau_x)}\left[-\eta_n u_l\left(c_r, \frac{w_r}{w_n} l_r\right) \frac{w_r l_r}{w_n} - \mu \pi_n w_n l_n \frac{\tau_x}{1+\tau_x} \frac{1-\alpha}{\alpha}\right]
\end{aligned}$$

Because $\mu > 0$ then if $\tau_x \leq 0$ we obtain that

$$W'(\tau_x) > 0,$$

so that the planner always improves by marginally increasing $\tau_x$. Furthermore, since optimality implies that $W'(\tau_x) = 0$ then the optimal tax on robots verifies that

$$\frac{\tau_x}{1+\tau_x} = \frac{\alpha}{1-\alpha} \frac{\eta_n \left( -u_l \left( c_r, \frac{w_r}{w_n} l_r \right) \frac{w_r l_r}{w_n} \right)}{\mu w_n l_n}$$

The first order condition with respect to $l_r$ implies that

$$-\frac{\eta_n}{\mu} u_l \left( c_r, \frac{w_r}{w_r} l_r \right) \frac{w_r l_r}{w_n} = \frac{\widetilde{\omega}_r \pi_r u_l \left( c_r, l_r \right) l_r + \frac{\pi_r w_r l_r}{1+\tau_x}}{\pi_n} = \frac{\pi_r \phi l_r}{\pi_n} \left[ 1 - \frac{\widetilde{\omega}_r \left( -u_l \left( c_r, l_r \right) \right)}{\phi} \right]$$

where $\widetilde{\omega}_r = \omega_r / \mu$. Replacing this in the optimal condition for $\tau_x$ we obtain

$$\frac{\tau_x}{1+\tau_x} = \frac{\alpha}{1-\alpha} \frac{\pi_r \phi l_r}{\pi_n w_n l_n} \left[ 1 - \frac{\widetilde{\omega}_r \left( -u_l \left( c_r, l_r \right) \right)}{\phi} \right].$$

### A.3.1 The full automation case ($m = 1$, $l_r = 0$)

If the optimal plan features $l_r = 0$ then it must be that $l_n > 0$. This result implies that $\psi = 0$. From the envelope condition we can see that

$$W'(\tau_x) = -\frac{\mu}{\alpha \left( 1 + \tau_x \right)} \pi_n w_n l_n \frac{\tau_x}{1+\tau_x} \frac{1-\alpha}{\alpha} = 0 \Leftrightarrow \tau_x = 0. \tag{57}$$

## A.4 Model with simple taxes

The competitive equilibrium for this economy is characterized by the following set of equations

$$c_n = \lambda \left( w_n l_n \right)^{1-\gamma} = \lambda \left( \frac{\alpha Y}{\pi_r} \right)^{1-\gamma}, \tag{58}$$

$$\frac{-u_l(c_n, l_n) l_n}{u_c(c_n, l_n)} = \lambda \left( 1 - \gamma \right) \left( w_n l_n \right)^{1-\gamma}, \tag{59}$$

$$c_r = \lambda \left( w_r l_r \right)^{1-\gamma} = \lambda \left( \frac{(1-\alpha)(1-m) Y}{\pi_r} \right)^{1-\gamma}, \tag{60}$$

$$\frac{-u_l(c_r, l_r) l_r}{u_c(c_r, l_r)} = \lambda \left( 1 - \gamma \right) \left( w_r l_r \right)^{1-\gamma}, \tag{61}$$

$$m = \max \left\{ 1 - \left( \frac{\phi(1+\tau_x)}{(1-\alpha) A} \right)^{1/\alpha} \frac{\pi_r l_r}{\pi_n l_n}, 0 \right\}, \tag{62}$$

$$Y = \left\{ \begin{array}{c} A \left( \pi_r l_r \right)^{1-\alpha} \left( \pi_n l_n \right)^{\alpha}, \text{ if } m = 0 \\ \frac{w_n}{\alpha} \pi_n l_n, \text{ if } m > 0 \end{array} \right\} \tag{63}$$

$$\pi_n c_n + \pi_r c_r + G \leq \begin{Bmatrix} A\left(\pi_r l_r\right)^{1-\alpha}\left(\pi_n l_n\right)^{\alpha}, & \text{if } m = 0 \\ w_n \pi_n l_n \frac{\tau_x+\alpha}{\alpha(1+\tau_x)} + \frac{w_r \pi_r l_r}{1+\tau_x}, & \text{if } m > 0 \end{Bmatrix}, \tag{64}$$

where $w_r$ and $w_n$ are given by (15) and (14), respectively.

Taking the ratio between equations (58) and (60), we can see that a necessary condition is

$$\frac{c_r}{c_n} = \left[\frac{(1-\alpha)(1-m)}{\alpha}\frac{\pi_n}{\pi_r}\right]^{1-\gamma} \Leftrightarrow c_r = c_n\left[\frac{(1-\alpha)}{\alpha}\frac{\pi_n}{\pi_r}\left(\frac{\phi(1+\tau_x)}{(1-\alpha)A}\right)^{1/\alpha}\frac{\pi_r l_r}{\pi_n l_n}\right]^{1-\gamma}. \tag{65}$$

The conditions (**??**), (**??**), (**??**) and (20) are necessary and sufficient for an interior automation equilibrium in terms the allocations $\{c_r, l_r, c_n, l_n, G\}$ and the tax parameters $\{\tau_x, \gamma\}$. They are necessary because they follow from the equilibrium conditions. They are sufficient because, given a solution for $\{c_r, l_r, c_n, l_n, G\}$ and $\{\tau_x, \gamma\}$ which satisfies the constraints, the other remaining conditions can be satisfied by the choice of the remaining variables. In particular, equations (15) and (14) can be satisfied by the choice of $w_n$ and $w_r$, respectively. We can set $\lambda$ such that

$$\lambda = \frac{1}{(1-\gamma)\left(w_n l_n\right)^{1-\gamma}}\frac{-u_l(c_n, l_n) l_n}{u_c(c_n, l_n)},$$

which satisfies (59). This choice of $\lambda$ combined with (**??**) also satisfies (58). Choosing $\lambda$ in this way and combined with (**??**) implies that (61) is satisfied. Satisfying (61) with this choice of $\lambda$ also implies that (60) is satisfied. The conditions (62) and (63) are used to solve for $m$ and $Y$. The condition (64) is the same as (20).

We now derive equation (**??**). The Ramsey planner solves the following problem

$$\max \omega_r \pi_r u(c_r, l_r) + \omega_n \pi_n u(c_n, l_n) + v(G),$$

subject to

$$\left[\frac{\eta}{c_r}\right] \quad c_r = c_n\left[\frac{(1-\alpha)\left(\frac{\phi(1+\tau_x)}{(1-\alpha)A}\right)^{1/\alpha}\frac{l_r}{l_n}}{\alpha}\right]^{1-\gamma}$$

$$[\lambda_r] \;\; u_c(c_r, l_r) c_r + \frac{u_l(c_r, l_r) l_r}{1-\gamma} = 0,$$

$$[\lambda_n] \;\; u_c(c_n, l_n) c_n + \frac{u_l(c_n, l_n) l_n}{1-\gamma} = 0,$$

$$[\mu] \;\; \pi_r c_r + \pi_n c_n + G \leq w_n \pi_n l_n \frac{\tau_x+\alpha}{\alpha(1+\tau_x)} + \phi \pi_r l_r.$$

