

Robots, Trade, and Luddism: A Sufficient Statistic Approach to Optimal Technology Regulation*

Arnaud Costinot
MIT

Iván Werning
MIT

This Version: April 2022

Abstract

Technological change, from the advent of robots to expanded trade opportunities, creates winners and losers. How should government policy respond? We provide a general theory of optimal technology regulation in a second-best world, with rich heterogeneity across households, linear taxes on the subset of firms affected by technological change, and a nonlinear tax on labor income. Our first set of results consists of optimal tax formulas, with minimal structural assumptions, involving sufficient statistics that can be implemented using evidence on the distributional impact of new technologies, such as robots and trade. Our final results are comparative static exercises illustrating, among other things, that while distributional concerns create a rationale for non-zero taxes on robots and trade, the magnitude of these taxes may decrease as the process of automation and globalization deepens and inequality increases.

*We thank Pol Antràs, Pablo Fajgelbaum, Matt Grant, Marc Melitz, Chris Palmer, Pascual Restrepo, Andrés Rodríguez-Clare, Florian Scheuer, Uwe Thueemmel and seminar and conference participants for helpful comments. Masao Fukui, Tishara Garg, Mary Gong, Danny O'Connor, John Sturm, and Martina Uccioli provided excellent research assistance. Finally, we thank HAL for its unrelenting discouragement.

1 Introduction

Robots and artificial intelligence technologies are on the rise. So are imports from China and other developing countries. These changes create opportunities for some workers, destroy opportunities for others, and generate significant distributional consequences, as documented in the recent empirical work of [Autor, Dorn and Hanson \(2013\)](#) and [Acemoglu and Restrepo \(2017b\)](#) for the United States.

Should any policy response be in place? And if so, how should we manage new technologies? Should we become more Luddite as machines become more efficient or more protectionist as trade opportunities expand? The goal of this paper is to provide a general second-best framework to help address these and other related questions.

Answers to these questions necessarily depend on the range of available policy instruments. At one extreme, if lump-sum transfers are available, as in the Second Welfare Theorem, or if linear taxes are available on all goods and factors, as in [Diamond and Mirrlees \(1971a,b\)](#) and [Dixit and Norman \(1980\)](#), then redistribution can be done without distorting production. In such cases, production efficiency implies the optimality of zero taxes on robots and free trade. At another extreme, in the absence of any policy instrument, whenever technological progress creates at least one loser, a welfare criterion must be consulted and the status quo may be preferred.

Here, we focus on a more realistic situation where tax instruments are available, but more limited than those ensuring production efficiency. We restrict the set of taxes that can be imposed on households' labor supply to be a function of their income, but not on their labor type. This creates a canonical trade-off between redistribution and efficiency. We consider two sets of technologies, which we refer to as old and new. For instance, firms using the new technology may be producers of robots or traders that export some goods in exchange for others. Our main focus is on the optimal regulation of the new technology captured by ad-valorem taxes on firms using that technology. Key to our analysis is the idea that, in addition to redistribution using nonlinear income-taxation, there may be *predistribution* using taxes on new technology firms, e.g. taxes on robots or trade, in order to affect wages across the income distribution.

Our first set of results characterizes the structure of optimal taxes on new technology firms that best complements optimal income taxation. In a two-type environment, [Naito \(1999\)](#) has proven that governments seeking to redistribute income from high- to low-skill workers may have incentives to depart from production efficiency. Doing so manipulates relative wages, which cannot be taxed directly, and relaxes incentive compatibility constraints. Our general analysis goes beyond this qualitative insight by allowing for rich

heterogeneity across households and deriving optimal tax formulas expressed in terms of sufficient statistics that are, at least in principle, empirically measurable.

Specifically, we provide two novel optimal tax formulas, each based on the general observation that starting at an optimum, any small change in taxes should have zero first-order welfare effects. The first of our formulas focuses on changes in the taxes on new technology firms, while holding fixed the shape of the income tax schedule. The second of our formulas accompanies the former tax changes with a reform of the income tax schedule that holds fixed the distribution of utility across quantiles of the income distribution. Each formula provides different insights and involves its own set of sufficient statistics, but they both give a central role to the impact on the wage distribution. Although the response of wages to robots or trade is of obvious empirical interest for descriptive reasons, our formulas show how it also provides a sufficient statistic for optimal policy design. Given knowledge of this statistic, the specific structure of the economy leading to a change in wages can be left in the background. For example, it is not necessary to take a stand on how robots and workers may be combined to perform different tasks, or on how production processes may get fragmented across countries. While these features may be critical in shaping the impact of new technologies on wages, all that is needed according to our formulas is knowledge of this impact, not how it comes about. This distinguishes our approach from other more structural and quantitative explorations.

We illustrate the usefulness of our approach by exploring the magnitude of optimal taxes on robots and trade. We focus on our second formula, which can be implemented without taking a stand on preferences for inequality, since it does not involve social welfare weights. Using the reduced-form evidence of [Acemoglu and Restrepo \(2017b\)](#) on the impact of robots in the United States, we find efficient taxes on robots ranging from 3.78% to 6.42%. In contrast, the evidence of [Chetverikov, Larsen and Palmer \(2016\)](#) on the impact of Chinese imports on U.S. inequality points towards even smaller efficient tariffs, between 0.02% to 0.12%. While the estimated impact of robots and Chinese imports on wages is of similar magnitude, robots account for a much smaller share of the U.S. economy. According to our formula, this calls for a smaller tax on trade than robots.

While the previous results shed light on optimal taxes on robots and trade based on recent empirical evidence, a distinct question is how future changes in the pace of automation and globalization may affect optimal technology regulation. Our final results are comparative static exercises designed to explore this issue. We do so in the context of a simple economy in which the government puts higher social welfare weight on the poorest households and where cheaper machines, either robots or imported machines from China, increase inequality. Despite these two features, we first show that improvements

in new technologies are associated with lower taxes on firms using those technologies. Thus, as the process of automation and globalization deepens, more inequality may best be met with lower Luddism and less protectionism.

We also show through a numerical example that when the economy is calibrated to match existing reduced-form evidence on the impact of robots and Chinese imports on relative wages, the optimal tax on machines accompanying optimal income taxes remains small for all values of social preferences. When income taxes differ significantly from optimal ones, large taxes or subsidies on machines meant to undo the distortions associated with income taxes are possible, but the welfare gains from such second-best interventions remain minimal.

Related Literature

Our paper makes two distinct contributions to the existing literature. The first one is a general characterization of the structure of optimal taxes in environments where different factors (i.e. labor skill types) are subject to the same income tax. In so doing, we fill a gap between the general analysis of [Diamond and Mirrlees \(1971a,b\)](#) and [Dixit and Norman \(1980\)](#), which assumes that linear taxes on all factors are available, and specific examples, typically with two goods and two labor skill types, in which only income taxation is available, as in the original work of [Naito \(1999\)](#), and subsequent work by [Guesnerie \(1998\)](#), [Spector \(2001\)](#), [Naito \(2006\)](#), [Slavik and Yazici \(2014\)](#), and [Jacobs \(2015\)](#).¹ On the broad spectrum of restrictions on available policy instruments, one can also view our analysis as an intermediate step between the work of [Diamond and Mirrlees \(1971a,b\)](#) and [Dixit and Norman \(1980\)](#) and the trade policy literature where it is common to assume that the only instruments available for redistribution are trade taxes. In fact, our first formula is a strict generalization of the formulas reviewed by [Helpman \(1997\)](#), including [Grossman and Helpman's \(1994\)](#) tariff formula.

Our second contribution is to offer a more specific application of our general formulas to the taxation of robots and trade. In recent work, [Guerreiro, Rebelo and Teles \(2017\)](#) and [Thuemmel \(2018\)](#) have studied models with both heterogeneous workers as well as robots. Assuming factor-specific taxes are unavailable, they find a non-zero tax on robots to be generally optimal, in line with [Naito \(1999\)](#). Although we share the same rationale

¹In the first three papers, like in [Dixit and Norman \(1980\)](#), the new technology is international trade. In another related trade application, [Feenstra and Lewis \(1994\)](#) study an environment where governments cannot subject different worker types to different taxes, but can offer subsidies to workers moving from one industry to another in response to trade. They provide conditions under which such a trade adjustment assistance program are sufficient to guarantee Pareto gains from trade, as in [Dixit and Norman \(1980\)](#).

for finding nonzero taxes on robots, based on [Naito \(1999\)](#), our main goal is not to sign the tax on robots, nor to explore a particular production structure, but instead to offer tax formulas highlighting key sufficient statistics needed to determine the level of taxes, with fewer structural assumptions. In this way, our formulas provide a foundation for empirical work as well as the basis for novel comparative static results.² In another recent contribution, [Hosseini and Shourideh \(2018\)](#) analyze a multi-country Ricardian model of trade with input-output linkages and imperfect mobility of workers across sectors. Although sector-specific taxes on labor are not explicitly allowed, these missing taxes can be perfectly mimicked by the available tax instruments. By implication, their economy provides an alternative implementation but fits [Diamond and Mirrlees \(1971b,a\)](#) and [Dixit and Norman \(1980, 1986\)](#), where households face a complete set of linear taxes, including sector-specific taxes on labor. Production efficiency and free trade then follow, just as they did in [Diamond and Mirrlees \(1971a,b\)](#).³

2 Environment

We consider an economy with an arbitrary number of goods and a continuum of heterogeneous households supplying labor. Households have the same preferences, but differ in their skills. We allow this heterogeneity to be multi-dimensional, unlike the classical one-dimensional Mirrleesian model. For instance, a household may be more productive at some tasks, but less productive at others, as in a Roy model. Households sell their labor in competitive labor markets and pay nonlinear taxes on their earnings to the government. Production is carried out by competitive firms. The government may linearly tax transactions between firms and households as well as the transactions that take place between firms, inducing production inefficiency. This is the focus of our analysis.

²Related theoretical work on the employment and growth effects of automation include [Benzell, Kotlikoff, LaGarda and Sachs \(2015\)](#) and [Acemoglu and Restrepo \(2017a\)](#).

³A separate line of work, e.g. [Itskhoki \(2008\)](#), [Antras, de Gortari and Itskhoki \(2017\)](#) and [Tsyvinski and Werquin \(2018\)](#), studies technological changes such as trade or robots, without considering taxes on these new technologies, but instead focusing on how the income tax schedule may respond to these changes.

2.1 Preferences

All households have identical and weakly separable preferences between goods and labor. The utility of household θ is given by

$$\begin{aligned}U(\theta) &= u(C(\theta), n(\theta)), \\ C(\theta) &= v(c(\theta)),\end{aligned}$$

where $C(\theta)$ is the sub-utility that household θ derives from consuming goods, $n(\theta)$ is her labor supply, $c(\theta) \equiv \{c_i(\theta)\}$ is her vector of good consumption. Throughout our analysis, we assume that the upper-level utility function $u(C, n)$ is quasi-concave and strictly increasing in C and decreasing in n , while the lower-level utility function $v(c)$ is quasi-homothetic, quasi-concave and strictly increasing in c .⁴

2.2 Technology

Households are distinguished by their skill $\theta \in \Theta \subseteq \mathbb{R}^K$ with distribution F . Each skill type θ provides a distinct labor input for use in production. We assume that, for at least one of the elements of $\theta = (\theta_1, \theta_2, \dots, \theta_K)$, higher values are associated with higher productivity (thus, commanding higher wages).

We divide technologies into two types, which we refer to as *old* and *new*, each associated with a distinct production set. In our applications, the old technology is how most production takes place, while the new technology captures either trade with the rest of the world or the production of machines, like robots. The dichotomy between old and new technologies is what allows us to consider the taxation of transactions between firms and the resulting aggregate production inefficiency. Without such taxation we could consolidate technology into a single aggregate production set.

Old Technology. Let $y \equiv \{y_i\}$ denote the vector of total net output by old technology firms and let $n \equiv \{n(\theta)\}$ denote the schedule of their total labor demand. Positive values for y_i represent output, while negative y_i represent inputs. The production set associated with the old technology corresponds to all production plans (y, n) such that

$$G(y, n) \leq 0,$$

⁴Quasi-homotheticity is defined by the requirement that Marshallian demands are linear in total spending on the c goods. It nests the homothetic and quasilinear specifications as special cases. Though this assumption simplifies parts of our analysis, it is not required for our main results, as shown in [Costinot and Werning \(2018\)](#).

where G is a strictly increasing, convex, and homogeneous function of (y, n) . Homogeneity of G implies constant returns to scale.

Except for constant returns to scale, we impose no restriction on the old technology. This allows for arbitrary production networks and global supply chains. For instance, old technology firms may be able to produce a final good by executing a continuum of tasks, with each task chosen to be performed by either workers or robots, as in [Acemoglu and Restrepo \(2017a\)](#), or by domestic or foreign workers, as in [Grossman and Rossi-Hansberg \(2008\)](#). In such environments, the production possibility frontier G can be derived from a subproblem that solves the optimal assignment of workers and other inputs to tasks. The commodity vector y then consists of the final good produced and the intermediate goods demanded, i.e., the robots or foreign labor services supplied by new technology firms, but omits tasks as they become subsumed in the definition of G . Appendix [B.1](#) provides the formal mapping between production functions that explicitly model tasks and our general production possibility frontier G .

New Technology. Let $y^* \equiv \{y_i^*\}$ denote the vector of total net output by new technology firms. The production set associated with the new technology corresponds to all production plans y^* such that

$$G^*(y^*; \phi) \leq 0,$$

where G^* is a strictly increasing, convex, and homogeneous function of y^* and $\phi > 0$ is a productivity parameter.

Unlike the old technology, the new technology does not employ labor directly. This assumption fits well our applications to robots and trade. In the first case, new technology firms may be robot-producers that transform a composite of all other goods in the economy, call it gross output, into robots. This is the standard way to model capital accumulation in a neoclassical growth model.⁵ In the second case, new technology firms may be traders who export and import goods,

$$G^*(y^*; \phi) = \sum_i \bar{p}_i(\phi) y_i^*,$$

where $\bar{p}_i(\phi)$ denotes the world price of good i . A change in ϕ corresponds to a terms-of-trade shock that may be due to a change in transportation costs or productivity in the rest

⁵Due to constant returns to scale, profits are zero for the owners of new technology firms. In our framework, one can capture windfall gains of initial owners by introducing endowments of robots. This leads to issues similar to those found in the literature on capital taxation, with the optimum possibly calling for expropriatory levels of taxation of such initial wealth.

of the world.

Abstracting from labor in the new technology is convenient, as it implies that wages are determined by the old technology. New technology has an effect on wages, through its effect on the structure of production within the old technology, but not directly through employment. For a fixed value of ϕ , this restriction is without loss of generality since the new technology sector can always be defined as the last stage of production where taxation is imposed, as described in Appendix B.2. For comparative static exercises, the omission of labor from G^* implicitly restricts attention to changes in ϕ that are labor-neutral in the sense that they do not induce changes in wages for given prices faced by the old technology firms.

Resource Constraint. For all goods, the demand by households is equal to the supply by old and new technology firms,

$$\int c(\theta)dF(\theta) = y + y^*.$$

2.3 Prices and Taxes

Factors. Let $w \equiv \{w(\theta)\}$ denote the schedule of wages faced by firms. Because of income taxation, a household with ability θ and labor earnings $w(\theta)n(\theta)$ retains

$$w(\theta)n(\theta) - T(w(\theta)n(\theta)),$$

where $T(w(\theta)n(\theta))$ denotes its total tax payment. Crucially, the income tax schedule T is the same for all households. This rules out agent-specific lump-sum transfers, in contrast to the Second Welfare Theorem, as well as factor-specific linear taxes, in contrast to the analysis of [Diamond and Mirrlees \(1971a,b\)](#) and [Dixit and Norman \(1980\)](#).

Goods. Let $p^* \equiv \{p_i^*\}$ denote the vector of good prices faced by new technology firms. Because of ad-valorem taxes $t^* \equiv \{t_i^*\}$, these prices may differ from the vector of good prices $p \equiv \{p_i\}$ faced by old technology firms and households. Non-arbitrage implies

$$p_i = (1 + t_i^*)p_i^*, \text{ for all } i.$$

Production inefficiency arises if $t^* \neq 0$ because of the wedge created between the two technologies. In the robot context the tax in question might be a tax on robots produced by the new technology and employed in the old technology. In a trade context, an import

tariff or an export subsidy on good i corresponds to $t_i^* > 0$, whereas an import subsidy or an export tax corresponds to $t_i^* < 0$. Since demand and supply only depend on relative prices, we can normalize prices and taxes such that $p_1 = p_1^* = 1$ and $t_1^* = 0$. We maintain this normalization throughout.⁶

3 Equilibrium, Social Welfare, and Government Problem

We now define the equilibrium for this economy, introduce our general social welfare criterion, and describe the government problem.