The first-order condition with respect to $\tau_x$ is given by

$$
\begin{aligned}
0 &= \frac{\eta}{c_r} c_n \left[ \frac{(1-\alpha)\left(\frac{\phi(1+\tau_x)}{(1-\alpha)A}\right)^{1/\alpha} \frac{l_r}{l_n}}{\alpha} \right]^{1-\gamma} \frac{1-\gamma}{\alpha(1+\tau_x)} + \mu \left[ \frac{dw_n}{d\tau_x} \pi_n l_n \frac{\tau_x + \alpha}{\alpha(1+\tau_x)} + w_n \pi_n l_n \frac{1-\alpha}{\alpha(1+\tau_x)^2} \right] \\
0 &= \frac{\eta}{c_r} c_r \frac{1-\gamma}{\alpha(1+\tau_x)} + \mu \frac{w_n \pi_n l_n}{\alpha(1+\tau_x)^2} \left[ \frac{d \log w_n}{d \log (1+\tau_x)} (\tau_x + \alpha) + 1 - \alpha \right] \\
0 &= \frac{\eta}{c_r} c_r (1-\gamma) + \mu \frac{w_n \pi_n l_n (1-\alpha)}{(1+\tau_x)} \left[ -\frac{\tau_x + \alpha}{\alpha} + 1 \right] \\
\eta (1-\gamma) &= \mu w_n \pi_n l_n \frac{1-\alpha}{\alpha} \frac{\tau_x}{1+\tau_x} \Leftrightarrow \frac{\tau_x}{1+\tau_x} = \frac{\alpha}{1-\alpha} \frac{\eta (1-\gamma)}{\mu w_n \pi_n l_n}.
\end{aligned}
$$

## A.5 Simple taxes with a lump-sum transfer

The conditions are necessary as they follow from manipulations of the necessary conditions for an equilibrium. Sufficiency is established as follows. Let $\{c_r, l_r, c_n, l_n, G\}$ and $\{\tau_x, \gamma, \omega\}$ denote some allocation that satisfies the conditions

$$u_c(c_r, l_r)(c_r - \Omega) + \frac{u_l(c_r, l_r) l_r}{1-\gamma} = 0,$$

$$u_c(c_n, l_n)(c_n - \Omega) + \frac{u_l(c_n, l_n) l_n}{1-\gamma} = 0,$$

$$c_r - \Omega = (c_n - \Omega) \left[ \frac{(1-\alpha)}{\alpha} \frac{\pi_n}{\pi_r} \left( \frac{\phi(1+\tau_x)}{(1-\alpha)A} \right)^{1/\alpha} \frac{\pi_r l_r}{\pi_n l_n} \right]^{1-\gamma},$$

$$\pi_r c_r + \pi_n c_n + G \leq w_n \pi_n l_n \frac{\tau_x + \alpha}{\alpha(1+\tau_x)} + \frac{w_r \pi_r l_r}{1+\tau_x}.$$

First, let us set $w_n$ and $w_r$ according to their definitions (14) and (15), respectively. Now set $Y$, $\lambda$, $\Omega$ and $m$ such that

$$
\begin{aligned}
Y &= \frac{\pi_n w_n l_n}{\alpha}, \\
\lambda &= \frac{-u_l(c_n, l_n) l_n}{u_c(c_n, l_n)(1-\gamma)(w_n l_n)^{1-\gamma}}, \\
m &= \max \left\{ 1 - \left[ \frac{\phi(1+\tau_x)}{A(1-\alpha)} \right]^{1/\alpha} \frac{\pi_r l_r}{\pi_n l_n}, 0 \right\}.
\end{aligned}
$$

and note that

$$Y=\frac{\pi_n w_n l_n}{\alpha}=\frac{\pi_r w_r l_r}{\alpha}\frac{\pi_n w_n l_n}{\pi_r w_r l_r}=\pi_r w_r l_r\left[\frac{1}{(1-\alpha)\left[\frac{\phi(1+\tau_x)}{A(1-\alpha)}\right]^{1/\alpha}\frac{\pi_r l_r}{\pi_n l_n}}\right]=\frac{\pi_r w_r l_r}{(1-\alpha)\,(1-m)}.$$

To show that the conditions for optimality of non-routine households are satisfied, we note that

$$u_c(c_n,l_n)\,(c_n-\Omega)+\frac{u_l(c_n,l_n)l_n}{1-\gamma}=0$$

$$c_n=-\frac{u_l(c_n,l_n)l_n}{u_c(c_n,l_n)\,(1-\gamma)}+\Omega$$

and by definition of $\lambda$ we obtain

$$c_n=\lambda(w_n l_n)^{1-\gamma}+\Omega.$$

Now to show that the conditions for optimality of the routine household are satisfied, we note that the no-discrimination constraint implies that

$$c_r-\Omega=(c_n-\Omega)\frac{\left(\frac{(1-\alpha)}{\pi_r}\left(\frac{\phi(1+\tau_x)}{(1-\alpha)A}\right)^{1/\alpha}\frac{\pi_r l_r}{\pi_n l_n}\right)^{1-\gamma}}{\left(\frac{\alpha}{\pi_n}\right)^{1-\gamma}}=(c_n-\Omega)\frac{\left(\frac{(1-\alpha)(1-m)Y}{\pi_r}\right)^{1-\gamma}}{\left(\frac{\alpha Y}{\pi_n}\right)^{1-\gamma}}$$

$$c_r-\Omega=\lambda(w_n l_n)^{1-\gamma}\frac{\left(\frac{(1-\alpha)(1-m)Y}{\pi_r}\right)^{1-\gamma}}{\left(\frac{\alpha Y}{\pi_n}\right)^{1-\gamma}}=\lambda(w_r l_r)^{1-\gamma}$$

which shows that the budget constraint is satisfied. Furthermore, from the implementability constraint

$$u_c(c_r,l_r)\,(c_r-\Omega)+\frac{u_l(c_r,l_r)l_r}{1-\gamma}=0$$

$$\Leftrightarrow\quad c_r-\Omega=-\frac{u_l(c_r,l_r)l_r}{u_c(c_r,l_r)\,(1-\gamma)}$$

and using what we have found above

$$-\frac{u_l(c_r,l_r)l_r}{u_c(c_r,l_r)\,(1-\gamma)}=\lambda(w_r l_r)^{1-\gamma}.$$

which shows that the budget marginal condition is also satisfied.

The problem of the government is

$$\max \omega_r \pi_r u(c_r, l_r) + \omega_n \pi_n u(c_n, l_n) + v(G),$$

subject to

$$[\lambda_j] \quad u_c(c_j, l_j)\,(c_r - \Omega) + \frac{u_l(c_j, l_j) l_j}{1-\gamma} = 0,$$

$$\left[\frac{\eta}{c_r}\right] \quad c_r - \Omega = (c_n - \Omega)\,\frac{\left(\frac{(1-\alpha)}{\pi_r}\left(\frac{\phi(1+\tau_x)}{(1-\alpha)A}\right)^{1/\alpha}\frac{\pi_r l_r}{\pi_n l_n}\right)^{1-\gamma}}{\left(\frac{\alpha}{\pi_n}\right)^{1-\gamma}},$$

$$[\phi] \quad u\left(c_j, l_j\right) \geq u\left(\Omega, 0\right)$$

$$[\mu] \quad \pi_r c_r + \pi_n c_n + G \leq w_n \pi_n l_n \frac{\tau_x + \alpha}{\alpha(1+\tau_x)} + \phi \pi_r l_r.$$

The first order condition with respect to $\tau_x$ is given by

$$\begin{aligned} 0 &= \frac{\eta}{c_r}\,(c_n - \Omega)\,\frac{\left(\frac{(1-\alpha)}{\pi_r}\left(\frac{\phi(1+\tau_x)}{(1-\alpha)A}\right)^{1/\alpha}\frac{\pi_r l_r}{\pi_n l_n}\right)^{1-\gamma}}{\left(\frac{\alpha}{\pi_n}\right)^{1-\gamma}}\frac{1-\gamma}{\alpha}\frac{1}{1+\tau_x} \\ &\quad +\mu\left[\frac{dw_n}{d\tau_x}\pi_n l_n \frac{\tau_x+\alpha}{\alpha(1+\tau_x)} + w_n \pi_n l_n \frac{1-\alpha}{\alpha(1+\tau_x)^2}\right] \end{aligned}$$

$$\begin{aligned} \eta\frac{1-\gamma}{\alpha}\left(\frac{c_r-\Omega}{c_r}\right) &= \mu\frac{w_n \pi_n l_n\,(1-\alpha)}{(1+\tau_x)}\tau_x \\ \frac{\tau_x}{1+\tau_x} &= \frac{\alpha}{1-\alpha}\frac{\eta\,(1-\gamma)}{\mu w_n \pi_n l_n}\left(\frac{c_r-\Omega}{c_r}\right). \end{aligned}$$

## A.6 Simple Taxes with lump-sum transfers - The household's problem with regressivity

In this section of the appendix we discuss the problem of the household when the income tax function is regressive. Under the proposed tax function with a lump-sum transfer, a household which has a wage rate $w$ solves the following problem

$$\max u(c,l) \text{ subject to } c \leq \lambda\,(wl)^{1-\gamma} + \Omega.$$

For simplicity assume that preferences are given by

$$u(c,l) = \log c - \zeta\frac{l^{1+\nu}}{1+\nu},$$

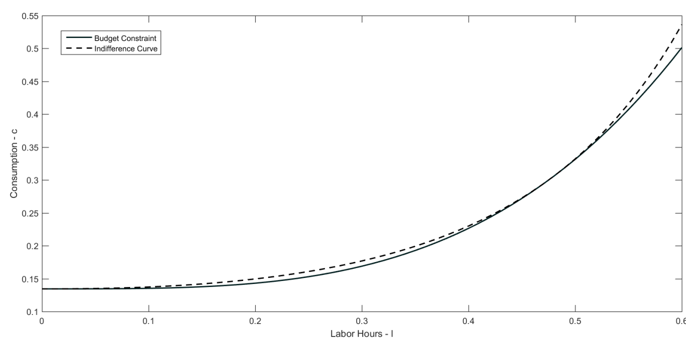
for $\zeta, \nu > 0$. The solution to this problem satisfies the following conditions:

$$c\zeta l^{\nu} = (1-\gamma)\,\lambda w^{1-\gamma} l^{-\gamma}, \tag{67}$$

$$c = \lambda\,(wl)^{1-\gamma} + \Omega. \tag{68}$$

Note that if the tax system is regressive, $\gamma < 0$, and lump-sum transfers are positive, $\Omega > 0$, then as $l \to 0$ both the right- and left-hand sides of (67) converge to zero. As a result, a corner solution may be the optimal choice.

This case actually happens in our solutions to the simple income taxes with lump-sum transfers problem. Indeed, when the routine worker drops out of the labor force, it is optimal to set the lump-sum transfer up to a level in which the non-routine worker is exactly indifferent between the corner solution, with $c = \Omega$ and $l = 0$, and the interior solution, with $c > \Omega$ and $l > 0$. This is easiest seen in the following figure. In this figure we plot both the budget constraint for this case, and the indifference curve for the non-routine worker with the highest associated level of utility.




## A.7 Necessity and sufficiency in the dynamic model

TO BE WRITTEN

## A.8 Proof of lemma 1

First, note that the extensive margin incentive compatibility constraints can be equivalently written as

$$U_t(\theta) \geq U_t(\theta') + \left(\theta' - \theta\right) s_t(\theta') \tag{69}$$

for all $t$ and $\theta, \theta' \in \Theta$.

First, suppose that (69) are satisfied. Then, take $\theta, \theta' \in \Theta_{r,t}$, i.e. such that $s_t(\theta) = s_t(\theta') = 0$. As a result, those conditions imply

$$U_t(\theta) \geq U_t(\theta')$$
$$U_t(\theta') \geq U_t(\theta)$$

which is equivalent to $U_t(\theta) = U_t(\theta')$. This must hold for all $\theta, \theta' \in \Theta_{r,t}$. Let's define $U_{r,t} = U_t(\theta)$ for $\theta \in \Theta_r$. Next, let's take $\theta, \theta' \in \Theta_{n,t}$. Then,

$$U_t(\theta) \geq U_t(\theta') + (\theta' - \theta)$$
$$U_t(\theta') + (\theta' - \theta) \geq U_t(\theta)$$

which implies that for all $\theta, \theta' \in \Theta_{n,t}$ we have $U_t(\theta) = U_t(\theta') + (\theta' - \theta)$. Then, for some $\theta' \in \Theta_{n,t}$ we can define $U_{n,t} = U_t(\theta') + \theta'$, and obtain $U_t(\theta) = U_{n,t} - \theta$ for all $\theta \in \Theta_{n,t}$. Finally, define $\theta_t^* \equiv U_{n,t} - U_{r,t}$. This implies that for all $\theta < \theta_t^*$ we have

$$U_{n,t} - \theta > U_{r,t} \tag{70}$$

which implies that $s_t(\theta) = 1$. For all $\theta > \theta^*$ we have

$$U_{n,t} - \theta < U_{r,t} \tag{71}$$

which implies that $s_t(\theta) = 0$.

To show the reverse implication suppose that the conditions in the lemma hold. Then, for all $\theta \in \Theta_{n,t}$ we have

$$U_t(\theta) = U_{n,t} - \theta = U_t(\theta') + \theta' - \theta, \quad \forall \theta' \in \Theta_{n,t}$$
$$U_t(\theta) = U_{n,t} - \theta \geq U_{n,t} - \theta^* = U_{r,t} = U_t(\theta'), \quad \forall \theta' \in \Theta_{r,t}.$$

Instead, if $\theta \in \Theta_{r,t}$, then

$$U_t(\theta) = U_{r,t} = U_t(\theta'), \quad \forall \theta' \in \Theta_{r,t}$$
$$U_t(\theta) = U_{r,t} = U_{n,t} - \theta^* \geq U_{n,t} - \theta = U_t(\theta') + \theta' - \theta, \quad \forall \theta' \in \Theta_{n,t}.$$

As a result, the allocation is extensive margin incentive compatible, i.e. it satisfies (69).

## A.9 Proof of lemma 2

With the simplification provided by lemma 1, we can write the planning problem as follows:

$$\max \sum_{t=0}^{\infty} \beta^t \left[ U_{n,t} \int_{-\infty}^{\theta_t^*} \omega_t(\theta)\lambda(\theta)d\theta + U_{r,t} \int_{\theta_t^*}^{\infty} \omega_t(\theta)\lambda(\theta)d\theta - \int_{-\infty}^{\theta_t^*} \omega_t(\theta)\lambda(\theta)\theta d\theta \right]$$

$$\begin{aligned}
[\eta^i_{r,t}(\theta)] \quad & U_{n,t} \geq U_{r,t} + \psi\left(l_t(\theta)\right) - \psi\left(\frac{F_{r,t}}{F_{n,t}} l_t(\theta)\right), \quad \forall t \geq 1, \theta > \theta_t^* \\
[\eta^i_{n,t}(\theta)] \quad & U_{r,t} \geq U_{n,t} + \psi\left(l_t(\theta)\right) - \psi\left(\frac{F_{n,t}}{F_{r,t}} l_t(\theta)\right), \quad \forall t \geq 1, \theta < \theta_t^* \\
[\xi_t] \quad & \theta_t^* = U_{n,t} - U_{r,t}, \quad \forall t \geq 1 \\
[\eta^e_{n,t}(\theta)] \quad & U_{n,t} = u(c_t^y(\theta)) - \psi(l_t(\theta)) + v(G_t) + \beta[u(c^o_{t+1}(\theta)) + v(G_{t+1})], \quad \forall t \geq 1, \theta < \theta_t^* \\
[\eta^e_{r,t}(\theta)] \quad & U_{r,t} = u(c_t^y(\theta)) - \psi(l_t(\theta)) + v(G_t) + \beta[u(c^o_{t+1}(\theta)) + v(G_{t+1})], \quad \forall t \geq 1, \theta > \theta_t^* \\
[\mu_t] \quad & \sum_{a=y,o} \int_\theta c_t^a(\theta)\lambda(\theta)d\theta + G_t + \phi_t\left[X_{t+1} - (1-\delta)X_t\right] \leq F\left(X_t, N_{r,t}, N_{n,t}\right), \quad \forall t \geq 1
\end{aligned}$$

where in parenthesis we have written the Lagrange multipliers for each constraint. We will prove this result by means of a contradiction. For a given candidate solution, and for each $t \geq 1$, let us define $j_t^+ \equiv \arg\max_{j\in\{r,n\}} F_{j,t}$ and $j_t^- \equiv \arg\min_{j\in\{r,n\}} F_{j,t}$.