3.1 Equilibrium

An equilibrium consists of an allocation, $c \equiv \{c(\theta)\}$, $n \equiv \{n(\theta)\}$, $C \equiv \{C(\theta)\}$, $y \equiv \{y_i\}$, and $y^* \equiv \{y_i^*\}$, prices and wages, $p \equiv \{p_i\}$, $p^* \equiv \{p_i^*\}$, and $w \equiv \{w(\theta)\}$, as well as an income tax schedule, T , and taxes on new technology firms, $t^* \equiv \{t_i^*\}$, such that: (i) households maximize their utility, (ii) firms maximize profits, (iii) markets clear, (iv) prices satisfy the non-arbitrage condition, and (v) the government's budget is balanced. All these equilibrium conditions are standard. We collect them in Appendix C.

The equilibrium determination of wages is central to our analysis. As shown in Appendix C, profit maximization by old technology firms implies a wage schedule

$$w(p, n; \theta)$$

that depends on prices p and labor n . By affecting the labor demand of old technology firms, changes in p affect wages. Given the limited ability of the government to tax different factors differently, this creates a pecuniary motive for taxing goods. This redistribution is the key mechanism at play in our optimal tax formulas.

3.2 Social Welfare

We consider a general social welfare criterion that depends on the distribution of individual well-beings, not the particular well-being of certain agents. Any consumption and labor supply schedule $(c, n) \equiv \{c(\theta), n(\theta)\}$ is associated with a utility schedule $U \equiv \{U(\theta)\}$. This, in turn, induces a cumulative distribution over utilities, summarized

⁶Section 4.3 discusses the generalization of our results to environments where ad-valorem taxes on old technology firms, $t \equiv \{t_i\}$, are also available. In such case, the government may also create a wedge between the prices $p \equiv \{p_i\}$ faced by old technology firms and those faced by consumers, $q \equiv \{q_i\}$.

by the utility levels $\bar{U} \equiv \{\bar{U}(z)\}$ associated with each quantile $z \in [0, 1]$. The social welfare objective is assumed to be a strictly increasing function of this induced distribution,

$$\mathcal{W}(\bar{U}).$$

When θ is one dimensional and higher- θ households achieve higher utility, as in the standard Mirrleesian setup, this nests the special case of a weighted utilitarian objective. When θ is multidimensional, our assumption about \mathcal{W} only restricts Pareto weights to be the same for all households θ earning the same wage, since they obtain the same utility.

3.3 Government Problem

The government problem is to select a competitive equilibrium with taxes that maximizes social welfare. A compact statement of the government problem is as follows:

$$\max_{(c, n, y, y^*, p, p^*, w, T, t^*, \bar{U}) \in \Omega} \mathcal{W}(\bar{U})$$

subject to

$$G^*(y^*; \phi) = 0,$$

where the feasible set Ω imposes equilibrium conditions (i)-(iv), except for the requirement that new technology firms are on their production possibility frontier, $G^*(y^*; \phi) = 0$, as also described in Appendix C. By Walras' Law, $G^*(y^*; \phi) = 0$ if and only if the government's budget balance condition (v) holds. With this in mind, we sometimes refer to this constraint as the government's budget balance condition.

4 Optimal Technology Regulation

Our first set of results characterizes the structure of taxes on new technology firms. Specifically, we provide optimal tax formulas expressed in terms of sufficient statistics and using minimal structural assumptions.

4.1 Efficiency vs. Redistribution

Our tax formulas are derived by starting from an initial equilibrium with taxes (t^*, T) and engineering marginal changes δt^* in the taxes on the new technology firms and changes in the nonlinear tax schedule δT , such that all the equilibrium conditions are met except,

potentially, $G^*(y^*; \phi) = 0$. These marginal tax changes, in turn, induce general equilibrium marginal adjustments in prices δp , wages δw , quantities δy^* , labor δn , as well as other variables and, ultimately, social welfare.⁷ Our formulas differ in how the nonlinear tax is adjusted.

We start with an intermediate result that encompasses all cases, providing a condition that the marginal tax changes δt^* and δT as well as the marginal adjustments δp , δw , δy^* and δn must satisfy so that welfare is not improved by the variation. For any household at the quantile $z \in [0, 1]$ of both the utility and wage distribution, let $\bar{w}(z)$, $\bar{n}(z)$, $\bar{x}(z)$, and $\bar{c}(z)$ denote the common wage, labor supply, earnings, and consumption vector, respectively, and let $\tau(z) \equiv T'(\bar{x}(z))$ denote the marginal income tax rate. Using the previous notation, our first optimal tax result can be stated as follows.

Lemma 1. *Suppose that taxes (t^*, T) are optimal. Then for any variation $(\delta t^*, \delta T)$,*

$$\begin{aligned} & - \sum_i t_i^* (p_i^* y_i^*) \delta \ln y_i^* - \int \tau(z) \bar{x}(z) \delta \ln \bar{n}(z) dz \\ & = \int [\bar{\lambda}(z) - 1] \bar{x}(z) [(1 - \tau(z)) \delta \ln \bar{w}(z) - \frac{\delta T(z)}{\bar{x}(z)} - \sum_i \frac{p_i \bar{c}_i(z)}{\bar{x}(z)} \delta \ln p_i] dz, \quad (1) \end{aligned}$$

where $\bar{\lambda}(z)$ measures the social marginal benefit of allocating income to households at quantile z , as described in Appendix D.1.

Lemma 1 captures the trade-off between efficiency and redistribution. It states that for a marginal change in taxes not to improve welfare, its marginal costs in terms of efficiency should be equal to its marginal benefit in terms of redistribution. The formal proof is based on a standard variational argument similar to those used, for instance, in Saez (2001) and Tsyvinski and Werquin (2018) to characterize properties of the income tax schedule. The novelty here is to use Lemma 1 to study the structure of optimal commodity taxes, thereby opening up the door for generalizations and empirical applications of Naito's (1999) original insights.

Efficiency considerations are reflected in two fiscal externalities on the left-hand side of equation (1). The first term represents the change in revenues from the linear tax t^* , also equal to the marginal increase in the deadweight burden or “Harberger triangle”; the second term captures the change in revenue from the non-linear income tax schedule. These changes in revenue are not internalized by private agents and thus represent a change in efficiency.

⁷These marginal adjustments are general equilibrium adjustments in the sense that δp , δw , δy^* , δn are total derivatives that take into account both the direct effect of the change in taxes on each of these variables as well as their indirect effect through changes in all other prices and quantities.

Distributional considerations are represented by the right-hand side. It evaluates the change in utility in monetary terms directly perceived by a household, weighted by $\bar{\lambda}(z) - 1$. By an envelope argument, the percentage change in real income for z is given by $(1 - \tau(z)) \delta \ln \bar{w}(z) - \sum_i \frac{p_i \bar{c}_i(z)}{\bar{x}(z)} \delta \ln p_i - \frac{\delta T(z)}{\bar{x}(z)}$. The first two terms, $(1 - \tau(z)) \delta \ln \bar{w}(z) - \frac{\delta T(z)}{\bar{x}(z)}$, capture the percentage change in income, due to both the change in before-tax wages as well as the change in the tax schedule. The final term $-\sum_i \frac{p_i \bar{c}_i(z)}{\bar{x}(z)} \delta \ln p_i$ adjusts this change in income by a household-specific measure of inflation.

It is also important to note that all variables and adjustments in equation (1) are expressed in terms of the quantile z , not the underlying skills θ . This implies that one can collapse heterogeneity and proceed *as if* there were a single dimension of heterogeneity. From a theoretical standpoint, this feature derives both from our assumption that the government's social welfare function only depends on the distribution of utility levels $\{\bar{U}(z)\}$ and uses a mathematical result that equates the change in the wage $\bar{w}(z)$ at a given quantile z with the average change in the wage for households originally at this quantile.⁸

From a practical standpoint, this is critical. It implies that researchers can focus measurement on changes in the distribution of wages $\{\bar{w}(z)\}$, as is often done in practice for positive, descriptive purposes. Individual workers may move across the wage distribution, switching quantiles, as panel data would reveal, but a repeated cross-section of wages is sufficient. In so doing, we provide a normative rationale for quantile wage regressions.

4.2 Optimal Tax Formulas

We now explore two feasible tax variations, each leading to a novel optimal tax formula. Both variations lead to changes in y^* in any desired direction, but differ with respect to the nonlinear labor income tax schedule. In particular, we consider variations with:

- i. no change in (the shape of) the income tax schedule, $\delta T = 0$;
- ii. no change in (the distribution of) utility, $\delta \bar{U} = 0$.

To obtain a formula for the tax on good $i \neq 1$ in each of these cases, we focus on variations in the taxes on new technology firms δt^* such that $\delta y_i^*, \delta y_1^* \neq 0$ and $\delta y_j^* = 0$ for $j \neq i, 1$.

⁸The formal argument, which is related to Reynolds Transport Theorem, can be found in Appendix D.1. This reduction in dimensionality due to the equivalence between these two measures of wage changes has no counterpart in the context of the standard Mirrlees optimal income-tax analysis (e.g. Saez, 2001), where wages are assumed to be fixed.

No change in (the shape of) the income tax schedule. Our first variation is the simplest. It leaves the labor income tax schedule unchanged, up to some uniform lump-sum transfer that maintains the government's budget balance, $G^*(y^*; \phi) = 0$. As demonstrated in Appendix D.2, substituting for the value of this lump-sum transfer in equation (1) leads to the following formula.

Proposition 1 ($\delta T = 0$). *The optimal tax on good $i \neq 1$ satisfies*

$$t_i^* = \int \left[1 - \frac{\bar{\lambda}(z)}{\int \bar{\lambda}(v) dv} \right] \frac{\bar{x}(z)}{p_i^* y_i^*} \left[(1 - \tau(z)) \frac{\delta \ln \bar{w}(z)}{\delta \ln y_i^*} \Big|_{\delta T=0} - \sum_j \frac{p_j \bar{c}_j(z)}{\bar{x}(z)} \frac{\delta \ln p_j}{\delta \ln y_i^*} \Big|_{\delta T=0} \right] dz - \int \tau(z) \frac{\bar{x}(z)}{p_i^* y_i^*} \frac{\delta \ln \bar{n}(z)}{\delta \ln \bar{w}(z)} \Big|_{\delta T=0} \frac{\delta \ln \bar{w}(z)}{\delta \ln y_i^*} \Big|_{\delta T=0} dz. \quad (2)$$

While Proposition 1 follows mechanically from Lemma 1, it offers a strict generalization of, as well as a new perspective on, the tax formulas found in the political economy of trade literature. The models discussed in Helpman's (1997) review, for instance, focus on the special case with quasi-linear preferences, with inelastic labor supply, without labor income taxation, and with sector-specific factors of production. Under quasi-linear preferences, $u(v(c), n) = c_1 + \sum_{i \neq 1} v_i(c_i) - h(n)$, our formula becomes

$$t_i^* = \int \left\{ \left[1 - \frac{\bar{\lambda}(z)}{\int \bar{\lambda}(v) dv} \right] [1 - \tau(z)] - \tau(z) \varepsilon(z) \right\} \frac{\bar{x}(z)}{p_i^* y_i^*} \frac{\delta \ln \bar{w}(z)}{\delta \ln y_i^*} \Big|_{\delta T=0} dz, \quad (3)$$

where $\varepsilon(z) \equiv d \ln h'(n(z)) / d \ln n$ denotes the elasticity of labor supply with respect to the wage at quantile z , i.e. the percentage change in labor supply caused by a one-percent change in the wage of any household with earnings initially in that quantile, holding marginal tax rates fixed. Under the other three restrictions, with each z now corresponding to the index of a specific factor, so that $\frac{\delta \bar{w}(z)}{\delta y_i^*} \Big|_{\delta T=0} = \frac{d \bar{w}(z)}{d y_i^*} = 0$ for $z \neq i$, one obtains the following corollary of Proposition 1.

Corollary 1. *Suppose that preferences are quasi-linear, labor supply is inelastic, labor income cannot be taxed, and factors are sector-specific. Then the optimal tax on good $i \neq 1$ satisfies*

$$t_i^* = \left(\frac{\bar{\lambda}(i)}{\int \bar{\lambda}(v) dv} - 1 \right) \times \left(- \frac{1}{p_i^*} \frac{d \bar{x}(i)}{d y_i^*} \right).$$

This expression coincides with the tariff formula derived by Helpman (1997) for various political-economy models, including the lobbying model of Grossman and Helpman (1994), where tariffs are expressed as a function of the the ratio of domestic output to

imports, the elasticity of import demand, and a dummy variable that captures whether sectors are politically organized or not.

Our alternative way of expressing optimal tariffs succinctly captures the essence of trade protection as predistribution. Corollary 1 states that protection in a given sector i , as measured by the gap between the domestic and foreign prices, $p_i - p_i^* = t_i^* p_i^*$, should be equal to the Pareto weight that the government puts on workers from sector i (relative to other sectors) times the marginal impact of a decrease in imports on the earnings of these workers. Different political-economy models—direct democracy (Mayer, 1984), political support function (Hillman, 1982), tariff formation function (Findlay and Wellisz, 1982), electoral competition (Magee et al., 1989), and influence-driven contributions (Grossman and Helpman, 1994)—simply correspond to different Pareto weights, $\bar{\lambda}(i)$.

No change in (the distribution of) utility. Our second variation complements changes in taxes on new technology firms δt^* with changes in the income tax schedule δT that keeps the distribution of utility unchanged. This amounts to setting $\delta T(z)$ equal to the change in real incomes $[1 - \tau(z)]\bar{n}(z)\delta w(z) - \sum_j \bar{c}_j(z)\delta p_j$ at all quantiles z of the income distribution.⁹ As demonstrated in Appendix D.3, this second variation leads after simplifications to the following optimal tax formula.

Proposition 2 ($\delta \bar{U} = 0$). *The optimal tax on good $i \neq 1$ satisfies*

$$t_i^* = \int \tau(z) \frac{\bar{x}(z)}{p^* y_i^*} \frac{\varepsilon_H(z)}{\varepsilon_M(z) + 1} \frac{\delta \ln \omega(z)}{\delta \ln y_i^*} \Big|_{\delta \bar{U}=0} dz, \quad (4)$$

with $\omega(z) \equiv \bar{w}'(z)/\bar{w}(z)$ the growth rate of the wage, $\varepsilon_H(z)$ the Hicksian labor supply elasticity, and $\varepsilon_M(z)$ the Marshallian labor supply elasticity.

The main attractive feature of Proposition 2 is that it does not require any information on the welfare weights $\bar{\lambda}(z)$. This is because, by construction, the distribution of utility is unaffected by our variation; thus, social welfare is unaffected. As a result, distributional considerations vanish and only efficiency considerations remain, captured by the labor fiscal externality. The key step in the proof of Proposition 2 is to show that this fiscal externality takes an extremely simple form, with the change in labor supply equal to

$$\delta \ln \bar{n}(z) = -\frac{\varepsilon_H(z)}{\varepsilon_M(z) + 1} \delta \ln \omega(z). \quad (5)$$

⁹In accompanying the tax reform of interest, here changes in taxes on new technology firms, with a distributively offsetting adjustment to income taxes, we follow the same general strategy as Kaplow (2010). A detailed analysis of the properties of such adjustments can be found in Tsyvinski and Werquin (2018).

To understand why changes in labor supply must be related in this way to changes in the growth rate of wages, it is convenient to start from the incentive compatibility constraint,

$$\bar{U}(z) = \max_{z'} u(\bar{C}(z'), \bar{n}(z') \frac{\bar{w}(z')}{\bar{w}(z)}), \quad (6)$$

where $\bar{C}(z)$ is the indirect utility from consumption, given income $\bar{w}(z)\bar{n}(z) - T(\bar{w}(z)\bar{n}(z))$ and prices p . The Envelope Theorem implies $\bar{U}'(z) = -\bar{u}_n(z)\omega(z)\bar{n}(z)$ with $\bar{u}_n(z) \equiv u_n(\bar{C}(z), \bar{n}(z))$. Since our variation holds the distribution of utility unchanged, $\delta\bar{U}(z) = 0$, it must also satisfy $\delta\bar{U}'(z) = \delta[-\bar{u}_n(z)\omega(z)\bar{n}(z)] = 0$, or in logs,

$$\delta \ln \bar{n}(z) = -\delta \ln \omega(z) - \delta \ln(-\bar{u}_n(z)).$$

In the special case where preferences are quasi-linear, $u(v(c), n) = c_1 + \sum_{i \neq 1} v_i(c_i) - h(n)$, $\delta \ln(-\bar{u}_n(z)) = \frac{1}{\varepsilon(z)} \delta \ln \bar{n}(z)$, with $\varepsilon(z) \equiv d \ln h'(n(z)) / d \ln n$, and equation (5) follows. The same logic extends to general preferences provided that differences between Hicksian and Marshallian labor supply elasticities are accounted for, as shown in Appendix D.3.