Take first the occupation $j_t^+$ and suppose there exists $\theta, \theta' \in \Theta_{j_t^+,t}$ such that $l_t(\theta) \neq l_t(\theta')$. Then, define $l_{j_t^+} \equiv \inf_{\theta\in\Theta_{j_t^+,t}} l_t(\theta)$. Note that, because $\psi'' > 0$, then

$$\psi(l_{j_t^+}) - \psi\left(\frac{F_{j_t^+,t}}{F_{j_t^-,t}} l_{j_t^+}\right) > \psi(l_t(\theta)) - \psi\left(\frac{F_{j_t^+,t}}{F_{j_t^-,t}} l_t(\theta)\right).$$

This in turn implies that $\eta^i_{j_t^+,t}(\theta) = 0$ for all $\theta$ such that $l_t(\theta) > l_{j_t^+,t}$. Suppose first that $l_{j_t^+,t} < l_t(\theta)$ for all $\theta \in \Theta_{j_t^+,t}$. As a result, the first order conditions with respect to the allocations are

$$\begin{aligned}
& u_c\left(c_t^y(\theta)\right)\eta^e_{j_t^+,t}(\theta) = \mu_t\lambda(\theta) \\
& \beta u_c\left(c^o_{t+1}(\theta)\right)\eta^e_{j_t^+,t}(\theta) = \mu_{t+1}\lambda(\theta) \\
& \psi_l(l_t(\theta))\eta^e_{j_t^+,t}(\theta) - \int_{\Theta_{j_t^-,t}} \psi_l\left(\frac{F_{j_t^-,t}}{F_{j_t^+,t}} l_t(\hat\theta)\right) \eta^i_{j_t^-,t}(\hat\theta) l_t(\hat\theta) d\hat\theta \frac{d\frac{F_{j_t^-,t}}{F_{j_t^+,t}}}{dN_{j_t^+,t}} \lambda(\theta) = \mu_t F_{j_t^+,t}\lambda(\theta)
\end{aligned}$$

which implies that we can solve for $c_t^y(\theta), c_{t+1}^o(\theta), l_t(\theta)$ for all $\theta$ with the following conditions

$$\frac{u_c\left(c_t^y(\theta)\right)}{\beta u_c\left(c_{t+1}^o(\theta)\right)} = \frac{\mu_t}{\mu_{t+1}}$$

$$\frac{\psi_l(l_t(\theta))}{u_c(c_t^y(\theta))} = \frac{\mu_t F_{j_t^+,t} + \int_{\Theta_{j_t^-,t}} \psi_l\left(\frac{F_{j_t^-,t}}{F_{j_t^+,t}} l_t(\hat{\theta})\right) \eta^i_{j_t^-,t}(\hat{\theta}) l_t(\hat{\theta}) d\hat{\theta} \frac{d\frac{F_{j_t^-,t}}{F_{j_t^+,t}}}{dN_{j_t^+,t}}}{\mu_t}$$

$$U_{j_t^+,t} = u(c_t^y(\theta)) - \psi(l_t(\theta)) + v(G_t) + \beta[u(c_{t+1}^o(\theta)) + v(G_{t+1})]$$

In turn, these three conditions imply that the allocations $c_t^y(\theta), c_{t+1}^o(\theta)$, and $l_t(\theta)$ must be constant across $\theta$. This is a contradiction of our original assumption.

If there exists $\theta \in \Theta_{j_t^+,t}$ such that $l_t(\theta) = l_{j_t^+,t}$, then we must modify those consitions. Define $\Theta^1_{j_t^+,t} = \{\theta \in \Theta_{j_t^+,t} : l_t(\theta) = l_{j_t^+,t}\}$ and $\Theta^0_{j_t^+,t} = \Theta_{j_t^+,t} - \Theta^1_{j_t^+,t}$. It is still true that $\eta^i_{j_t^+,t}(\theta) = 0$ for all $\theta \in \Theta^0_{j_t^+,t}$. The first order conditions now become:

$$u_c\left(c_t^y(\theta)\right)\eta^e_{j_t^+,t}(\theta) = \mu_t \lambda(\theta), \quad \forall \theta \in \Theta_{j_t^+,t}$$

$$\beta u_c\left(c_{t+1}^o(\theta)\right)\eta^e_{j_t^+,t}(\theta) = \mu_{t+1} \lambda(\theta), \quad \forall \theta \in \Theta_{j_t^+,t}$$

$$\psi_l(l_t(\theta))\eta^e_{j_t^+,t}(\theta) = \left[\mu_t F_{j_t^+,t} + \Gamma_{j_t^+,t}\right]\lambda(\theta), \quad \forall \theta \in \Theta^0_{j_t^+,t}$$

$$\psi_l(l_t(\theta))\left[\eta^e_{j_t^+,t}(\theta) + \eta^i_{j_t^+,t}(\theta)\left(1 - \frac{\psi_l\left(\frac{F_{j_t^+,t}}{F_{j_t^-,t}} l_t(\theta)\right)\frac{F_{j_t^+,t}}{F_{j_t^-,t}}}{\psi_l\left(l_t(\theta)\right)}\right)\right] = \left[\mu_t F_{j_t^+,t} + \Gamma_{j_t^+,t}\right]\lambda(\theta), \quad \forall \theta \in \Theta^1_{j_t^+,t}$$

where $\Gamma_{j_t^+,t} \equiv -\int_{\theta \in \Theta^1_{j_t^+,t}} \psi_l\left(\frac{F_{j_t^+,t}}{F_{j_t^-,t}} l_t(\hat{\theta})\right)\eta^i_{j_t^+,t}(\hat{\theta}) l_t(\hat{\theta}) d\hat{\theta} \frac{d\frac{F_{j_t^+,t}}{F_{j_t^-,t}}}{dN_{j_t^+,t}} - \int_{\Theta_{j_t^-,t}} \psi_l\left(\frac{F_{j_t^-,t}}{F_{j_t^+,t}} l_t(\hat{\theta})\right)\eta^i_{j_t^-,t}(\hat{\theta}) l_t(\hat{\theta}) d\hat{\theta} \frac{d\frac{F_{j_t^-,t}}{F_{j_t^+,t}}}{dN_{j_t^+,t}}$.
As a result, we obtain the following conditions

$$\frac{u_c\left(c_t^y(\theta)\right)}{\beta u_c\left(c_{t+1}^o(\theta)\right)} = \frac{\mu_t}{\mu_{t+1}}, \quad \forall \theta \in \Theta_{j_t^+,t}$$

$$\frac{\psi_l(l_t(\theta))}{u_c(c_t^y(\theta))} = \frac{\mu_t F_{j_t^+,t} + \Gamma_{j_t^+,t}}{\mu_t}, \quad \forall \theta \in \Theta^0_{j_t^+,t}$$

$$\frac{\psi_l(l_t(\theta))}{u_c(c_t^y(\theta))}\frac{\eta^e_{j_t^+,t}(\theta) + \eta^i_{j_t^+,t}(\theta)\left(1 - \frac{\psi_l\left(\frac{F_{j_t^+,t}}{F_{j_t^-,t}} l_t(\theta)\right)\frac{F_{j_t^+,t}}{F_{j_t^-,t}}}{\psi_l(l_t(\theta))}\right)}{\eta^e_{j_t^+,t}(\theta)} = \frac{\mu_t F_{j_t^+,t} + \Gamma_{j_t^+,t}}{\mu_t}, \quad \forall \theta \in \Theta^1_{j_t^+,t}$$

$$U_{j_t^+,t} = u(c_t^y(\theta)) - \psi(l_t(\theta)) + v(G_t) + \beta[u(c_{t+1}^o(\theta)) + v(G_{t+1})], \quad \forall \theta \in \Theta_{j_t^+,t}$$