For purposes of implementation, the fact that welfare weights do not enter the formula in Proposition 2 must be welcomed. It opens up the possibility of a purely empirical evaluation, without any subjective choice over the social welfare objective. Our formula effectively rules out a first-order dominant improvement in the distribution of utilities. Our general strategy is reminiscent of the one used by Dixit and Norman (1986) to show the existence of Pareto gains from trade by constructing commodity taxes such that all households are kept at the same utility level under free trade, while simultaneously increasing the fiscal revenues of the government. Here, we show that unless the formula in Proposition 2 holds, one can also construct changes in taxes that increase fiscal revenues, while holding utility fixed at all quantiles of the wage distribution.¹⁰

4.3 Discussion

Wage Manipulation as Predistribution. Despite their differences, all our formulas give center stage to the change in the wage schedule, as either captured by the change in the wage level w , as in Proposition 1 and Corollary 1, or wage growth ω , as in Proposition 2. Our formulas make clear that changes in the quantiles of the wage distribution, which

¹⁰The existence of a nonlinear income tax schedule T plays a crucial role, as evidenced by the presence of the marginal tax rates $\tau(z)$. Everything else being equal, higher marginal taxes $\tau(z)$ potentiate fiscal externalities and demand larger t_i^* . To take an extreme case, if marginal taxes were zero, $\tau(z)$, then the formula immediately implies $t_i^* = 0$. This should come as no surprise: the First Welfare Theorem holds in our environment, so the absence of taxation leads to a Pareto optimum that cannot be improved upon.

may be of empirical interest for descriptive reasons, is actually a sufficient statistic for optimal policy design. Given knowledge of this statistic, the underlying structure of the economy leading to the change in wages can be left in the background.

Unlike in [Diamond and Mirrlees \(1971a,b\)](#) and [Dixit and Norman \(1980\)](#), the government here cannot achieve its distributional objectives by taxing workers of different types at different rates. To achieve the same objectives, it is now forced to predistribute by manipulating wages before taxes. It does so through taxes t^* that affect the prices p that firms face, and, in turn, their demand for workers of different types, as in [Naito \(1999\)](#).

Note that the fact that only the change in the growth rate $\omega(z) = w'(z)/w(z)$, a measure of inequality, rather than the level $w(z)$, appears in [Proposition 2](#) immediately implies that a zero tax is optimal if the impact on wages is proportionally uniform across z . This reflects the fact that only relative wages matter for incentives. To see this more formally, consider again the incentive compatibility constraint (6). In [Propositions 2](#), $\omega(z)$ is the local counterpart to $\frac{\bar{w}(z')}{\bar{w}(z)}$. It captures the fact that a household of type z that earns the same amount as one of type z' must work $\hat{n}(z, z')$ where $\bar{w}(z)\hat{n}(z, z') = \bar{w}(z')\bar{n}(z')$. Hence, changes in relative wages may tighten or loosen incentive compatibility constraints, not changes in the overall wage level. This is the same mechanism at play as in [Stiglitz \(1982\)](#) and [Rotschild and Scheuer \(2013\)](#), albeit in environments without commodity taxation.

Taxes on Old and New Technology Firms. Our previous formulas have assumed that only taxes on new technology firms were available. What if taxes on old technology firms were available as well? In a trade context, this would mean the possibility of imposing production taxes rather than import tariffs or export taxes. We now describe how our results extend to environments where the government can create wedges between the prices faced by old technology firms, $p \equiv \{p_i\}$, the prices faced by new technology firms, $p^* \equiv \{p_i^*\}$, and the prices faced by consumers, $q \equiv \{q_i\}$, with $p_i = (1 + t_i)p_i^*$ and $q_i = (1 + t_i^*)p_i^*$. In such environments, the trade-off between efficiency and redistribution described in [Lemma 1](#) generalizes to

$$\begin{aligned} & - \int \sum_i t_i^*(p_i^* \bar{c}_i(z)) \delta \ln \bar{c}_i(z) dz + \sum_i t_i(p_i^* y_i) \delta \ln y_i - \int \tau(z) \bar{x}(z) \delta \ln \bar{n}(z) dz \\ & = \int [\bar{\lambda}(z) - 1] \bar{x}(z) [(1 - \tau(z)) \delta \ln \bar{w}(z) - \frac{\delta T(z)}{\bar{x}(z)} - \sum_i \frac{q_i \bar{c}_i(z)}{\bar{x}(z)} \delta \ln q_i] dz. \quad (7) \end{aligned}$$

Compared to [equation \(1\)](#), the first term on the left-hand side is now split into two, reflecting the fact that the fiscal externality associated with changes in consumption and output are now different. In addition, the price deflator on the right-hand side is now given by

$\sum_i \frac{q_i \bar{c}_i(z)}{\bar{x}(z)} \delta \ln q_i$, reflecting the fact that consumer prices are now given by q rather than p .

Using the standard [Atkinson and Stiglitz's \(1976\)](#) logic, one can show that for a feasible variation not to improve welfare, q and p^* should be equalized. That is, there should only be a wedge between old and new technology firms, but not households and new technology firms. One can view the absence of the second wedge as an expression of the Targeting Principle. Since wages depend on the labor demand of those firms, redistribution through wage manipulation is best achieved by manipulating p , without introducing any additional consumption distortion by manipulating q .¹¹

Given the equality between q and p^* , one can use the same steps as in Section 4.2 to go from equation (7) to each of our formulas. The only difference between our old formulas and the new ones is that the differentials on the right-hand side should now be taken with respect to δy rather than δy^* , reflecting again the fact that only the output of old technology firms is being distorted, not the consumption of households.¹²

A Pigouvian Interpretation. Interestingly, our formulas provide a direct expression for the tax rate. This differs from the optimal linear tax literature (e.g. [Diamond and Mirrlees, 1971b](#)), which usually derives a system of simultaneous equations, with the entire set of tax rates on the left hand side. We could have stated our formulas in such forms by focusing on price variations such that $\delta p_i \neq 0$ and $\delta p_j = 0$ for $j \neq i, 1$. In vector and matrix notation, the formula in Proposition 2, for instance, would then become

$$[D_p y^*]_{\delta \bar{U}=0} (p^* - p) = \int \tau(z) \bar{x}(z) \frac{\varepsilon_H(z)}{\varepsilon_M(z) + 1} [\nabla_p \ln \omega(z)]_{\delta \bar{U}=0} dz$$

with $[D_p y^*]_{\delta \bar{U}=0} \equiv \{(\delta y_j^* / \delta p_i) |_{\delta \bar{U}=0}\}$ and $[\nabla_p \omega(z)]_{\delta \bar{U}=0} \equiv \{(\delta \omega(z) / \delta p_i) |_{\delta \bar{U}=0}\}$. Our formulas, which focus on quantity variations, are more akin to the Pigouvian tax literature, which provides an explicit expression for the tax on each good in terms of its externality. Indeed, we favor a Pigouvian interpretation of our formulas, as correcting for distributional externalities: if an extra unit of y_i^* is produced, then this has an impact on the wage schedule that, in turn, affects distribution and social welfare; the tax asks agents to pay

¹¹[Mayer and Riezman \(1987\)](#) establish a similar result in a trade context with inelastic factor supply and no income taxation. If both producer and consumer taxes are available, they show that only the former should be used. This result, however, requires preferences to be homothetic, as discussed in [Mayer and Riezman \(1989\)](#). Our result does not require this restriction. In fact, the quasi-homotheticity of the subutility v is not necessary either. This reflects the fact that we have access to non-linear income taxation and that preferences are weakly separable. Hence, absent the wage manipulation motive, there is no rationale for commodity taxation, as in [Atkinson and Stiglitz \(1976\)](#).

¹²The counterpart of Corollary 1 provides a strict generalization—to an environment with endogenous labor supply, nonlinear income taxation, as well as general preferences and technology—of the optimal production taxes in [Dixit \(1996\)](#).

for these marginal effects.

Full versus Partial Optimality. Provided that the government optimally chooses both the taxes on new technologies t^* and the income tax schedule T , the taxes on new technology firms given by Propositions 1 and 2 must coincide, for any given welfare function. In this situation, both formulas highlight the best way to complement income taxation with predistribution in order to achieve some redistributive objectives. The fact that social preferences explicitly enter the first formula but not the second reflects the fact that in the latter, the welfare weights $\bar{\lambda}(z)$ are implicitly revealed by the marginal tax rates $\tau(z)$ after controlling for the distortionary cost of redistribution, as measured by the labor supply elasticities $\varepsilon_H(z)$ and $\varepsilon_M(z)$.¹³

Away from an optimum, there is a priori no reason for the taxes on new technologies offered by our two formulas to coincide. There should also be no presumption as to which formula is more useful. In such situations, each formula highlights an alternative way to increase welfare through a different marginal change in taxes. Specifically, if taxes on good i observed in equilibrium are lower than those predicted by the right-hand side of (2), then a small change in t^* designed to lower y_i^* , while holding the shape of the income tax schedule fixed and rebating the proceed in a lump-sum fashion, increases social welfare $\mathcal{W}(\bar{U})$.¹⁴ Likewise, if taxes on good i observed in equilibrium are lower than those predicted by the right-hand side of (4), then a small change in t^* that lowers y_i^* , while reforming income taxes to keep the distribution of utility fixed, can be used to create a fiscal surplus (that can later be rebated in a lump-sum fashion in order to increase $\bar{U}(z)$ for all z). As usual, if the sufficient statistics appearing on the right-hand side of a formula were to be unaffected by these marginal tax reforms, then the optimal tax on good i could still be read directly from that formula, despite the economy not being initially at an optimum.¹⁵

Absent full optimality, it should also be clear that any of our formulas may fail to detect welfare-improving changes in t^* . For instance, consider the extreme case where there are no commodity taxes and no income taxes observed in equilibrium. In this situation, Proposition 2 leads to $t_i^* = 0$, which is what a utilitarian government would have found

¹³This is the idea behind [Werning's \(2007\)](#) test of whether an income tax schedule is Pareto optimal. Namely, it is if the inferred Pareto weights are all positive.

¹⁴Recall that any of our tax variations also holds fixed the output of other goods by new technology firms: $y_j^* = 0$ for all $j \neq i$. In the general environment that we consider, lowering y_i^* may not necessarily consist in solely raising t_i^* , but instead a linear combination of such taxes. The mapping from changes in output to changes in taxes is simply given by the inverse of $D_p y^*$.

¹⁵Of course, starting from a situation where the formulas in Propositions 1 and 2 do not coincide, the right-hand side of at least one of these two formulas would have to vary along the path to an optimum.

optimal. But suppose that the government has Rawlsian preferences, so that the absence of taxes is undesirable. In this case, the government enjoys many ways of improving welfare. It could change the nonlinear tax schedule only, or it could also change the linear tax t^* along the lines of Proposition 1. In this case, the formula in Proposition 1 may detect an improvement for a particularly chosen welfare function, even if the formula in Proposition 2 does not, because, in point of fact, there is no Pareto improvement. While the formula in Proposition 2 is the one that fails to detect an improvement in this example, it is easy to construct examples where the opposite happens.¹⁶

Finally, it is worth stressing the obvious: if the formulas in Propositions 1 and Proposition 2 do not coincide, then the income tax schedule is not locally optimized. Thus, rather than change t^* to improve welfare, one may prefer to simply change the income tax T . As we show in an example in Section 6.3, this may have much larger welfare gains.

5 Putting the Formulas to Work

We now illustrate through two examples, robots and trade, how our theoretical results can be combined with existing reduced-form evidence to provide estimates of optimal taxes.

5.1 Preliminary

We restrict attention to the formula displayed in Proposition 2. The main benefit of this formula is that it which allows us to dispense with any assumption on welfare weights. The main potential drawback is that the wage elasticity appearing on the right-hand side, $\frac{\delta \ln \omega(z)}{\delta \ln y_i^*} \Big|_{\delta \bar{U}=0}$, a priori depends on the details of the variation of the income tax schedule that holds utility fixed.¹⁷ Since this particular variation is unlikely to have been observed

¹⁶Suppose, for instance, that preferences are quasi-linear and that the marginal income tax schedule is such that $\frac{\tau(z)}{1-\tau(z)} = \frac{1}{\varepsilon(z)} \left[1 - \frac{\bar{\lambda}(z)}{\int \bar{\lambda}(v) dv} \right] \neq 0$. As can be seen from equation (3), the formula in Proposition 1 would lead to $t_i^* = 0$, which is what a government unable to affect the shape of its income tax schedule and with welfare weights $\{\bar{\lambda}(z)\}$ would have chosen. The formula in Proposition 2, however, shows that a small change in change in t^* that lowers y_i^* , accompanied by the proper income tax reform, could raise utility at all quantiles of the income distribution. Mathematically, the general issue illustrated by these two examples is that a derivative of the Lagrangian associated with the government's problem can be zero, even if social welfare is not maximized.

¹⁷This same issue does not affect the labor supply elasticities, $\varepsilon_H(z)$ and $\varepsilon_M(z)$, which only depend on the structure of preferences, as described in Appendix D.2. This stands in contrast to Proposition 1, where the elasticity $(\delta \ln \bar{n}(z) / \delta \ln \bar{w}(z)) \Big|_{\delta T=0}$ may also depend on these considerations, unless preferences are quasi-linear, in which case $(\delta \ln \bar{n}(z) / \delta \ln \bar{w}(z)) \Big|_{\delta T=0} = \varepsilon(z) = \varepsilon_H(z) = \varepsilon_M(z)$.

in practice, a natural question is whether there are other wage elasticities that could be used to implement Proposition 2. The next proposition answers in the affirmative.

Proposition 3 ($\delta\bar{U} = 0$, first-order approximation). *The optimal tax on good $i \neq 1$ satisfies*

$$t_i^* = \int \tau(z) \frac{\bar{x}(z)}{p_i^* y_i^*} \frac{\varepsilon_H(z)}{\varepsilon_M(z) + 1} \frac{\delta \ln \omega(z)}{\delta \ln y_i^*} \Big|_{\delta G^*=0} dz + O(\bar{\varepsilon}^2)$$

with $\delta G^* = 0$ a budget-balanced variation and $\bar{\varepsilon}$ such that $|\varepsilon_H(z)|, |\varepsilon_M(z)| < \bar{\varepsilon}$ for all $z \in [0, 1]$.

The formal proof is contained in Appendix E.1, but the basic idea is as follows. The difference in the wage responses across all budget-balanced variations stems from possible differences in labor supply. When labor supply is unchanged the two variations generate the exact same change in prices and wages. If labor supply changes are small, the differences are small. More precisely, the differences in labor supply are proportional to labor supply elasticities. In terms of the optimal tax rate, note that labor supply elasticities enter the formula in Proposition 2 multiplicatively. This then implies that substituting one wage elasticity for another creates a difference in labor supply of only second-order in terms of labor supply elasticities. As a result, to a first-order approximation the formula in Proposition 2 can ignore the differences in these indirect effects and, thus, employ wage elasticities derived under *any* budget-balanced variation, including our earlier $\delta T = 0$ variation.