If $\eta^i_{j_t^+,t}(\theta) = 0$, again these conditions imply that the allocations must be the same. Otherwise, note that because $F_{j_t^+,t}/F_{j_t^-,t} > 1$ and $\eta^e_{j_t^+,t}(\theta), \eta^i_{j_t^+,t}(\theta) > 0$ then the these marginal conditions imply that

$$\frac{\psi_l(l_t(\theta))}{u_c(c_t^y(\theta))} < \frac{\psi_l(l_t(\theta'))}{u_c(c_t^y(\theta'))} \tag{72}$$

for $\theta \in \Theta^0_{j_t^+,t}$ and $\theta' \in \Theta^1_{j_t^+,t}$. However, the assumption that $l_t(\theta) > l_t(\theta')$ combined with the fact that

$$\frac{u_c\left(c_t^y(\theta)\right)}{\beta u_c\left(c_{t+1}^o(\theta)\right)} = \frac{u_c\left(c_t^y(\theta')\right)}{\beta u_c\left(c_{t+1}^o(\theta\prime)\right)}$$

and

$$u(c_t^y(\theta)) - \psi(l_t(\theta)) + \beta u(c_{t+1}^o(\theta)) = u(c_t^y(\theta')) - \psi(l_t(\theta')) + \beta u(c_{t+1}^o(\theta'))$$

imply that $c_t^y(\theta) > c_t^y(\theta')$. As a result, by convexity of $\psi$ and concavity of $u$ it must be that

$$\frac{\psi_l(l_t(\theta))}{u_c(c_t^y(\theta))} > \frac{\psi_l(l_t(\theta'))}{u_c(c_t^y(\theta'))} \tag{73}$$

which is a contradiction of (72).

The proof for $\Theta_{j_t^-,t}$ is the analogous, by taking $l_{j_t^-,t} \equiv \sup_{\theta\in\Theta_{j_t^-,t}} l_t(\theta)$.

## A.10 Proof of proposition 2

Let us define $\beta^t\mu_t$ the multiplier for period $t$ resource constraint, and $\beta^t\eta_{j,t}$ the multiplier for the incentive constraint of workers of skill $j = r, n$. The first order conditions with respect to consumption are

$$\begin{aligned}
[c^y_{j,t}] \quad & u_c\left(c^y_{j,t}\right)\left[\tilde{\Lambda}_{j,t}\omega_{j,t} + \eta_{j,t} - \eta_{-j,t}\right] = \mu_t\Lambda_{j,t} \\
[c^o_{j,t}] \quad & u_c\left(c^o_{j,t}\right)\left[\tilde{\Lambda}_{j,t-1}\omega_{j,t-1} + \eta_{j,t-1} - \eta_{-j,t}\right] = \mu_t\Lambda_{j,t-1}.
\end{aligned}$$

Taking the ratio of these conditions implies that

$$\frac{u_c\left(c^y_{j,t}\right)}{\beta u_c\left(c^o_{j,t+1}\right)} = \frac{\mu_t}{\beta\mu_{t+1}}$$

for all $t \geq 1$ and $j = r, n$.

If the incentive constraint of routine workers does not bind, then $\eta_{r,t} = 0$. In this case, the first order condition with respect to $X_{t+1}$ is

$$\phi_t \mu_t = \beta \mu_{t+1} \left[ F_{x,t+1} + \phi_{t+1}(1-\delta) \right] + \beta \eta_{n,t+1} \psi_l \left( \frac{F_{r,t+1}}{F_{n,t+1}} l_{r,t+1} \right) \frac{F_{r,t+1}}{F_{n,t+1}} \frac{l_{r,t+1}}{X_{t+1}} \frac{d \log \left( \frac{F_{r,t+1}}{F_{n,t+1}} \right)}{d \log(X_{t+1})}$$

$$\frac{\mu_t}{\beta \mu_{t+1}} = \frac{F_{x,t+1} + \phi_{t+1}(1-\delta)}{\phi_t} + \frac{\eta_{n,t+1}}{\phi_t \mu_{t+1} X_{t+1}} \psi_l \left( \frac{F_{r,t+1}}{F_{n,t+1}} l_{r,t+1} \right) \frac{F_{r,t+1}}{F_{n,t+1}} l_{r,t+1} \mathcal{E}_{t+1}$$

Combining this condition with the previous result yields

$$\frac{u_c \left( c^y_{j,t} \right)}{\beta u_c \left( c^o_{j,t+1} \right)} = \frac{F_{x,t+1} + \phi_{t+1}(1-\delta)}{\phi_t} + \frac{\eta_{n,t+1}}{\phi_t \mu_{t+1} X_{t+1}} \psi_l \left( \frac{F_{r,t+1}}{F_{n,t+1}} l_{r,t+1} \right) \frac{F_{r,t+1}}{F_{n,t+1}} l_{r,t+1} \mathcal{E}_{t+1}$$

and since $\eta_{n,t+1}, \mu_{t+1} > 0$ and $\mathcal{E}_{t+1} < 0$ then as long as $l_{r,t+1} > 0$:

$$\frac{u_c \left( c^y_{j,t} \right)}{\beta u_c \left( c^o_{j,t+1} \right)} < \frac{F_{x,t+1} + \phi_{t+1}(1-\delta)}{\phi_t}.$$

## A.11 Normalizing dynamic model

It is useful to normalize the variables to remove the asymptotic growth trend. Let us define $\overline{c}^y_{i,t} \equiv c^y_{i,t} / \exp\left\{ \frac{1-\alpha}{\alpha} g_\phi (t-1) \right\}$, $\overline{c}^o_{i,t} \equiv c^o_{i,t} / \exp\left\{ \frac{1-\alpha}{\alpha} g_\phi (t-1) \right\}$, $\overline{G}_t \equiv G_t / \exp\left\{ \frac{1-\alpha}{\alpha} g_\phi (t-1) \right\}$, and $\overline{X}_t \equiv \phi_t X_t / \exp\left\{ \frac{1-\alpha}{\alpha} g_\phi (t-2) \right\}$.

As a result, in the optimum we need only find the allocations for agents that acquire routine skills and for the agents that acquire non-routine skills. We can find the optimum by maximizing welfare

$$\omega_0 U_0 + \sum_{t=1}^{\infty} \beta^t \left\{ \sum_{j=n,r} \tilde{\Lambda}_{j,t}(\theta^*_t) \overline{U}_{j,t} - \int_{-\infty}^{\theta^*_t} \theta \lambda(\theta) \omega(\theta) d\theta \right\} \tag{74}$$

where $\overline{U}_{j,t} \equiv u(\overline{c}^y_{j,t}) - \psi(l_{j,t}) + v(\overline{G}_t) + \beta(u(\overline{c}^o_{j,t+1})) + v(\overline{G}_{t+1})$, $\tilde{\Lambda}_{n,t}(\theta^*_t) = \int_{-\infty}^{\theta^*_t} \lambda(\theta)\omega(\theta) d\theta$, and $\tilde{\Lambda}_{r,t}(\theta^*_t) = \int_{\theta^*_t}^{\infty} \lambda(\theta)\omega(\theta) d\theta$, subject to the constraint on $\theta^*_t$

$$\theta^*_t = \overline{U}_{n,t} - \overline{U}_{r,t}, \tag{75}$$

two intensive margin incentive compatibility constraints

$$u(\overline{c}^y_{n,t}) + \beta u(\overline{c}^y_{n,t+1}) - \psi(l_{n,t}) \geq u(\overline{c}^y_{r,t}) + \beta u(\overline{c}^y_{r,t+1}) - \psi\left( \frac{F_{r,t}}{F_{n,t}} l_{r,t} \right), \tag{76}$$

$$u(\overline{c}^y_{r,t}) + \beta u(\overline{c}^y_{r,t+1}) - \psi(l_{r,t}) \geq u(\overline{c}^y_{n,t}) + \beta u(\overline{c}^y_{n,t+1}) - \psi\left( \frac{F_{n,t}}{F_{r,t}} l_{n,t} \right) \tag{77}$$

and the resource constraint

$$\Lambda(\theta_t^*)\overline{c}_{n,t}^y + [1-\Lambda(\theta_t^*)]\overline{c}_{r,t}^y + \Lambda(\theta_{t-1}^*)\overline{c}_{n,t}^o + [1-\Lambda(\theta_{t-1}^*)]\overline{c}_{r,t}^o + \overline{G}_t$$
$$+ \overline{X}_{t+1} - \left(1-\overline{\delta}\right)\overline{X}_t = \overline{A}\left[\overline{X}_t^{\frac{\varepsilon-1}{\varepsilon}} + \left(\overline{\phi}_t\left(1-\Lambda(\theta_t^*)\right)l_{r,t}\right)^{\frac{\varepsilon-1}{\varepsilon}}\right]^{\frac{\varepsilon}{\varepsilon-1}(1-\alpha)} \left(\Lambda(\theta_t^*)l_{n,t}\right)^{\alpha}, \quad (78)$$

for $t \geq 2$, and the resource constraint at $t = 1$ which is similar but there is only one $c_1^o$. Here the modified parameters are given by $\overline{\delta} \equiv 1-(1-\delta)e^{-g_\phi/\alpha}$, $\overline{A} \equiv A\left(\exp\{-(1-\alpha)/\alpha g_\phi\}/\phi^{1-\alpha}\right)$, and $\overline{\phi}_t = \phi\exp\{-g_\phi(t-1)-(1-\alpha)/\alpha g_\phi(t-2)\}$.