Since estimates of labor supply elasticities are small, as reviewed recently in Chetty (2012), we will employ the first-order approximation in Proposition 3 to explore the magnitude of an optimal ad-valorem tax t_m^* on either robots or imports,

$$t_m^* \simeq \int \tau(z) \frac{\bar{x}(z)}{p_m^* y_m^*} \frac{\varepsilon_H(z)}{\varepsilon_M(z) + 1} \frac{\delta \ln \omega(z)}{\delta \ln y_m^*} \Big|_{\delta G^*=0} dz. \quad (8)$$

5.2 Quantitative Example (I): Robots

To calculate the efficient tax on robots in the United States using equation (8), the key input is the elasticity of relative wages with respect to the number of robots, $\frac{\delta \ln \omega(z)}{\delta \ln y_m^*} \Big|_{\delta G^*=0}$. To recover that elasticity, the ideal experiment would engineer a marginal change in taxes that increases the number of robots in the entire United States by one unit, while holding the government's budget balance, and record the differential changes in wages between consecutive quantiles of the income distribution. As a first proxy for such an ideal experiment, we propose to use the empirical estimates from Acemoglu and Restrepo (2017b). Using a difference-in-difference strategy, the previous authors have estimated the effect

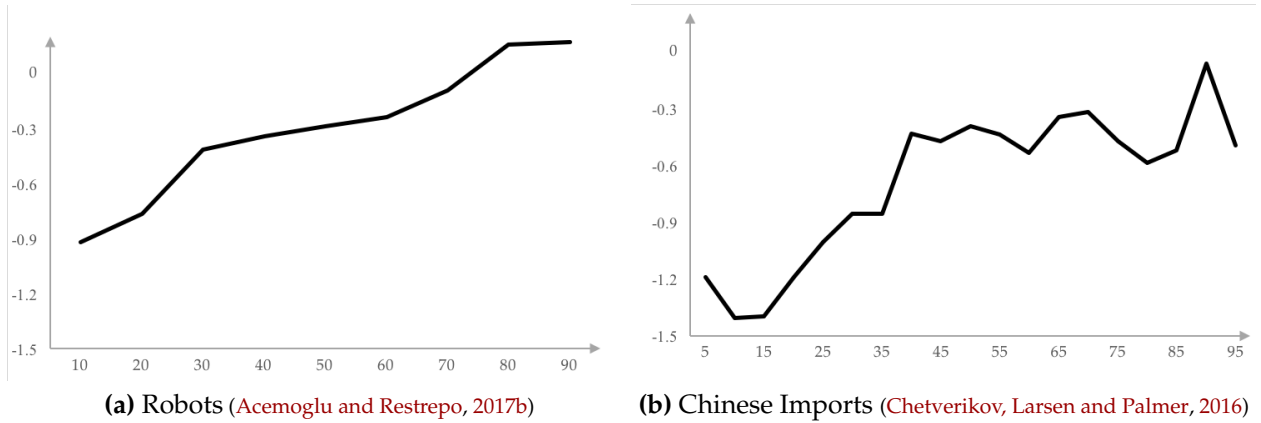


Figure 1: Semi-Elasticity of wages, $\frac{\delta \ln \omega(z)}{\delta y_m^*} \times 100$, across quantiles of US wage distribution.

of industrial robots, defined as “an automatically controlled, reprogrammable, and multipurpose [machine]” on different quantiles of the wage distribution between 1990 and 2007 across US commuting zones. We interpret their estimates as the semi-elasticity of wages with respect to robots, $\eta_{AR}(z) \simeq \frac{\delta \ln \bar{\omega}(z)}{\delta y_m^*} |_{\delta G^*=0}$, where y_m^* is expressed as number of robots per thousand workers. These estimates are reported in Figure 1a.

Under the previous interpretation, the elasticity that we are interested in is given by

$$\frac{\delta \ln \omega(z)}{\delta \ln y_m^*} |_{\delta G^*=0} = \frac{y_m^*}{\Delta \ln \bar{\omega}(z)} \times \Delta \eta_{AR}(z), \quad (9)$$

where $\Delta \ln \bar{\omega}(z)$ and $\Delta \eta_{AR}(z)$ denote changes between consecutive deciles of the wage distribution. In the United States in 2007, the number of robots per thousand workers is slightly greater than one, $y_m^* \simeq 1.2$, as reported by Acemoglu and Restrepo (2017b). This leads to an average elasticity $\frac{\delta \ln \omega(z)}{\delta \ln y_m^*} |_{\delta G^*=0}$ across deciles around 0.5%.

Given estimates of the previous elasticities, the only additional information required to evaluate the optimal tax on robots given by equation (8) is: (i) the marginal income tax rates, $\tau(z)$; (ii) the ratio of labor earnings, $\bar{x}(z)$, to total spending on robots, $p_m^* y_m^*$; and (iii) the Hicksian and Marshallian labor supply elasticities, $\varepsilon_H(z)$ and $\varepsilon_M(z)$. Table E.1 in Appendix E.2 describes how we obtain estimates each of these variables. For our baseline computation, we use US statutory marginal tax rates, as reported in Guner, Kaygusuz and Ventura (2014), and US labor earnings from the World Wealth and Income Database together with US spending on robots from Graetz and Michaels (2018). This leads to a ratio of total labor earnings to total spending on robots equal to 245. Finally, we set the Hicksian labor supply elasticities equal to 0.5, consistent with Chetty (2012), and impose

	Baseline	$\varepsilon_M = 0$	Effective τ	Fitted $\frac{\delta \ln \bar{w}}{\delta y_m^*}$	Fitted $\frac{\delta \ln \omega}{\delta \ln y_m^*}$
	(1)	(2)	(3)	(4)	(5)
Robots Tax:					
t_m^*	4.28%	6.42%	3.83%	5.38%	4.49%
Chinese Imports Tax:					
t_m^*	0.07%	0.10%	0.12%	0.02%	0.19%

Table 1: Optimal Taxes on Robots and Chinese Imports

Notes: Column (1) reports the optimal taxes on robots and Chinese imports obtained from equation (8) under our baseline calibration. Column (2)-(5) report the same taxes using instead high values of the labor supply elasticities ε_H and ε_M , effective marginal tax rates $\tau(z)$, fitted values of the semi-elasticities $\frac{\delta \ln \bar{w}}{\delta y_m^*}$ and elasticities $\frac{\delta \ln \omega}{\delta \ln y_m^*}$, respectively, as further described in Appendix E.

the same value for the Marshallian labor supply elasticity.

Column (1) of Table 1 reports the value of the optimal tax on robots under our baseline assumptions. Despite the negative impact of robots on wage inequality documented by [Acemoglu and Restrepo \(2017b\)](#), it is fairly small with $t_m^* \simeq 4.28\%$. By Slutsky, if leisure is a normal good, $\varepsilon_M(z) \leq \varepsilon_H(z)$; so setting $\varepsilon_M(z) = \varepsilon_H(z) = 0.5$ would bias downward our estimate of the optimal tax on robots. Column (2) of Table 1 considers the extreme case in which we instead set $\varepsilon_M(z) = 0$. This leads to $t_m^* \simeq 6.42\%$. We also obtain similar estimates when instead of using statutory marginal tax rates, we use the effective marginal rates reported in [Guner, Kaygusuz and Ventura \(2014\)](#), as can be seen from Column (3). Lastly, we explore the extent to which mismeasurement in the semi-elasticity of wages, $\delta \ln \bar{w}(z) / \delta y_m^*$, or the associated elasticity of relative wages, $\delta \ln \omega(z) / \delta \ln y_m^*$, may be driving our conclusions. To do so, we start by projecting $\delta \ln \bar{w}(z) / \delta y_m^*$ on a flexible polynomial, as further described in Appendix E.2, and then use the fitted values to compute $\delta \ln \omega(z) / \delta \ln y_m^*$ using equation (9). As shown in Column (4), the optimal tax on robots is only slightly higher than in our baseline. Our final exercise, in Column (5), directly projects the elasticity of relative wages $\delta \ln \omega(z) / \delta \ln y_m^*$ from our baseline computations on a flexible polynomial before implementing equation (8). In this case, the tax on our robots goes down to $t_m^* \simeq 4.49\%$.

5.3 Quantitative Example (II): Chinese Imports

Our second example focuses on Chinese imports. Like in the case of robots, we propose to use estimates obtained from a difference-in-difference strategy as a proxy for the elasticity that we are interested in, $\frac{\delta \ln \omega(z)}{\delta \ln y_m^*} |_{\delta G^*=0}$. Using the same empirical strategy as in [Autor, Dorn and Hanson \(2013\)](#), [Chetverikov, Larsen and Palmer \(2016\)](#) have estimated the effect on log wages of a \$1,000 increase in Chinese imports per worker at different percentiles

of the wage distribution, as described in Figure 1b. Following the same approach as in the case of robots, we can transform the previous semi-elasticities into elasticities using $y_m^* \simeq 2.2$ as the value of Chinese imports, in thousands of US dollars, per worker for the United States in 2007.

Interestingly, the average value of the relative wage elasticity, $\frac{\delta \ln \omega(z)}{\delta \ln y_m^*} |_{\delta G^*=0}$, is of the same order of magnitude as the one implied by the estimates of Acemoglu and Restrepo (2017b), around 0.5%. Compared to the robot example, however, the ratio of total labor earnings to total Chinese imports in 2007 is only 26.4, an order of magnitude smaller than the ratio to total spending on robots. This leads to an optimal tax on Chinese imports that is even smaller than the tax on robots in our baseline computation, $t_m^* \simeq 0.07\%$, as reported in Column (1) of Table 1, and remains an order of magnitude smaller in all our sensitivity exercises, as can be seen in Columns (2)-(5).

6 Comparative Statics

Our final results focus on comparative static issues in a special case of the environment presented in Section 3 where equilibrium variables and optimal taxes can be solved for in closed-form. In this environment, we first study how optimal taxes on new technologies vary with technological progress. We then explore how the previous taxes vary with social preferences, both when income taxes are set optimally and when they are fixed.

6.1 A Simple Economy with Heterogeneous Households and Machines

There is one final good, indexed by f , and one intermediate good, indexed by m , which could be a robot produced domestically or a machine imported from abroad.

Preferences. Households have one-dimensional skills θ , uniformly distributed over $[0, 1]$, and quasi-linear preferences,

$$U(\theta) = C(\theta) - \frac{(n(\theta))^{1+1/\varepsilon}}{1+1/\varepsilon}, \quad (10)$$

with $C(\theta)$ the consumption of the unique final good, which we use as our numeraire, $p_f = p_f^* = 1$, and ε the constant labor supply elasticity.

Technology. Old technology firms produce the final good, $y_f \geq 0$, using workers, $n \equiv \{n(\theta)\}$, and machines, $y_m \leq 0$, as an input. Their production set is given by

$$G(y_f, y_m, n) = y_f - \max_{\{\tilde{y}_m(\theta)\}} \left\{ \int g(\tilde{y}_m(\theta), n(\theta); \theta) dF(\theta) \mid \int \tilde{y}_m(\theta) dF(\theta) \leq -y_m \right\}, \quad (11)$$

with $g(\tilde{y}_m(\theta), n(\theta); \theta)$ a Cobb-Douglas production function,

$$g(\tilde{y}_m(\theta), n(\theta); \theta) = \exp(\alpha(\theta)) \cdot \left(\frac{\tilde{y}_m(\theta)}{\beta(\theta)} \right)^{\beta(\theta)} \left(\frac{n(\theta)}{1 - \beta(\theta)} \right)^{1 - \beta(\theta)}, \quad (12)$$

where $\tilde{y}_m(\theta)$ represents the number of machines combined with workers of type θ to produce the final good, $\alpha(\theta) \equiv \frac{\alpha \ln(1-\theta)}{\beta \ln(1-\theta) - 1}$, and $\beta(\theta) \equiv \frac{\beta \ln(1-\theta)}{\beta \ln(1-\theta) - 1}$, with $\alpha, \beta > 0$. New technology firms produce machines, $y_m^* \geq 0$, using the final good, $y_f^* \leq 0$,

$$G^*(y_f^*, y_m^*; \phi) = \phi y_f^* + y_m^*, \quad (13)$$

where ϕ measures the productivity of machine producers.

Equilibrium Wage Schedule. Let p_m and p_m^* denote the price of robots faced by old and new technology firms. Profit maximization by new technology firms implies

$$p_m^* = 1/\phi,$$

whereas profit maximization by old technology firms implies

$$w(p_m; \theta) = (1 - \theta)^{-1/\gamma(p_m)},$$

with $\gamma(p_m) \equiv 1/(\alpha - \beta \ln p_m)$. Under the restriction that $\gamma(p_m) > 0$, which we maintain throughout, wages are increasing in θ and Pareto distributed with shape parameter equal to $\gamma(p_m)$ and lower bound equal to 1. Note that since workers' skill θ is uniformly distributed over $[0, 1]$, wages are strictly increasing in θ , and labor supplies are strictly increasing in wages, the index θ also corresponds to a worker's quantile in the distribution of earnings.

By construction, more skilled workers tend to use machines relatively more, since $\beta(\theta)$ is increasing in θ . So an increase in the price of machines tends to lower their wages relatively more, which decreases inequality,

$$\frac{d \ln \omega(\theta)}{d \ln p_m} = -\frac{d \ln \gamma(p_m)}{d \ln p_m} = -\beta \gamma(p_m) < 0.$$

Here, because of additive separability in production, machines directly affect inequality by affecting the relative marginal products of workers of different skills, but not indirectly through further changes in their relative labor supply, as in [Stiglitz \(1982\)](#).

Social Welfare. The government aims to maximize a linear social welfare function,

$$\mathcal{W}(U) = \int U(\theta) d\Lambda(\theta), \quad (14)$$

$$\Lambda(\theta) = \lambda + \theta(1 - \lambda). \quad (15)$$

The social preference parameter $\lambda \in [0, 1]$ controls the government's preference for redistribution. At one extreme, $\lambda = 0$, Pareto weights are uniformly distributed, and the government is utilitarian. At the other extreme, $\lambda = 1$, the distribution of Pareto weights is a Dirac at $\theta = 0$, and the government is Rawlsian.

6.2 Technological Progress and Technology Regulation

Our first comparative static exercise studies how the optimal tax on machines t_m^* varies with the productivity of machine produces ϕ . For comparative static purposes, a limitation of our previous formulas is that they involve the entire schedule of marginal income tax rates. These are themselves endogenous objects that will respond to changes in ϕ . In [Online Appendix F.1](#), we demonstrate how to solve for $\tau(\theta)$ and obtain the following formula for the optimal tax on machines,

$$\frac{t_m^*}{1 + t_m^*} = \frac{\frac{\varepsilon}{\varepsilon+1} \frac{d \ln \omega}{d \ln y_m^*} \tau^*}{1 - \frac{\varepsilon}{\varepsilon+1} \frac{d \ln \omega}{d \ln y_m^*} \tau^*} \frac{1 - s_m}{s_m}, \quad (16)$$

where the elasticity of relative wages, $\frac{d \ln \omega}{d \ln y_m^*} \equiv -\beta \gamma(p_m) \frac{\partial \ln p_m}{\partial \ln |y_m(p_m, n)|}$, is now constant across agents; $\tau^* \equiv \frac{\varepsilon+1}{\varepsilon+1+\varepsilon \gamma(p_m)/\lambda}$ corresponds to the optimal marginal tax rate that would be imposed in the absence of a tax on machines, as in [Diamond \(1998\)](#), [Saez \(2001\)](#), and [Scheuer and Werning \(2017\)](#); and $s_m \equiv \frac{p_m y_m^*}{\int x(\theta) dF(\theta) + p_m y_m^*}$ measures the share of machines in gross output. After expressing the three previous statistics, $\frac{d \ln \omega}{d \ln y_m^*}$, τ^* , and s_m , as functions of t_m^* and ϕ , we can apply the Implicit Function Theorem to determine the monotonicity of the optimal tax on machines, as we do in [Online Appendix F.2](#). This leads to the following proposition.

Proposition 4. *In a simple economy where equations (10)-(15) hold, the optimal tax on machines t_m^* is decreasing with the productivity ϕ of new technology firms.*

By construction, more machines always increase inequality in this simple economy, $\frac{d \ln \omega}{d \ln y_m^*} > 0$. So one should always tax new technology firms. For comparative static purposes, however, the relevant question is whether this effect gets exacerbated as the new technology improves. Here, one can check that $\frac{\partial}{\partial \phi} \frac{d \ln \omega}{d \ln y_m^*} < 0$ both because relative wages are becoming less responsive to the price of machines, $\frac{\partial}{\partial \phi} \left| \frac{d \ln \omega}{d \ln y_m^*} \right| < 0$, and because the demand for machines is becoming more elastic, $\frac{\partial}{\partial \phi} \left| \frac{\partial \ln p_m}{\partial \ln |y_m(p_m, n)|} \right| < 0$, due to the increase in the labor supply of high-skilled workers whose demand for machines is more elastic. One can also check that these two effects dominate the increase in the marginal tax rate, $\frac{\partial \tau^*}{\partial \phi} > 0$, in response to greater inequality. For a given share of machines s_m , this implies that the total fiscal externality associated with new machines decreases. Since the share of machines increase with improvements in the new technology, $\frac{\partial s_m}{\partial \phi} > 0$, the fiscal externality per machine a fortiori decreases and so does the tax on machines.

As this simple example illustrates, cheaper robots may lead to a higher share of robots in the economy, more inequality, but a *lower* optimal tax on robots. Likewise, more imports and more inequality, in spite of the government having extreme distributional concerns and imports causing inequality, may be optimally met with less trade protection. This decline in t_m^* does not derive from redistribution becoming more costly as the economy gets more open.¹⁸ Here, the elasticity of labor supply is fixed and the marginal tax rate τ^* increases with ϕ . This decline also does not derive from redistribution through income taxation becoming more attractive. Everything else being equal, an increase in τ^* raises the tax on imports. Rather the decline in t_m^* predicted by Proposition 4 captures a standard Pigouvian intuition: as ϕ increases, the total fiscal externality associated with imports increases, but the marginal impact on fiscal revenues does not, leading to a lower value for the optimal tax.

6.3 Social Preferences, Income Taxes, and Technology Regulation

Our second comparative static exercise studies how the optimal tax on machines varies with the social preference parameter λ . We do so under two baseline scenarios. In the first one, consistent with our earlier analysis, we assume that the government can choose both the tax on machines t_m and the income tax schedule T in order to maximize social welfare. In the second scenario, we assume instead that income taxes are exogenously set at $T = T^c$, with T^c a linear tax schedule with constant marginal tax rate $\tau^c \in [0, 1]$.

We refer to the taxes on machines associated with these two scenarios as the optimal

¹⁸This is the point emphasized by [Itskhoki \(2008\)](#) and [Antras, de Gortari and Itskhoki \(2017\)](#) in an economy where entrepreneurs can decide whether to export or not. This makes labor supply decisions more elastic in an open economy, which may reduce redistribution at the optimum.

and constrained optimal taxes, t_m^* and t_m^c , respectively. As before, the optimal tax on machines t_m^* is given by equation (16). Under this scenario, since the government is at a full optimum, t_m^* is consistent with the formulas in either Propositions 1 or 2, as discussed in Section 4.3. In contrast, when income taxes are exogenously set so that $T = T^c \neq T^*$, the constrained optimal tax t_m^c has to be derived exclusively using the formula in Proposition 1, since the variation underlying Proposition 2's formula is no longer feasible. As shown in Online Appendix F.3, the constrained optimal tax on machines can be expressed as

$$\frac{t_m^c}{1 + t_m^c} = \frac{[\lambda(1 - \tau^c) - \tau^c \varepsilon][1 - (1 + \varepsilon)/\gamma(p_m)]}{2\beta(1 + \varepsilon) + 1 - (1 + \varepsilon)/\gamma(p_m)}, \quad (17)$$

Using equations (16) and (17), one can establish our final analytical result, whose formal proof can be found in Appendix F.3.

Proposition 5. *In a simple economy where equations (10)-(15) hold, the optimal and constrained optimal taxes on machines t_m^* and t_m^c are increasing with the social preference parameter λ .*

Since more machines increase inequality, the more the government cares about the poor, in the sense of a higher social preference parameter λ , the higher the tax on machines will be. This is true regardless of whether or not non-linear income taxes are available. In the case in which the income tax schedule is exogenously given, though, the constrained optimal tax on machines t_m^c may very well be a subsidy. From equation (17), we see that $t_m^c < 0$ if $\tau^c/(1 - \tau^c) > \lambda/\varepsilon$, in which case the role of a subsidy on machines is to undo the redistributive consequences of a too progressive income tax schedule.

To explore further how social preferences as well as inefficiencies in income taxes may shape the tax on machines as well as its welfare consequences, we conclude with numerical simulations. We start from a baseline economy that has no taxes on machines $t_m = 0$, a constant marginal income tax rate $\tau^c = 27\%$, a labor supply elasticity $\varepsilon = 0.5$, and technological parameters $\alpha = 0.57$, $\beta = 0.003$, and $\phi = 1$.¹⁹ We define the welfare gains from optimal taxation as the gains of moving from $(t_m = 0, T = T^c)$ to $(t_m = t_m^*, T = T^*)$, whereas the welfare gains from constrained optimal taxation are equal to the gains of moving from $(t_m = 0, T = T^c)$ to $(t_m = t_m^c, T = T^c)$.

Figure 2 describes how the optimal and constrained optimal taxes on machines t_m^* and t_m^c vary with social preferences, in Panel (a), as well as the welfare gains from optimal and constrained optimal taxation, in Panel (b). To help interpret magnitudes, the x-axis does not report λ itself, but rather the optimal marginal income tax rate τ^* that a government

¹⁹This implies that the elasticity of the relative wage $d \ln \omega / d \ln y_m^*$ is equal to 0.5%, as in the two previous quantitative examples of Section 5, and the share of machines s_m is equal to 2%, the average between the shares in the robots and trade example from Section 5. See Appendix F.4 for details.

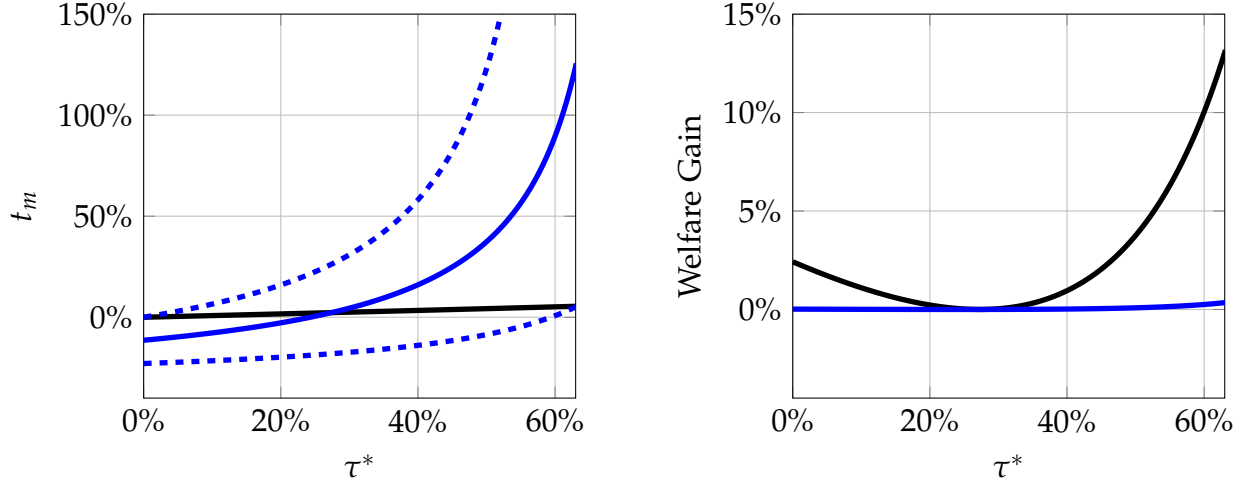


Figure 2: Social Preferences, Income Taxes, and Technology Regulation

Notes: Panel (a) plots the optimal ad-valorem tax t_m^* (in black) and the constrained optimal tax t_m^c (in blue) as a function of the optimal marginal income tax rate $\tau^* \equiv \frac{\varepsilon+1}{\varepsilon+1+\varepsilon\gamma(1)/\lambda}$. Panel (b) plots the welfare gains from optimal taxation (in black) and constrained optimal taxation (in blue), both in percentage of initial GDP. Preference and technological parameters are set such that $\varepsilon = 0.5$, $\alpha = 0.57$, $\beta = 0.003$, and $\phi = 1$. Marginal income tax rate is set to $\tau^c = 27\%$ when computing t_m^c and associated welfare gains.

with social preference parameter $\lambda \in [0, 1]$ and only access to income taxation would have chosen. This amounts to a change of variable, with $\tau^* \equiv \frac{\varepsilon+1}{\varepsilon+1+\varepsilon\gamma(1)/\lambda}$.

In Panel (a), we see that the optimal tax on machines t_m^* (in black) remains fairly stable and small, going from $t_m^* = 0\%$ in the utilitarian case $\tau^* = 0$ to $t_m^* = 5.43\%$ when optimal marginal income tax rate reaches $\tau^* = 63\%$ in the Rawlsian case. In contrast, the constrained optimal tax t_m^c (in blue) is very sensitive to the underlying preference parameter, going from $t_m^c = -11.39\%$ when $\tau^* = 0$ to $t_m^c = 125.63\%$ when $\tau^* = 63\%$.²⁰ This reflects the fact that taxes on machines, and only taxes on machines, may be used as an imperfect instrument to achieve the same redistributive objectives. For the same reason, the constrained optimal taxes on machines would be even larger in a baseline economy without income taxes, $\tau^c = 0$, and the constrained optimal subsidies even larger in a baseline economy with substantial income taxes, $\tau^c = 63\%$, as can be seen from the upper and

²⁰In the special case where the marginal income tax rate observed in the baseline economy reveals its social preferences for redistribution in the sense that λ is such that $\tau^* = \tau_c = 27\%$, we also see from Panel (a) that the fully and constrained optimal taxes on machines are very close: $t_m^* \simeq 2.28\%$ and $t_m^c \simeq 2.26\%$. The only source of discrepancy between the two taxes arises because at a full optimum, the income tax schedule should be adjusted to take into account the fact that the tax on machines is non-zero, which the constrained optimal tax ignores. Since the tax on machines is small, however, this extra round of adjustments in income taxes has a minor impact on the value of t_m^* relative to t_m^c . Likewise, since the baseline economy is very close to a full optimum for $\tau^* = \tau_c = 27\%$ and $t_m = 0$, one can check that using the formulas in Propositions 2 and 3, again ignoring that income taxes would have to be adjusted to reach the full optimum, leads to very similar prescriptions: $(t_m)_{\delta\bar{U}=0} \simeq 2.28\%$ and $(t_m)_{\delta\bar{U}=0, f.o.a} \simeq 2.23\%$, as shown in Appendix F.4.

lower blue dashed lines, respectively. In terms of welfare, though, Panel (b) shows that the returns to such second-best interventions (in blue) are orders of magnitude smaller than those of first-best interventions that would also reform income taxes (in black).

7 Concluding Remarks

How should government policy respond to technological change? Our answer is that in second-best environments—where income taxation is available, but taxes on specific factors are not—there is a case for taxing new technology firms, with each of our formulas offering a precise answer to what optimal taxes on new technology firms should be as a function of a few sufficient statistics.

Although our formulas differ in important ways, they all give center stage to changes in the wage schedule. This reflects a general Pigouvian motive for optimal technology regulation to correct distributional externalities. When one extra unit is produced using the new technology, either in the form of a robot or imports from abroad, this has an impact on the wage schedule that, in turn, affects distribution and social welfare; the optimal tax asks agents to pay for these marginal effects.

Perhaps surprisingly, we have also provided an example showing that more robots or more trade may go hand in hand with more inequality and lower taxes, despite robots or trade being responsible for the rise in inequality, and governments having preferences for redistribution. Although there is always a distributional externality to be corrected, the marginal impact of either robots or trade is what matters for the magnitude of the tax, and that marginal impact is always falling in this example. As a result, optimal taxes decrease as the process of automation and globalization deepens and inequality increases.

While we have focused on automation and globalization, our tax formulas could be used to shed light on a variety of salient social issues, from immigration to environmental regulations, that involve the same basic trade-off between efficiency and redistribution. We hope that our general theoretical analysis can help build bridges from existing empirical estimates of the redistributive consequences of such shocks to their ultimate policy implications.

References

- Acemoglu, Daron and Pascual Restrepo**, “The Race Between Machine and Man: Implications of Technology for Growth, Factor Shares and Employment,” *NBER Work. Pap.* 22252, 2017.
- **and** —, “Robots and Jobs: Evidence from US Local Labor Markets,” *NBER Working Paper* 23285, 2017.
- Antras, Pol, Alonso de Gortari, and Oleg Itskhoki**, “Globalization, Inequality and Welfare,” *Journal of International Economics*, 2017, 108, 387–412.
- Atkinson, A.B. and J.E. Stiglitz**, “The Design of Tax Structure: Direct versus Indirect Taxation,” *Journal of Public Economics*, 1976, 6, 55–75.
- Autor, D., D. Dorn, and G. Hanson**, “The China syndrom: Local labor market effects of import competition in the United States,” *American Economic Review*, 2013, 103, 2121–2168.
- Benzell, Seth G., Laurence J. Kotlikoff, Guillermo LaGarda, and Jeffrey D. Sachs**, “Robots Are Us: Some Economics of Human Replacement,” *NBER Work. Pap.* 20941, 2015.
- Chetty, Raj**, “Bounds on Elasticities with Optimization Frictions: A Synthesis of Micro and Macro Evidence on Labor Supply,” *Econometrica*, 2012, 80 (3), 969–1018.
- Chetverikov, Denis, Bradley Larsen, and Christopher John Palmer**, “IV Quantile Regression for Group-Level Treatments, with an Application to the Effects of Trade on the Distribution of Wage,” *Econometrica*, 2016, 84 (2), 809–833.
- Costinot, Arnaud and Ivan Werning**, “Robots, Trade, and Luddism: A Sufficient Statistic Approach to Optimal Technology Regulation,” *NBER Working Paper* 25103, 2018.
- Diamond, Peter A.**, “Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates,” *American Economic Review*, 1998, 88 (1), 83–95.
- **and James A. Mirrlees**, “Optimal Taxation and Public Production I: Production Efficiency,” *American Economic Review*, 1971, 61 (1), 8–27.
- **and** —, “Optimal Taxation and Public Production II: Tax Rules,” *American Economic Review*, 1971, 61 (3), 261–278.

- Dixit, Avinash and Victor Norman**, “Gains from Trade without Lump-Sum Compensation,” *Journal of International Economics*, 1986, 21, 111–122.
- Dixit, Avinash K.**, “Special-Interest Lobbying and Endogenous Commodity Taxation,” *Eastern Economic Journal*, 1996, 22 (4), 375–388.
- **and Victor Norman**, *Theory of International Trade*, Cambridge University Press, 1980.
- Feenstra, Robert C. and Tracy R. Lewis**, “Trade adjustment assistance and Pareto gains from trade,” *Journal of International Economics*, 1994, 36, 201–222.
- Findlay, Ronald and Stanislaw Wellisz**, *Endogenous tariffs, the political economy of trade restrictions, and welfare* Import Competition and Response, Chicago: University of Chicago Press, 1982.
- Graetz, Georg and Guy Michaels**, “Robots at Work,” *mimeo LSE*, 2018.
- Grossman, Gene M. and Elhanan Helpman**, “Protection for Sale,” *American Economic Review*, 1994, 84 (4), 833–850.
- **and Esteban Rossi-Hansberg**, “Trading Tasks: A Simple Theory of Offshoring,” *American Economic Review*, 2008, 98 (5), 1978–1997.
- Guerreiro, Joao, Sergio Rebelo, and Pedro Teles**, “Should Robots Be Taxed?,” *mimeo Northwestern University*, 2017.
- Guesnerie, Roger**, “Peut-on toujours redistribuer les gains a la specialisation et a l’échange? Un retour en pointille sur Ricardo et Heckscher-Ohlin,” *Revue Economique*, 1998, 49 (3), 555–579.
- Guner, Nezh, Remzi Kaygusuz, and Gustavo Ventura**, “Income Taxation of U.S. Households: Facts and Parametric Estimates,” *Review of Economic Dynamics*, 2014, 17 (4), 559–581.
- Helpman, Elhanan**, *Politics and Trade Policy* Advances in Economics and Econometrics: Theory and Applications, New York: Cambridge University Press, 1997.
- Hillman, Arye L.**, “Declining Industries and Political-Support Protectionist Motives,” *American Economic Review*, 1982, 72 (1180-1187).
- Hosseini, Roozbeh and Ali Shourideh**, “Inequality, Redistribution, and Optimal Trade Policy: A Public Finance Approach,” *mimeo Carnegie Mellon University*, 2018.