## A.12 Proof of lemma 3

In what follows, we derive the steady state assuming that the allocations converge, i.e. assume that the allocations $\{\overline{c}_{n,t}^y, \overline{c}_{n,t}^o, l_{n,t}, \overline{c}_{r,t}^y, \overline{c}_{r,t}^o, l_{r,t}, \theta_t^*, \overline{X}_{t+1}, \overline{G}_t\} \to \{\overline{c}_n^y, \overline{c}_n^o, l_n, \overline{c}_r^y, \overline{c}_r^o, l_r, \theta^*, \overline{X}, \overline{G}\}$ are at a steady state.. We show first that if allocations converge then the multipliers on the constraints must also converge. In the long-run steady state growth path it must be that $\overline{\phi}_t \to 0$ and $F_{r,t}/F_{n,t} \to 0$.

**Lemma 5.** *Suppose that the allocations converge to a steady state growth path with interior automation, then $\eta_{r,t} \to 0$.*

First note that since $\overline{\phi}_t \to 0$ then the optimal labor supply by agents with routine skills is $l_{r,t} = 0$. This implies that the utility of a worker with routine skills converges to

$$U_{r,t} \to u(\overline{c}_r^y) + \beta u(\overline{c}_r^o) + (1+\beta)v(\overline{G}) \equiv U_r$$

while the utility from pretending to be a non-routine worker converges to $-\infty$ since it must be that $l_n > 0$:

$$\lim_{t\to\infty} u(\overline{c}_{n,t}^y) + \beta u(\overline{c}_{n,t+1}^o) + v(\overline{G}_t) + \beta v(\overline{G}_{t+1}) - \psi\left(\frac{F_{n,t}}{F_{r,t}}l_{r,t}\right) = -\infty.$$

The first order conditions with respect to consumption when young are given by:

$$u_c(\overline{c}_n^y)[\Lambda_n\omega_n + \eta_t^* + \eta_{n,t}] = \mu_t\Lambda_n$$
$$u_c(\overline{c}_r^y)[\Lambda_r\omega_r - \eta_t^* - \eta_{n,t}] = \mu_t\Lambda_r$$

These imply that

$$\frac{\Lambda_n\mu_t}{u_c(\overline{c}_n^y)} + \frac{\Lambda_r\mu_t}{u_c(\overline{c}_r^y)} = \Lambda_n\omega_n + \Lambda_r\omega_r$$

which requires that $\mu_t = \mu$ is constant over time. As a result, from the consumption first order conditions we also obtain $\eta_{n,t} \to \eta_n$.

The first order conditions with respect to consumption when old are given by:

$$u_c(\overline{c}_n^o)[\Lambda_n\omega_n + \eta_t^* + \eta_n] = \mu\Lambda_n$$
$$u_c(\overline{c}_r^o)[\Lambda_r\omega_r - \eta_t^* - \eta_n] = \mu\Lambda_r$$

This implies that, in the steady state growth path $\overline{c}_n^y = \overline{c}_n^o$ and $\overline{c}_r^y = \overline{c}_r^o$, and

$$u_c(\overline{c}_n^y)\frac{\Lambda_n\omega_n + \eta_n}{\Lambda_n} = u_c(\overline{c}_r^y)\frac{\Lambda_r\omega_r - \eta_n}{\Lambda_r}.$$

The first order conditions with respect to labor supply

$$l_{r,t} = 0$$
$$\psi_l(l_n)\left[\Lambda_j\omega_j + \eta_n\right] = \mu\alpha\overline{A}\ \overline{X}^{1-\alpha}\left(\Lambda_j l_n\right)^{\alpha-1}\Lambda_j$$

The optimality condition with respect to robots is simply given by the modified golden rule

$$\beta[(1-\alpha)\overline{A}\ \overline{X}^{-\alpha}\left(\Lambda_j l_n\right)^{\alpha} + 1 - \overline{\delta}] = 1 \Leftrightarrow \overline{X} = \left[\frac{(1-\alpha)\overline{A}}{\beta^{-1} - (1-\overline{\delta})}\right]^{\frac{1}{\alpha}}\Lambda_j l_n.$$

Note that combining the first order condition with respect to labor $l_n$ and with respect to consumption $c_n^y$ we obtain

$$\frac{\psi_l(l_n)}{u_c(\overline{c}_n^y)} = \alpha\overline{A}\ \left(\frac{\overline{X}}{\Lambda_j l_n}\right)^{1-\alpha} \Leftrightarrow c_n^y\psi_l(l_n) = \alpha\overline{A}\left[\frac{(1-\alpha)\overline{A}}{\beta^{-1} - (1-\overline{\delta})}\right]^{\frac{1-\alpha}{\alpha}}.$$

This is also the celebrated zero taxation on top result, which is true in this steady state growth path due to the disappearence of general equilibrium effects when $l_r = 0$.

## A.13 Proof of proposition 3

In the steady state growth path the golden rule holds:

$$\beta[(1-\alpha)\overline{A}\ \overline{X}^{-\alpha}\left(\Lambda_j l_n\right)^{\alpha} + 1 - \overline{\delta}] = 1 \Leftrightarrow \overline{X} = \left[\frac{(1-\alpha)\overline{A}}{\beta^{-1} - (1-\overline{\delta})}\right]^{\frac{1}{\alpha}}\Lambda_j l_n.$$

This implies that the distortions on capital should converge to zero if the allocations converge to the interior steady state.