- Itskhoki, Oleg**, "Optimal Redistribution in an Open Economy," *mimeo Princeton University*, 2008.
- Jacobs, Bas**, "Optimal Inefficient Production," *mimeo Erasmus University*, 2015.
- Kaplow, Louis**, *The Theory of Taxation and Public Economics*, Princeton University Press, 2010.
- Magee, Stephen P., William A. Brock, and Leslie Young**, *Black Hole Tariffs and Endogenous Policy Formation*, Cambridge: MIT Press, 1989.
- Mayer, Wolfgang**, "Endogenous Tariff Formation," *American Economic Review*, 1984, 74, 970–985.
- **and Raymond G. Riezman**, "Endogenous Choice of Trade Policy Instruments," *Journal of International Economics*, 1987, 23, 377–381.
- **and** —, "Tariff Formation in a Multidimensional Voting Model," *Economics and Politics*, 1989, 61–79.
- Naito, Hisairo**, "Re-examination of uniform commodity taxes under non-linear income tax system and its implications for production efficiency," *Journal of Public Economics*, 1999, 1 (2), 65–88.
- , "Redistribution, production inefficiency and decentralized efficiency," *International Tax Public Finance*, 2006, 13, 625–640.
- Rotschild, Casey and Florian Scheuer**, "Redistributive Taxation in the Roy Model," *Quarterly Journal of Economics*, 2013, 128 (2), 623–668.
- Saez, Emmanuel**, "Using Elasticities to Derive Optimal Income Tax Rates," *Review of Economic Studies*, 2001, 68 (1), 205–229.
- Scheuer, Florian and Ivan Werning**, "The Taxation of Superstars," *Quarterly Journal of Economics*, 2017, 132, 211–270.
- Slavik, Ctirad and Hakki Yazici**, "Machines, buildings, and optimal dynamic taxes," *Journal of Monetary Economics*, 2014, 66, 47–61.
- Spector, David**, "Is it possible to redistribute the gains from trade using income taxation?," *Journal of International Economics*, 2001, 55, 441–460.

Stiglitz, Joseph E., “Self-selection and Pareto efficient taxation,” *Journal of Public Economics*, 1982, 17 (2), 213–240.

Thuemmel, Uwe, “Optimal Taxation of Robots,” *mimeo University of Zurich*, 2018.

Tsyvinski, Aleh and Nicolas Werquin, “Generalized Compensation Principle,” *mimeo Yale University*, 2018.

Werning, Ivan, “Pareto Efficient Income Taxation,” *mimeo MIT*, 2007.

A Notation

Consider a function $h : \mathbb{R}^K \rightarrow \mathbb{R}^M$ such that

$$h(x_1, \dots, x_K) = (h_1(x_1, \dots, x_K), \dots, h_M(x_1, \dots, x_K)).$$

Throughout our appendix, we use the following notation

$$h_{x_i}(x_1, \dots, x_K) \equiv \frac{\partial h(x_1, \dots, x_K)}{\partial x_i} \text{ for all } i = 1, \dots, K,$$

$$h_{j,x_i}(x_1, \dots, x_K) \equiv \frac{\partial h_j(x_1, \dots, x_K)}{\partial x_i} \text{ for all } i = 1, \dots, K \text{ and } j = 1, \dots, M,$$

Whenever there is no risk of confusion, we also drop arguments from functions so that, for instance, h implicitly stands for $h(x_1, \dots, x_K)$.

B Section 2

B.1 Tasks in Old Technology

Our environment nests economies in which a final good is produced using a continuum of tasks. To see this formally, consider an economy that produces a final good f using a continuum of tasks indexed by j ,

$$y_f = g_f(\{y(j)\}),$$

with $y(j) \geq 0$ the output of task j and g_f a concave and constant returns to scale production function. Each task, in turn, is produced using domestic workers and robots, as in [Acemoglu and Restrepo \(2017a\)](#), or domestic and foreign workers, as in [Grossman and Rossi-Hansberg \(2008\)](#),

$$y(j) = g_j(y_m(j), \{n(\theta, j)\}), \text{ for all } j,$$

with $y_m(j) \geq 0$ the number of robots or foreign workers used to perform task j , $n(\theta, j) \geq 0$ the number of domestic workers, and g_j a concave and constant return to scale production function. The production possibility frontier is then given by

$$G(y_f, y_m, n) = y_f - g(y_m, n)$$

with $y_f \geq 0$ the output of the final good, $y_m \leq 0$ the total amount of robots or foreign workers demanded by old technology firms, and g such that

$$g(y_m, n) = \max_{\{y(j), y_m(j), n(\theta, j)\}} g_f(\{y(j)\})$$

subject to

$$y(j) \leq g_j(y_m(j), \{n(\theta, j)\}), \text{ for all } j,$$

$$n(\theta) f(\theta) \geq \int n(\theta, j) dj, \text{ for all } \theta,$$

$$-y_m \geq \int y_m(j) dj,$$

with $f(\theta)$ the pdf associated with the distribution of types F . The homogeneity and convexity of G follows from g_f and g_j being concave and constant returns.

B.2 Labor in New Technology

In Section 2, we have argued that it is without loss of generality to assume the new technology G^* does not employ labor. Here, we provide the formal argument.

Suppose that the production sets associated with the old and new technology are such that

$$\begin{aligned}\hat{G}(\hat{y}, \hat{n}) &\leq 0, \\ \hat{G}^*(\hat{y}^*, \hat{n}^*) &\leq 0.\end{aligned}$$

First, define

$$y = \begin{pmatrix} \hat{y} \\ \hat{y}^* \end{pmatrix}.$$

Next, define G so that the set of (y, n) satisfying $G(y, n) \leq 0$ coincides with the set of (y, n) for which there exists \hat{n} and \hat{n}^* such that

$$\begin{aligned}\hat{G}(\hat{y}, \hat{n}) &\leq 0, \\ \hat{G}^*(\hat{y}^*, \hat{n}^*) &\leq 0, \\ \hat{n} + \hat{n}^* &= n.\end{aligned}$$

Last, define G^* such that $G^*(y^*) \leq 0$ is satisfied if and only if

$$y^* = \begin{pmatrix} \hat{y} \\ \hat{y}^* \end{pmatrix}$$

is such that $\hat{y} \leq -\hat{y}^*$.

C Section 3

C.1 Competitive Equilibrium with Taxes

We provide a formal definition of a competitive equilibrium with taxes.

Demand. Households maximize utility taking prices p , wages w , and the income tax schedule T as given. Since preferences are weakly separable, the demand of any household θ is given by the two-step problem

$$c(\theta) \in \operatorname{argmax}_{\tilde{c}(\theta)} \{v(\tilde{c}(\theta)) | p \cdot \tilde{c}(\theta) \leq w(\theta)n(\theta) - T(w(\theta)n(\theta))\}, \quad (\text{C.1})$$

$$n(\theta) \in \operatorname{argmax}_{\tilde{n}(\theta)} \{u(C(p, w(\theta)\tilde{n}(\theta) - T(w(\theta)\tilde{n}(\theta))), \tilde{n}(\theta))\}, \quad (\text{C.2})$$

where $C(p, r) \equiv \max_{\tilde{c}} \{v(\tilde{c}) | p \cdot \tilde{c} \leq r\}$ is the indirect subutility of a household facing prices p with after-tax earnings r and \cdot denotes the inner product of two vectors.²¹

²¹In Section 2, we have used $C(\theta)$ to denote the level of the subutility of household θ . For ease of notation, we use C again here to denote the indirect subutility function, $C(p, r)$. By definition, we therefore

Supply. Firms maximize profits taking prices p and p^* and wages w as given,

$$y, n \in \operatorname{argmax}_{\tilde{y}, \tilde{n}} \{p \cdot \tilde{y} - \int w(\theta) \tilde{n}(\theta) dF(\theta) \mid G(\tilde{y}, \tilde{n}) \leq 0\}, \quad (\text{C.3})$$

$$y^* \in \operatorname{argmax}_{\tilde{y}^*} \{p^* \cdot \tilde{y}^* \mid G^*(\tilde{y}^*; \phi) \leq 0\}. \quad (\text{C.4})$$

Market Clearing. Demand equals supply for all goods,

$$\int c_i(\theta) dF(\theta) = y_i + y_i^* \quad \text{for all } i. \quad (\text{C.5})$$

Linear Taxation. Prices satisfy

$$p_i = (1 + t_i^*) p_i^* \quad \text{for all } i. \quad (\text{C.6})$$

Government's Budget Constraint. The sum of taxes paid by new technology firms and income taxes paid by households is zero,

$$\sum_i t_i^* p_i^* y_i^* + \int T(w(\theta) n(\theta)) dF(\theta) = 0. \quad (\text{C.7})$$

Equilibrium. A competitive equilibrium with taxes (T, t^*) is an allocation $c \equiv \{c(\theta)\}$, $n \equiv \{n(\theta)\}$, $y \equiv \{y_i\}$, $y^* \equiv \{y_i^*\}$, prices and wages $p \equiv \{p_i\}$, $p^* \equiv \{p_i^*\}$, and $w \equiv \{w(\theta)\}$, such that:

- i. households maximize their utility, condition (C.1) and (C.2);
- ii. firms maximize their profits, conditions (C.3) and (C.4);
- iii. good markets clear, condition (C.5);
- iv. prices satisfy the non-arbitrage condition (C.6);
- v. the government's budget is balanced, condition (C.7).

C.2 Wage Schedule

Let $y(p, n)$ denote the supply of old technology firms, that is the solution to

$$\max_{\tilde{y}} \{p \cdot \tilde{y} \mid G(\tilde{y}, n) \leq 0\}.$$

The first-order conditions of (C.3) imply a wage schedule $w(p, n) \equiv \{w(p, n; \theta)\}$ such that

$$w(\theta) f(\theta) = -G_{n(\theta)}(y(p, n), n) / G_{y_1}(y(p, n), n) \quad \text{for all } \theta.$$

have $C(\theta) = C(p, r(\theta))$, with $r(\theta) = w(\theta) n(\theta) - T(w(\theta) n(\theta))$ the after-tax earnings of household θ .

C.3 Government Problem

For any vector of consumption and labor supply, $c \equiv \{c(\theta)\}$ and $n \equiv \{n(\theta)\}$, the utility of quantile $z \in [0, 1]$ is given by

$$\bar{U}(z) = \inf\{u \in \mathbb{R} \mid z \leq \int_{u(v(c(\theta)), n(\theta)) \leq u} dF(\theta)\}. \quad (\text{C.8})$$

For any y^* , construct p^* such that

$$p_i^* = G_{y_i^*}^*(y^*; \phi) / G_{y_1^*}^*(y^*; \phi) \text{ for all } i. \quad (\text{C.9})$$

Define the feasible set

$$\Omega = \{(c, n, y, y^*, p, p^*, w, T, t^*, \bar{U}) \text{ such that equations (C.1)-(C.3), (C.5)-(C.6), and (C.8)-(C.9) hold}\}.$$

Noting that profit maximization by new technology firms, condition (C.4), is equivalent to (C.9) and $G^*(y^*; \phi) = 0$ and that, by Walras' Law, the government's budget constraint, condition (C.7), necessarily holds if the other equilibrium conditions do, the problem of selecting a competitive equilibrium with taxes that maximizes social welfare can then be expressed as

$$\max_{(c, n, y, y^*, p, p^*, w, T, t^*, \bar{U}) \in \Omega} \mathcal{W}(\bar{U})$$

subject to

$$G^*(y^*; \phi) = 0.$$

D Section 4

D.1 Proof of Lemma 1

Preliminaries. The next lemma will be used to reduce the dimensionality from θ to the percentile z in our optimality conditions.

Lemma 2. Suppose $\theta \in \Theta \subseteq \mathbb{R}^K$, $\epsilon \in \mathbb{R}$ and the function $w(\theta, \epsilon)$ is almost everywhere differentiable and increasing in some θ_i , define

$$Z(W, \epsilon) = \int_{w(\theta, \epsilon) \leq W} dF(\theta)$$

and let $\bar{w}(z, \epsilon)$ denote the inverse of Z with respect to W . Then

$$\bar{w}_\epsilon(z, \epsilon) = \mathbb{E}[w_\epsilon(\theta, \epsilon) \mid w(\theta, \epsilon) = \bar{w}(z, \epsilon)].$$

Proof. Denote the inverse of $W = w(\theta, \epsilon)$ for θ_i as $\hat{\theta}(\theta_{-i}, W, \epsilon)$. Then

$$Z(W, \epsilon) = \int F(\hat{\theta}(\theta_{-i}, W, \epsilon) \mid \theta_{-i}) dF(\theta_{-i}). \quad (\text{D.1})$$

Since $w(\theta_{-i}, \hat{\theta}(\theta_{-i}, W, \epsilon), \epsilon) = W$ it follows that

$$\begin{aligned}\hat{\theta}_W(\theta_{-i}, W, \epsilon) &= \frac{1}{w_{\theta_i}(\theta_{-i}, \hat{\theta}(\theta_{-i}, W, \epsilon), \epsilon)}, \\ \hat{\theta}_\epsilon(\theta_{-i}, W, \epsilon) &= -\frac{w_\epsilon(\theta_{-i}, \hat{\theta}(\theta_{-i}, W, \epsilon), \epsilon)}{w_{\theta_i}(\theta_{-i}, \hat{\theta}(\theta_{-i}, W, \epsilon), \epsilon)}.\end{aligned}$$

Since $\bar{w}(Z(W, \epsilon), \epsilon) = W$ it follows that

$$\bar{w}_\epsilon(z, \epsilon) = -\bar{w}_z(z, \epsilon)Z_\epsilon(\bar{w}(z, \epsilon), \epsilon) = -\frac{1}{Z_W(\bar{w}(z, \epsilon), \epsilon)}Z_\epsilon(\bar{w}(z, \epsilon), \epsilon).$$

Using these expressions and differentiating (D.1) gives

$$\begin{aligned}\bar{w}_\epsilon(z, \epsilon) &= \frac{1}{\int \frac{f(\hat{\theta}(\theta_{-i}, \bar{w}(z, \epsilon), \epsilon) | \theta_{-i})}{w_{\theta_i}(\theta_{-i}, \hat{\theta}(\theta_{-i}, \bar{w}(z, \epsilon), \epsilon), \epsilon)} dF(\theta_{-i})} \\ &\quad \times \int \frac{f(\hat{\theta}(\theta_{-i}, \bar{w}(z, \epsilon), \epsilon) | \theta_{-i})}{w_{\theta_i}(\theta_{-i}, \hat{\theta}(\theta_{-i}, \bar{w}(z, \epsilon), \epsilon), \epsilon)} w_\epsilon(\theta_{-i}, \hat{\theta}(\theta_{-i}, \bar{w}(z, \epsilon), \epsilon), \epsilon) dF(\theta_{-i}).\end{aligned}$$

To see that this establishes the desired equality, define

$$G(W | \theta_{-i}) = F(\hat{\theta}(\theta_{-i}, W, \epsilon) | \theta_{-i})$$

the c.d.f. for $W = w(\theta, \epsilon)$ conditional on θ_{-i} ; differentiating, one sees that

$$g(W | \theta_{-i}) = \frac{f(\hat{\theta}(\theta_{-i}, W, \epsilon) | \theta_{-i})}{w_{\theta_i}(\theta_{-i}, \hat{\theta}(\theta_{-i}, W, \epsilon), \epsilon)}$$

represents the associated conditional density. Noting that

$$g(\theta_{-i} | W) d\theta_{-i} = \frac{g(W | \theta_{-i}) dF(\theta_{-i})}{\int g(W | \theta_{-i}) dF(\theta_{-i})}$$

the result then follows: $\bar{w}_\epsilon(z, \epsilon) = \mathbb{E}[w_\epsilon(\theta, \epsilon) | w(\theta, \epsilon) = \bar{w}(z, \epsilon)]$. \square

The following simple corollary will be employed below, for any differentiable function $N(w, \epsilon)$ define $n(\theta, \epsilon) = N(w(\theta, \epsilon), \epsilon)$ and $\bar{n}(z, \epsilon) = N(\bar{w}(z, \epsilon), \epsilon)$ then

$$\bar{n}_\epsilon(z, \epsilon) = \mathbb{E}[n_\epsilon(\theta, \epsilon) | w(\theta, \epsilon) = \bar{w}(z, \epsilon)].$$

Proof of Lemma 1. We engineer a variation that ensures all the equilibrium conditions are met except $G^*(y^*; \phi) = 0$ and evaluate how this variation affects welfare $W(\bar{U})$ and $G^*(y^*; \phi)$.

The new tax schedule is given by

$$\begin{aligned}t_i^*(\epsilon) &= t_i^* + \epsilon \hat{t}_i^* \text{ for all } i, \\ T(x, \epsilon) &= T(x) + \epsilon \hat{T}(x) \text{ for all } x \geq 0,\end{aligned}$$

for some arbitrary vector of new technology taxes, \hat{t}^* , income tax schedule, \hat{T} , and $\epsilon \in \mathbb{R}$. We let $\{c(\theta, \epsilon)\}$, $\{n(\theta, \epsilon)\}$, $y(\epsilon) \equiv \{y_i(\epsilon)\}$, and $y^*(\epsilon) \equiv \{y_i^*(\epsilon)\}$ denote the associated equilibrium allo-

cation and we let $p(\epsilon) \equiv \{p_i(\epsilon)\}$, $p^*(\epsilon) \equiv \{p_i^*(\epsilon)\}$ and $\{w(\theta, \epsilon)\}$ denote the associated prices and wages. They are given by conditions (C.1)-(C.3) and (C.5)-(C.9). The only equilibrium condition that is not imposed is $G^*(y^*(\epsilon); \phi) = 0$.

Define the Lagrangian

$$\mathcal{L} = \mathcal{W}(\bar{U}) - \gamma G^*(y^*; \phi)$$

with Lagrange multiplier $\gamma > 0$. A necessary condition for such a feasible variation not to improve welfare is that

$$\left. \frac{d\mathcal{W}(\{\bar{U}(z, \epsilon)\})}{d\epsilon} \right|_{\epsilon=0} - \gamma \left. \frac{dG^*(y^*(\epsilon); \phi)}{d\epsilon} \right|_{\epsilon=0} = 0. \quad (\text{D.2})$$

We now compute each of these derivatives. Let $R(x, \epsilon) \equiv x - T(x, \epsilon)$ denote the retention function associated with the income tax schedule $T(x, \epsilon)$. Recall that since households have identical preferences, all households of a given quantile z of the utility distribution have wage $\bar{w}(z, \epsilon)$ and achieve utility

$$\bar{U}(z, \epsilon) = \max_{\bar{n}(\theta)} u(C(p(\epsilon), R(\bar{w}(z, \epsilon)\bar{n}(\theta), \epsilon)), \bar{n}(\theta)),$$

with labor supply $\bar{n}(z, \epsilon)$, consumption, $\bar{c}(z, \epsilon)$, and earnings, $\bar{x}(z, \epsilon) \equiv \bar{w}(z, \epsilon)\bar{n}(z, \epsilon)$. Let $\bar{U}_\epsilon(z, \epsilon) \equiv \partial \bar{U}(z, \epsilon) / \partial \epsilon$ denote the marginal change in utility at quantile z . The Envelope Theorem implies

$$\bar{U}_\epsilon(z, \epsilon) = \bar{u}_C(z, \epsilon) \bar{C}_R(z, \epsilon) [\bar{R}_x(z, \epsilon)\bar{n}(z, \epsilon)\bar{w}_\epsilon(z, \epsilon) + \bar{R}_\epsilon(z, \epsilon) - \bar{c}(z, \epsilon) \cdot p_\epsilon],$$

with

$$\begin{aligned} \bar{u}_C(z, \epsilon) &\equiv u_C(C, n)|_{C=C(p(\epsilon), R(z, \epsilon)), n=\bar{n}(z, \epsilon)}, \\ \bar{C}_R(z, \epsilon) &\equiv C_R(p, R)|_{p=p(\epsilon), R=\bar{R}(z, \epsilon)}, \\ \bar{R}(z, \epsilon) &\equiv R(\bar{x}(z, \epsilon), \epsilon), \\ \bar{R}_x(z, \epsilon) &\equiv R_x(x, \epsilon)|_{x=\bar{x}(z, \epsilon), \epsilon}, \\ \bar{R}_\epsilon(z, \epsilon) &\equiv R_\epsilon(x, \epsilon)|_{x=\bar{x}(z, \epsilon), \epsilon}. \end{aligned}$$

This further implies

$$\begin{aligned} \frac{d\mathcal{W}(\{\bar{U}(z, \epsilon)\})}{d\epsilon} &= \\ &\int \frac{\partial \mathcal{W}}{\partial \bar{U}(z)} \bar{u}_C(z, \epsilon) \bar{C}_R(z, \epsilon) [\bar{R}_x(z, \epsilon)\bar{n}(z, \epsilon)\bar{w}_\epsilon(z, \epsilon) - \bar{c}(z, \epsilon) \cdot p_\epsilon + \bar{R}_\epsilon(z, \epsilon)] dz. \end{aligned} \quad (\text{D.3})$$

Now consider the change in the cost of resources used by the new technology,

$$\frac{dG^*(y^*(\epsilon); \phi)}{d\epsilon} = \sum_i G_{y_i^*}^* y_{i, \epsilon}^*.$$

Using equation (C.9) and good market clearing, this can be rearranged as

$$\frac{dG^*(y^*(\epsilon); \phi)}{d\epsilon} = G_{y_1^*}^* \{(p^* - p) \cdot y_\epsilon^* + p \cdot \int c_\epsilon(\theta, \epsilon) dF(\theta) - p \cdot y_\epsilon\}. \quad (\text{D.4})$$

From the first-order conditions associated with (C.3), we know that

$$\begin{aligned} p \cdot \frac{\partial y(p, n)}{\partial p_i} &= 0 \text{ for all } i, \\ p \cdot \frac{\partial y(p, n)}{\partial n(\theta)} &= w(\theta) \text{ for all } \theta. \end{aligned}$$

It follows that

$$p \cdot y_\epsilon = p \cdot \left[\sum_i \frac{\partial y(p, n)}{\partial p_i} p_{i,\epsilon} + \int \frac{\partial y(p, n)}{\partial n(\theta)} n_\epsilon(\theta, \epsilon) dF(\theta) \right] = \int w(\theta) n_\epsilon(\theta, \epsilon) dF(\theta), \quad (\text{D.5})$$

From the budget constraint associated with (C.1), we also know that

$$\frac{d}{d\epsilon} [p(\epsilon) \cdot \int c(\theta, \epsilon) dF(\theta)] = \int \frac{dR(w(\theta, \epsilon)n(\theta, \epsilon), \epsilon)}{d\epsilon} dF(\theta).$$

It follows that

$$\begin{aligned} p \cdot \int c_\epsilon(\theta, \epsilon) dF(\theta) &= - \int p_\epsilon \cdot c(\theta, \epsilon) dF(\theta) + \int [R_\epsilon(w(\theta, \epsilon)n(\theta, \epsilon), \epsilon) \\ &\quad + R_x(w(\theta, \epsilon)n(\theta, \epsilon), \epsilon)(w_\epsilon(\theta, \epsilon)n(\theta, \epsilon) + w(\theta, \epsilon)n_\epsilon(\theta, \epsilon))] dF(\theta). \end{aligned} \quad (\text{D.6})$$

Combining (D.4)-(D.6) and using $w(\theta, \epsilon)n(\theta, \epsilon) = x(\theta, \epsilon)$, we get

$$\begin{aligned} \frac{dG^*(y^*(\epsilon); \phi)}{d\epsilon} &= G_{y_1^*}^* \left\{ (p^* - p) \cdot y_\epsilon^* - \int p_\epsilon \cdot c(\theta, \epsilon) dF(\theta) \right. \\ &\quad \left. + \int [R_\epsilon(x(\theta, \epsilon), \epsilon) + R_x(x(\theta, \epsilon), \epsilon)w_\epsilon(\theta, \epsilon)n(\theta, \epsilon) - (1 - R_x(x(\theta, \epsilon), \epsilon))w(\theta, \epsilon)n_\epsilon(\theta, \epsilon))] dF(\theta) \right\}. \end{aligned}$$

Applying the corollary of Lemma 2, we have

$$\bar{n}_\epsilon(z, \epsilon) = \int n_\epsilon(\theta, \epsilon) f(\theta | w(\theta, \epsilon) = \bar{w}(z, \epsilon)) d\theta.$$

Thus we can rearrange the previous expression as

$$\begin{aligned} \frac{dG^*(y^*(\epsilon); \phi)}{d\epsilon} &= G_{y_1^*}^* \left\{ (p^* - p) \cdot y_\epsilon^* + \int [\bar{R}_\epsilon(z, \epsilon) + \bar{R}_x(z, \epsilon)\bar{n}(z, \epsilon)\bar{w}_\epsilon(z, \epsilon) - \bar{c}(z, \epsilon) \cdot p_\epsilon] dz \right. \\ &\quad \left. - \int (1 - \bar{R}_x(z, \epsilon))\bar{w}(z, \epsilon)\bar{n}_\epsilon(z, \epsilon) dz \right\}. \end{aligned} \quad (\text{D.7})$$

Substituting equations (D.3) and (D.7) into equation (D.2), we obtain

$$\begin{aligned} (p^* - p) \cdot y_\epsilon^* &- \int \tau(z)\bar{w}(z)\bar{n}_\epsilon(z) dz \\ &= \int [\bar{\lambda}(z) - 1][(1 - \tau(z))\bar{n}(z)\bar{w}_\epsilon(z) - \bar{T}_\epsilon(z) - \bar{c}(z) \cdot p_\epsilon] dz, \end{aligned}$$

with

$$\begin{aligned}\bar{\lambda}(z) &\equiv \frac{(\partial\mathcal{W}/\partial\bar{U}(z))|_{\epsilon=0} \bar{u}_C(z,0) \bar{C}_R(z,0)}{\gamma G_{y_1^*}^*(y^*(0))}, \\ \tau(z) &\equiv 1 - \bar{R}_x(z,0), \\ \bar{T}_\epsilon(z) &\equiv -\bar{R}_\epsilon(z,0),\end{aligned}$$

as well as the obvious short hand notation, $\bar{n}(z) \equiv \bar{n}(z,0)$ etc.

D.2 Proof of Proposition 1

Consider a first variation such that

$$\begin{aligned}t_i^*(\epsilon) &= t_i^* + \epsilon \hat{t}_i^* \text{ for all } i, \\ T(x, \epsilon) &= T(x) - \epsilon \tau \text{ for all } x \geq 0,\end{aligned}$$

In the proof of Lemma 1, we have established that

$$\begin{aligned}\frac{dG^*(y^*(\epsilon); \phi)}{d\epsilon} &= G_{y_1^*}^* \{ (p^* - p) \cdot y_\epsilon^* + \int [\bar{R}_\epsilon(z, \epsilon) + \bar{R}_x(z, \epsilon) \bar{n}(z, \epsilon) \bar{w}_\epsilon(z, \epsilon) - \bar{c}(z, \epsilon) \cdot p_\epsilon] dz \\ &\quad - \int (1 - \bar{R}_x(z, \epsilon)) \bar{w}(z, \epsilon) \bar{n}_\epsilon(z, \epsilon) dz \}.\end{aligned}$$

For the uniform lump-sum transfer, $\tau = \bar{R}_\epsilon(z, \epsilon)$ for all z , to maintain the government's budget balance, it must therefore satisfy

$$\tau = -\{ (p^* - p) \cdot y_\epsilon^* + \int [\bar{R}_x(z, \epsilon) \bar{n}(z, \epsilon) \bar{w}_\epsilon(z, \epsilon) - \bar{c}(z, \epsilon) \cdot p_\epsilon] dz - \int (1 - \bar{R}_x(z, \epsilon)) \bar{w}(z, \epsilon) \bar{n}_\epsilon(z, \epsilon) dz \}.$$

Substituting this expression for $-\delta T(z)$ in equation (1), we obtain

$$\begin{aligned}[\int \{ (p^* - p) \cdot y_\epsilon^* - \int (1 - \bar{R}_x(z, \epsilon)) \bar{w}(z, \epsilon) \bar{n}_\epsilon(z, \epsilon) dz \}] \\ = \int \left[\frac{\bar{\lambda}(z)}{\int \bar{\lambda}(v) dv} - 1 \right] [(1 - \tau(z)) \bar{n}(z) \bar{w}_\epsilon(z, \epsilon) - \bar{c}(z) \cdot p_\epsilon] dz,\end{aligned}$$

where we again use the short hand notation, $\bar{n}(z) \equiv \bar{n}(z,0)$ etc. Now pick $\hat{t}_i^* \equiv \{\hat{t}_i^*\}$ such that $y_{1,\epsilon}^*, y_{i,\epsilon}^* \neq 0$ and $y_{j,\epsilon}^* = 0$ for all $j \neq 1, i$. The previous expression implies

$$\begin{aligned}p_i - p_i^* \\ = \int \left[\left(1 - \frac{\bar{\lambda}(z)}{\int \bar{\lambda}(v) dv} \right) \left((1 - \tau(z)) \bar{n}(z) \frac{\delta \bar{w}(z)}{\delta y_i^*} \Big|_{\delta T=0} - \bar{c}(z) \frac{\delta p}{\delta y_i^*} \Big|_{\delta T=0} \right) - \tau(z) \bar{w}(z) \frac{\delta \bar{n}(z)}{\delta y_i^*} \Big|_{\delta T=0} \right] dz.\end{aligned}$$

D.3 Proof of Proposition 2

Consider a variation

$$\begin{aligned} t_i^*(\epsilon) &= t_i^* + \epsilon \hat{t}_i^* \text{ for all } i, \\ T(x, \epsilon) &= T(x) + \epsilon \hat{T}(x) \text{ for all } x \geq 0, \end{aligned}$$

such that $\bar{U}_\epsilon(z, \epsilon) = 0$ for all z . From Lemma 1, we know that

$$\begin{aligned} (p^* - p) \cdot y_\epsilon^* - \int \tau(z) \bar{w}(z) \bar{n}_\epsilon(z) dz \\ = \int (\bar{\lambda}(z) - 1) [(1 - \tau(z)) \bar{n}(z) \bar{w}_\epsilon(z) - \bar{c}(z) \cdot p_\epsilon - T_\epsilon(\bar{x}(z))] dz, \end{aligned}$$

with the short hand notation, $\bar{n}(z) \equiv \bar{n}(z, 0)$ etc. In the proof of Lemma 1, we have already established that

$$\bar{U}_\epsilon(z) = \bar{u}_C(z) \bar{C}_R(z) [(1 - \tau(z)) \bar{n}(z) \bar{w}_\epsilon(z) - \bar{c}(z) \cdot p_\epsilon - T_\epsilon(\bar{x}(z))].$$

It follows that if $\bar{U}_\epsilon(z) = 0$ for all z , then

$$(p^* - p) \cdot y_\epsilon^* = \int \tau(z) \bar{w}(z) \bar{n}_\epsilon(z) dz. \quad (\text{D.8})$$

Let us now compute $\bar{n}_\epsilon(z)$. We use the following definitions,

$$\begin{aligned} \bar{C}(z, \epsilon) &\equiv C(p(\epsilon), R(\bar{w}(z, \epsilon) \bar{n}(z, \epsilon), \epsilon)), \\ MRS(C, n) &\equiv -\frac{u_n(C, n)}{u_C(C, n)}. \end{aligned}$$

Note that for any household at the quantile z of the utility distribution, we must have

$$\bar{U}(z, \epsilon) = \max_{z'} u(\bar{C}(z', \epsilon), \bar{n}(z', \epsilon) \frac{\bar{w}(z', \epsilon)}{\bar{w}(z, \epsilon)}).$$

By the Envelope Theorem, this further implies

$$\bar{U}_z(z, \epsilon) = -\bar{u}_n(z, \epsilon) \omega(z, \epsilon) \bar{n}(z, \epsilon), \quad (\text{D.9})$$

with

$$\begin{aligned} \bar{u}_n(z, \epsilon) &\equiv u_n(\bar{C}(z, \epsilon), \bar{n}(z, \epsilon)), \\ \omega(z, \epsilon) &\equiv \bar{w}_z(z, \epsilon) / \bar{w}(z, \epsilon). \end{aligned}$$

Since $\bar{U}_\epsilon(z) = 0$ for all z , we must also have $\bar{U}_{z\epsilon}(z) = 0$ for all z . Taking log and differentiating (D.9) with respect to ϵ we therefore obtain

$$\frac{\bar{u}_{n\epsilon}(z)}{\bar{u}_n(z)} + \frac{\omega_\epsilon(z)}{\omega(z)} + \frac{\bar{n}_\epsilon(z)}{\bar{n}(z)} = 0. \quad (\text{D.10})$$

Next let us show that if $\bar{U}_\epsilon(z) = 0$ for all z , then

$$\frac{\bar{u}_{n\epsilon}(z)}{\bar{u}_n(z)} = \frac{1}{\epsilon_C(z)} \frac{\bar{n}_\epsilon(z)}{\bar{n}(z)}, \quad (\text{D.11})$$

where $\epsilon_C(z) \equiv \frac{MRS(\bar{C}(z), \bar{n}(z))}{MRS_n(\bar{C}(z), \bar{n}(z))\bar{n}(z)}$ denotes the consumption-compensated elasticity of labor supply. Differentiating $\bar{u}_n(z, \epsilon)$ with respect to ϵ , we get

$$\frac{\bar{u}_{n\epsilon}(z)}{\bar{u}_n(z)} = \frac{u_{nC}}{u_n} \bar{C}_\epsilon(z, \epsilon) + \frac{u_{nn}}{u_n} \bar{n}_\epsilon(z).$$

Using the fact that $\bar{U}_\epsilon(z) = u_C \bar{C}_\epsilon(z, \epsilon) + u_n \bar{n}_\epsilon(z) = 0$, we can rearrange the previous expression as

$$\frac{\bar{u}_{n\epsilon}(z)}{\bar{u}_n(z)} = \left[-\frac{u_{nC}}{u_C} + \frac{u_{nn}}{u_n} \right] \bar{n}_\epsilon(z).$$

Noting that $\frac{\partial \ln MRS}{\partial \ln n} = \frac{u_{nn}}{u_n} - \frac{u_{nC}}{u_C}$, we obtain (D.11). Combining (D.10) and (D.11), we then get

$$\frac{\bar{n}_\epsilon(z)}{\bar{n}_\epsilon} = -\frac{\epsilon_C(z)}{1 + \epsilon_C(z)} \frac{\omega_\epsilon(z)}{\omega(z)}.$$

This can be rearranged equivalently as

$$\frac{\bar{n}_\epsilon(z)}{\bar{n}_\epsilon} = -\frac{\epsilon_H(z)}{1 + \epsilon_M(z)} \frac{\omega_\epsilon(z)}{\omega(z)},$$

where $\epsilon_H(z) \equiv \frac{1}{\bar{n}(z)MRS_C + \bar{n}(z)MRS_n/MRS}$ and $\epsilon_M(z) \equiv \frac{1 - \bar{n}(z)MRS_C}{\bar{n}(z)MRS_C + \bar{n}(z)MRS_n/MRS}$ denote the Hicksian and Marshallian labor supply elasticities, respectively.