# B Figures

Figure 1: Status-Quo Equilibrium

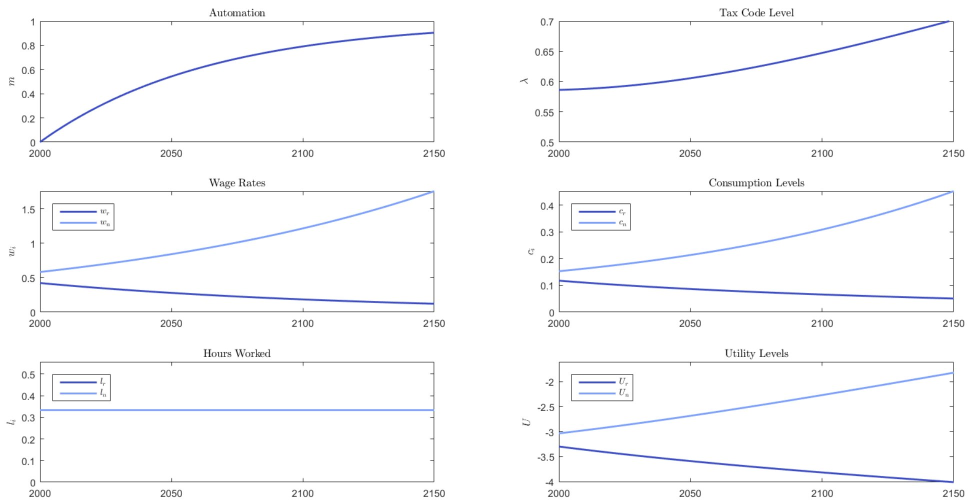



Figure 2: First Best

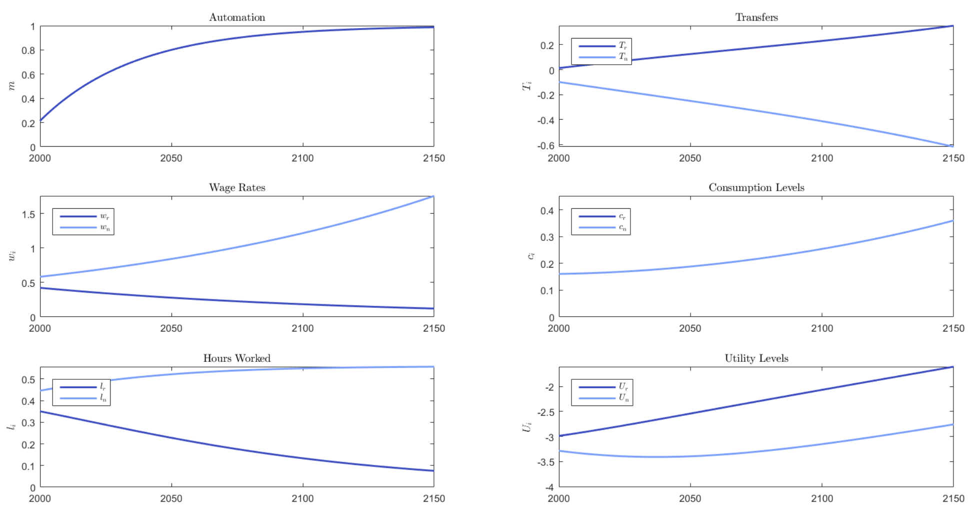


Figure 3: Mirrleesian Optimal Taxation

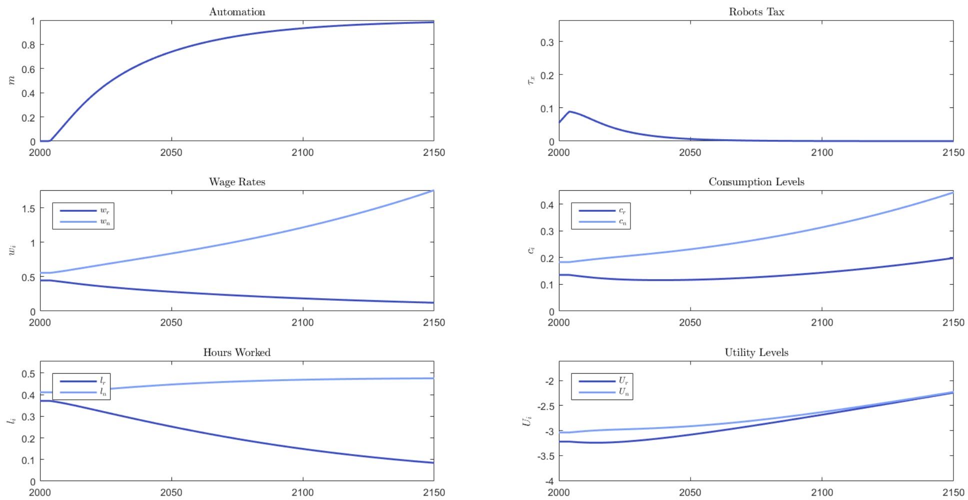



Figure 4: Simple Taxes - Panel A

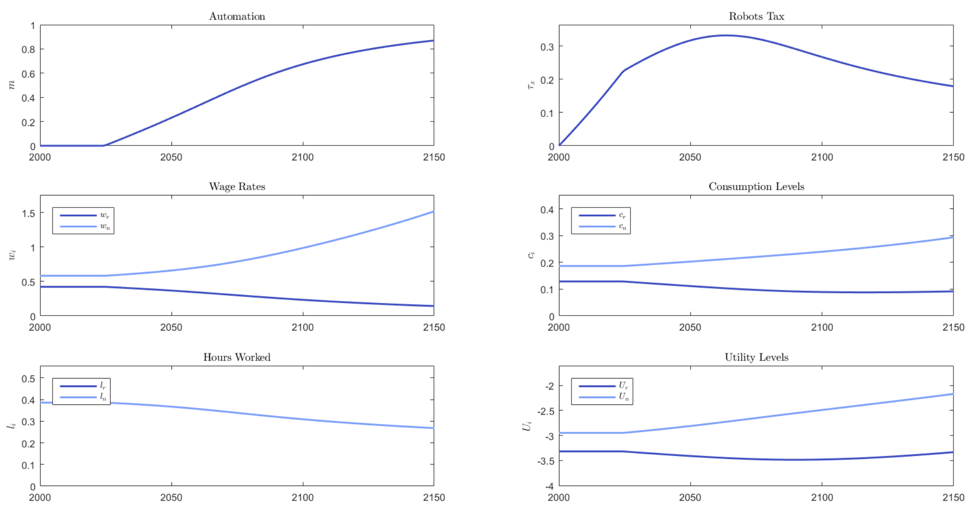


Figure 4: Simple Taxes - Panel B

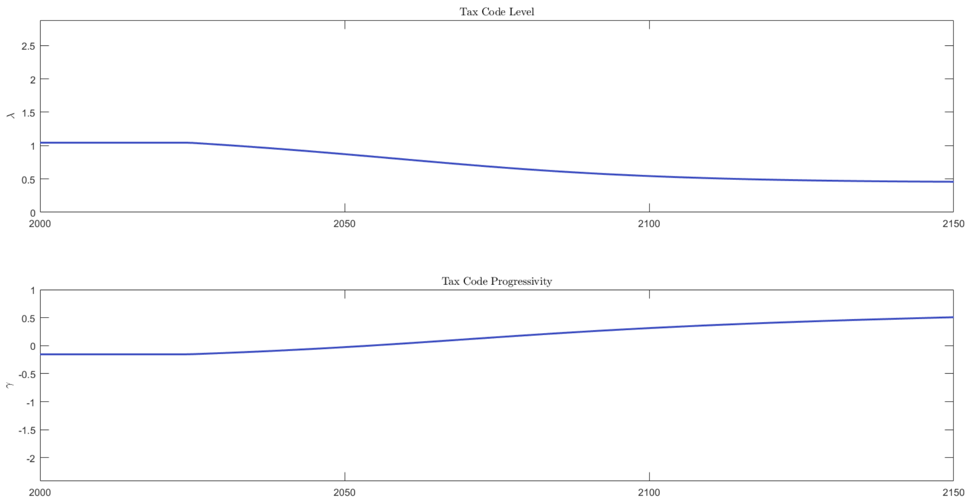



Figure 5: Simple Taxes & Lump Sum Rebate - Panel A

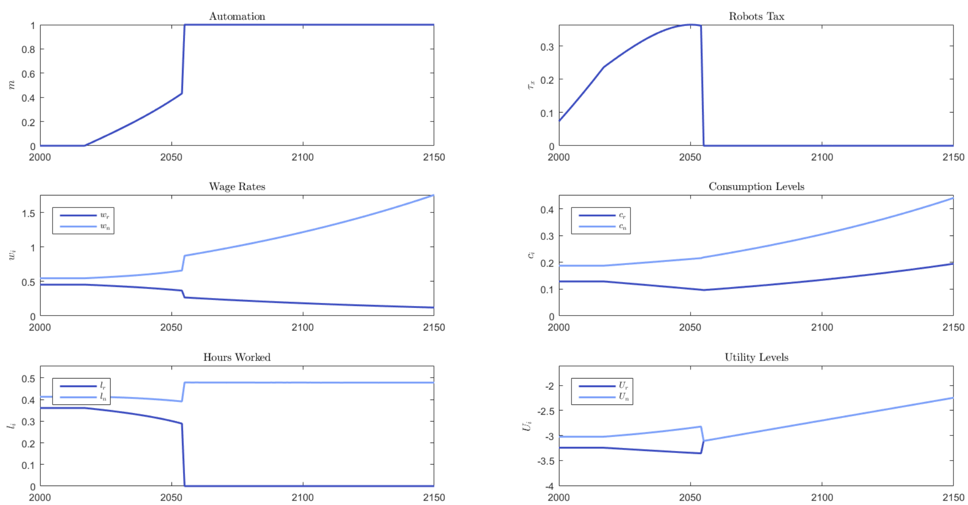


Figure 5: Simple Taxes & Lump Sum Rebate - Panel B

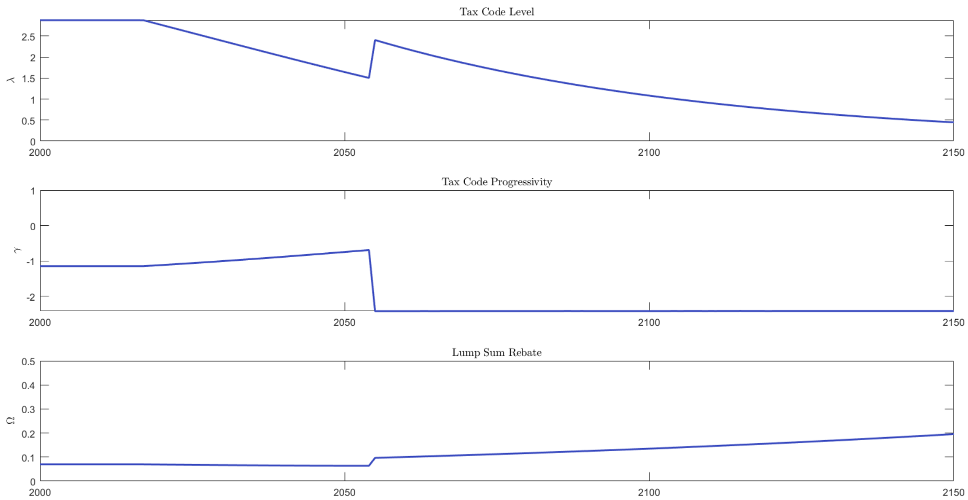



Figure 6: Welfare

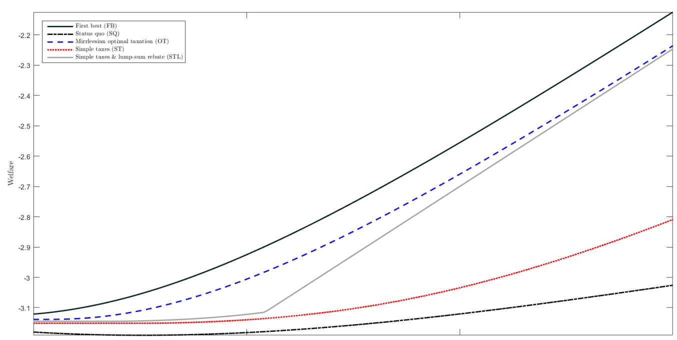


Figure 7: Consumption Equivalent

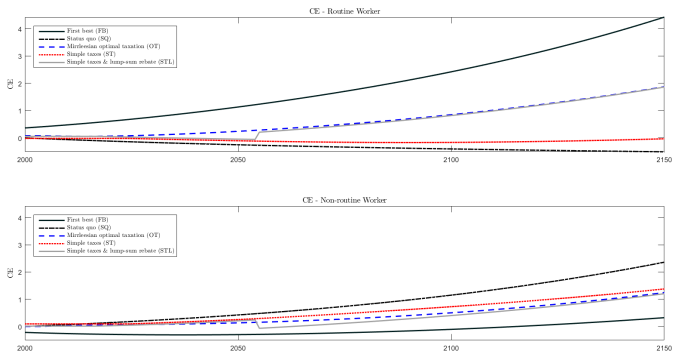



Figure 8: First Best with Occupational Choice

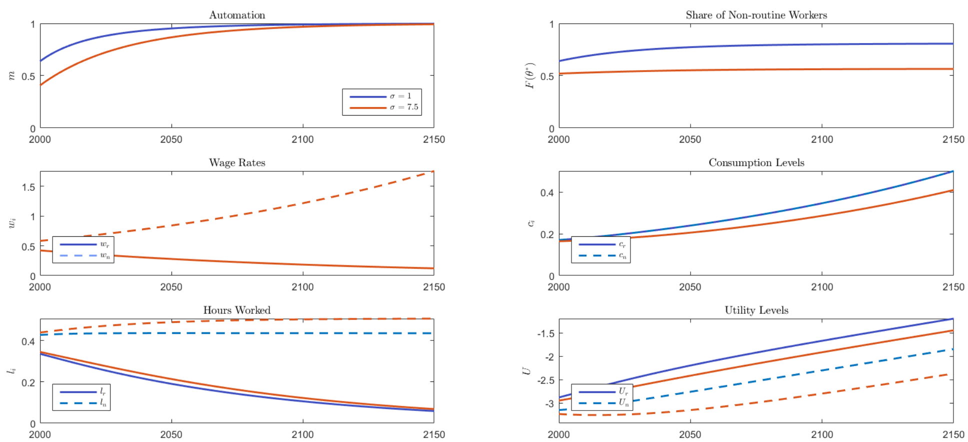


Figure 9: Mirrlees Optimal Taxation with Occupational Choice (Panel A)

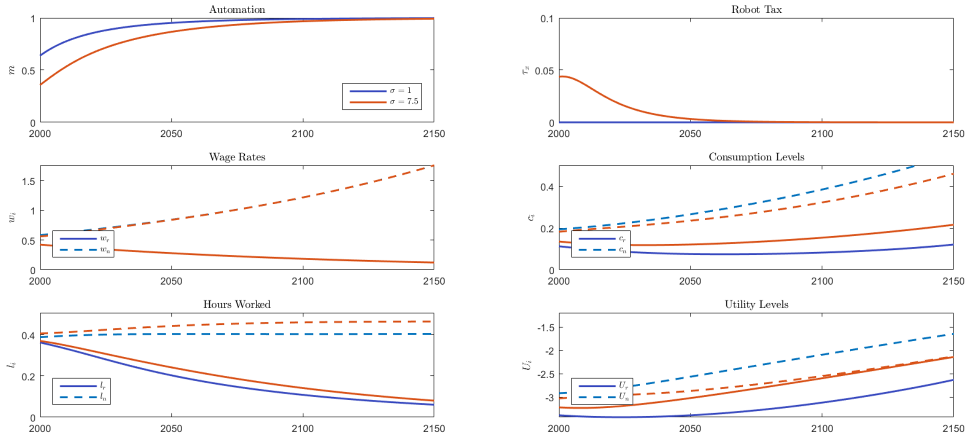



Figure 9: Mirrlees Second Best with Occupational Choice (Panel B)

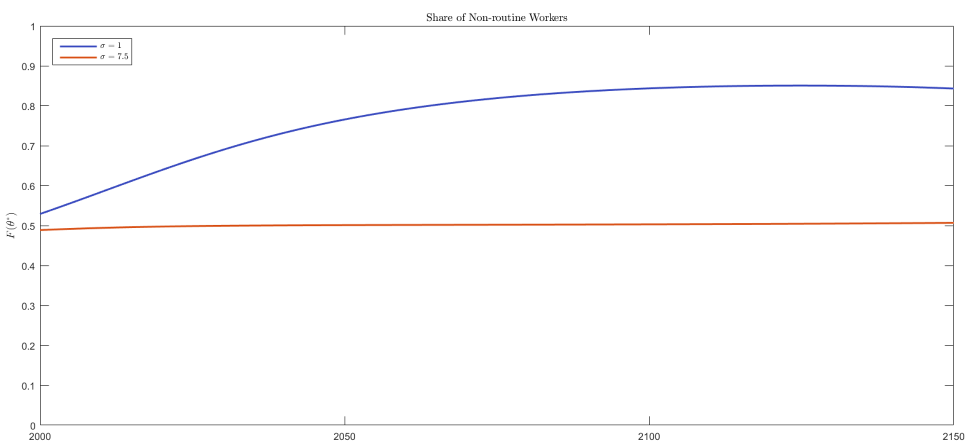


Figure 10: Mirrleesian Optimal Taxation - Dynamic Model