Substituting for $\bar{n}_\epsilon(z)$ into (D.8), we obtain

$$(p^* - p) \cdot y_\epsilon^* = - \int \tau(z) \bar{w}(z) \bar{n}(z) \frac{\epsilon_H(z)}{1 + \epsilon_M(z)} \frac{\omega_\epsilon(z)}{\omega(z)} dz.$$

Pick \hat{T} and $\hat{t}^* \equiv \{\hat{t}_i^*\}$ such that, in addition to $\bar{U}_\epsilon = 0$, $y_{1,\epsilon}^*, y_{i,\epsilon}^* \neq 0$ and $y_{j,\epsilon}^* = 0$ for all $j \neq 1, i$. Then

$$p_i - p_i^* = \int \tau(z) \bar{w}(z) \bar{n}(z) \frac{\epsilon_H(z)}{\epsilon_M(z) + 1} \frac{1}{\omega(z)} \frac{\delta \omega(z)}{\delta y_i^*} \Big|_{\delta \bar{U}=0} dz.$$

E Section 5

E.1 Proof of Proposition 3

Denote $(p_\epsilon)_{\delta \bar{U}=0}$, $(\{\omega_\epsilon(\theta)\})_{\delta \bar{U}=0}$, and $(y_\epsilon^*)_{\delta \bar{U}=0}$ the change in prices, wages and output by new technology firms associated with the variation $\delta \bar{U} = 0$. Now consider an alternative variation

$$\begin{aligned} t_i^*(\epsilon) &= t_i^* + \epsilon \hat{t}_i^* \text{ for all } i, \\ T(x, \epsilon) &= T(x) + \epsilon \hat{T}(x) \text{ for all } x \geq 0, \end{aligned}$$

with \hat{t} set such that $p_\epsilon = (p_\epsilon)_{\delta\bar{U}=0}$ and \hat{T} set such that $G_\epsilon^*(y^*) = 0$. We do not impose any other restriction on \hat{T} . We refer to such a variation as $\delta G^* = 0$.

Our goal is to show that if $|\varepsilon_H(z)|, |\varepsilon_M(z)| < \bar{\varepsilon}$, then

$$\begin{aligned}(\omega_\epsilon(z))_{\delta G^*=0} &= (\omega_\epsilon(z))_{\delta\bar{U}=0} + O(\bar{\varepsilon}); \\(y_\epsilon^*)_{\delta G^*=0} &= (y_\epsilon^*)_{\delta\bar{U}=0} + O(\bar{\varepsilon}).\end{aligned}$$

The two previous equations imply $\frac{\delta \ln \omega(z)}{\delta \ln y_i^*}|_{\delta G^*=0} = \frac{\delta \ln \omega(z)}{\delta \ln y_i^*}|_{\delta\bar{U}=0} + O(\bar{\varepsilon})$. Proposition 3 then follows from this observation and the fact that $\varepsilon_H(z) = O(\bar{\varepsilon})$ and $\varepsilon_M(z) = O(\bar{\varepsilon})$. Our proof proceeds in three steps.

Step 1: If $|\varepsilon_H(z)|, |\varepsilon_M(z)| < \bar{\varepsilon}$, then for any variation, $\bar{n}_\epsilon(z) = O(\bar{\varepsilon})$ for all z .

For any z , the household labor supply solves

$$MRS(\tilde{C}(\bar{w}(z, \epsilon)\bar{n}(z, \epsilon), \epsilon), \bar{n}(z, \epsilon)) = \bar{w}(z, \epsilon)\tilde{C}_x(\bar{w}(z, \epsilon)\bar{n}(z, \epsilon), \epsilon), \quad (\text{E.1})$$

with $\tilde{C}(x, \epsilon) \equiv C(p(\epsilon), R(x, \epsilon))$. Differentiating the previous expression with respect to ϵ we get

$$\bar{n}_\epsilon(z) = \frac{(\tilde{C}_{xx}\bar{w}(z)\bar{n}(z) + \tilde{C}_x)\bar{w}_\epsilon(z) + \tilde{C}_{x\epsilon}\bar{w}(z) - MRS_C(\tilde{C}_x\bar{n}(z)\bar{w}_\epsilon(z) + \tilde{C}_\epsilon)}{MRS_C\tilde{C}_x\bar{w}(z) + MRS_n - \tilde{C}_{xx}\bar{w}^2(z)}.$$

Using $\varepsilon_H(z) \equiv \frac{1}{\bar{n}(z)MRS_C + \bar{n}(z)MRS_n/MRS}$, $\varepsilon_M(z) \equiv \frac{1 - \bar{n}(z)MRS_C}{\bar{n}(z)MRS_C + \bar{n}(z)MRS_n/MRS}$, and $MRS = \bar{w}(z)\tilde{C}_x$, we can rearrange the previous expression as

$$\frac{\bar{n}_\epsilon(z)}{\bar{n}} = \frac{\varepsilon_H(z)\left(\frac{\tilde{C}_{xx}\bar{w}(z)\bar{n}(z)}{\tilde{C}_x} \frac{\bar{w}_\epsilon(z)}{\bar{w}(z)} + \frac{\tilde{C}_{x\epsilon}}{\tilde{C}_x} - \frac{\tilde{C}_\epsilon}{\tilde{C}_x\bar{w}(z)\bar{n}(z)}\right) + \varepsilon_M(z)\left(\frac{\bar{w}_\epsilon(z)}{\bar{w}(z)} + \frac{\tilde{C}_\epsilon}{\tilde{C}_x\bar{w}(z)\bar{n}(z)}\right)}{1 - \varepsilon_H(z)\frac{\tilde{C}_{xx}\bar{w}(z)\bar{n}(z)}{\tilde{C}_x}}. \quad (\text{E.2})$$

By assumption, we know that $\varepsilon_H(z) = O(\bar{\varepsilon})$ and $\varepsilon_M(z) = O(\bar{\varepsilon})$. Combining this observation with equation (E.2), we obtain $\bar{n}_\epsilon(z) = O(\bar{\varepsilon})$ for all z .

Step 2: If $|\varepsilon_H(z)|, |\varepsilon_M(z)| < \bar{\varepsilon}$, then $(\omega_\epsilon(z))_{\delta G^*=0} = (\omega_\epsilon(z))_{\delta\bar{U}=0} + O(\bar{\varepsilon})$ for all z .

By definition, we know that $w(\theta, \epsilon) = w(p(\epsilon), \{n(\theta, \epsilon)\})$. Differentiating the previous expression, we get

$$w_\epsilon(\theta) = w_p(\theta) \cdot p_\epsilon + \int w_{n(\theta')}(\theta)n_\epsilon(\theta')dF(\theta') \text{ for all } \theta.$$

Given Step 1, we must therefore have $w_\epsilon(\theta) = w_p(\theta) \cdot p_\epsilon + O(\bar{\varepsilon})$. Since $(p_\epsilon)_{\delta G^*=0} = (p_\epsilon)_{\delta\bar{U}=0}$, this further implies that $(w_\epsilon(\theta))_{\delta G^*=0} = (w(\theta))_{\delta\bar{U}=0} + O(\bar{\varepsilon})$ for all θ . Step 2 follows from this observation.

Step 3: If $|\varepsilon_H(z)|, |\varepsilon_M(z)| < \bar{\varepsilon}$, then $(y_\epsilon^*)_{\delta G^*=0} = (y_\epsilon^*)_{\delta\bar{U}=0} + O(\bar{\varepsilon})$.

Since preferences are quasi-homothetic, there exist $a(p)$ and $b(p)$ such that

$$\bar{c}(z, \epsilon) = a(p(\epsilon)) + b(p(\epsilon))\bar{C}(z, \epsilon) \text{ for all } z.$$

We can therefore express output by new technology firms as

$$y^*(\epsilon) = a(p(\epsilon)) + b(p(\epsilon)) \int \bar{C}(z, \epsilon) dz - y(p(\epsilon), \{n(\theta, \epsilon)\}).$$

Differentiating implies that for any variation,

$$\begin{aligned} y_\epsilon^* &= \\ & [a_p(p) + b_p(p) \int \bar{C}(z, \epsilon) dz - y_p] p_\epsilon - \int y_{n(\theta')} n_\epsilon(\theta') dF(\theta') + b(p) \int \bar{C}_\epsilon(z, \epsilon) dz \\ &= [a_p(p) + b_p(p) \int \bar{C}(z, \epsilon) dz - y_p] p_\epsilon + b(p) \int \bar{C}_\epsilon(z, \epsilon) dz + O(\bar{\epsilon}), \end{aligned} \quad (\text{E.3})$$

where the second equality derives from Step 1.

Next we show that $(\int \bar{C}_\epsilon(z, \epsilon) dz)_{\delta G^*=0} = O(\bar{\epsilon})$. We use the same type of arguments as in the proof of Lemma 1. Start from

$$G_\epsilon^* = (p^* - p) \cdot y_\epsilon^* + p \cdot \int \bar{c}_\epsilon(z) dz - p \cdot y_\epsilon = 0. \quad (\text{E.4})$$

Since firms maximize profits, we know that $p \cdot [y_p(p, \{n(\theta)\})] p_\epsilon = 0$ and $p \cdot [y_{n(\theta')}(p, \{n(\theta)\})] = w(\theta')$, which implies

$$p \cdot y_\epsilon = \int w(\theta) n_\epsilon(\theta) dF(\theta) = O(\bar{\epsilon}), \quad (\text{E.5})$$

where the second equality again follows from Step 1. Likewise, since consumers minimize expenditure, we know that $p \cdot [a_p(p) + b_p(p) \bar{C}(z, \epsilon)] p_\epsilon = 0$, which implies

$$p \cdot \int c_\epsilon(\theta, \epsilon) dF(\theta) = [p \cdot b(p)] [\int \bar{C}_\epsilon(z, \epsilon) dz]. \quad (\text{E.6})$$

From Proposition 2, we also know that

$$(p^* - p) \cdot y_\epsilon^* = \sum_i \left[\int \tau(z) \bar{w}(z) \bar{n}(z) \frac{\epsilon_H(z)}{\epsilon_M(z) + 1} \frac{\delta \ln \omega(z)}{\delta \ln y_i^*} \Big|_{\delta \bar{U}=0} dz \right] y_{i,\epsilon}^* = O(\bar{\epsilon}), \quad (\text{E.7})$$

where the second equality follows from $\epsilon_H(z) = O(\bar{\epsilon})$ and $\epsilon_M(z) = O(\bar{\epsilon})$. Combining equations (E.4)-(E.7), we get

$$\left(\int \bar{C}_\epsilon(z, \epsilon) dz \right)_{\delta G^*=0} = O(\bar{\epsilon}). \quad (\text{E.8})$$

Finally, we show that $(\int \bar{C}_\epsilon(z, \epsilon) dz)_{\delta \bar{U}=0} = O(\bar{\epsilon})$. Differentiating $\bar{U}(z, \epsilon) = u(\bar{C}(z, \epsilon), \bar{n}(z, \epsilon))$ and imposing $\bar{U}_\epsilon(z, \epsilon) = 0$ implies

$$(\bar{C}_\epsilon(z, \epsilon))_{\delta \bar{U}=0} = -(u_n / u_C) \bar{n}_\epsilon = O(\bar{\epsilon}) \text{ for all } z,$$

where the second equality derives from Step 1. The previous expression, in turn, implies

$$\left(\int \bar{C}_\epsilon(z, \epsilon) dz \right)_{\delta \bar{U}=0} = O(\bar{\epsilon}). \quad (\text{E.9})$$

Combining equations (E.3), (E.8), and (E.9) with the fact that $(p_\epsilon)_{\delta G^*=0} = (p_\epsilon)_{\delta \bar{U}=0}$, we conclude

that $(y_\epsilon^*)_{\delta G^*=0} = (y_\epsilon^*)_{\delta \bar{U}=0} + O(\bar{\epsilon})$.

E.2 Quantitative Implementation of Optimal Tax on Robots

To implement the formula in Proposition 3 in the case of robots, we start from the definition $\omega(z) = d \ln \bar{w}(z) / dz$ to express $\frac{\delta \ln \omega(z)}{\delta \ln y_m^*} \Big|_{\delta G^*=0}$ as a function of $\frac{\delta \ln \bar{w}(z)}{\delta y_m^*}$,

$$\frac{\delta \ln \omega(z)}{\delta \ln y_m^*} \Big|_{\delta G^*=0} = \frac{y_m^*}{\omega(z)} \frac{\delta \omega(z)}{\delta y_m^*} \Big|_{\delta G^*=0} = \frac{y_m^*}{\omega(z)} \frac{d}{dz} \left(\frac{\delta \ln \bar{w}(z)}{\delta y_m^*} \right) \Big|_{\delta G^*=0}.$$

Substituting into equation (8), we then obtain

$$t_m^* \simeq \int \tau(z) \frac{\bar{x}(z)}{p_m^* y_m^*} \frac{\varepsilon_H(z)}{\varepsilon_M(z) + 1} \frac{y_m^*}{\omega(z)} \frac{d}{dz} \left(\frac{\delta \ln \bar{w}(z)}{\delta y_m^*} \right) \Big|_{\delta G^*=0} dz.$$

We interpret [Acemoglu and Restrepo's \(2017b\)](#) estimates $\eta_{AR}(q)$ of the semi-elasticity of wages with respect to the number of robots across different deciles $q \in \mathcal{D} \equiv \{10, 20, \dots, 90\}$ (from their long-differences specification, in their Figure 13) as the empirical counterpart of $\left(\frac{\delta \ln \bar{w}(q)}{\delta y_m^*} \right) \Big|_{\delta G^*=0}$ for households at a decile q of the US income distribution.²² Using a discrete approximation, we can therefore express the optimal tax on robots as

$$t_m^* \simeq \sum_{q \in \mathcal{D} / \{10\}} \tau(q) \frac{\bar{x}(q)}{p_m^* y_m^*} \frac{\varepsilon_H(q)}{\varepsilon_M(q) + 1} \frac{y_m^* \Delta \eta_{AR}(q)}{\Delta \ln \bar{w}(q)},$$

where $\tau(q)$ denotes the marginal tax rate faced by households between the $(q - 10)$ -th and q -th deciles of the income distribution; $\bar{x}(q)$ denotes the labor earnings of individuals between the same two deciles; $\varepsilon_H(q)$ and $\varepsilon_M(q)$ denote their Hicksian and Marshallian labor supply elasticities of households at the q -th decile of the income distribution; and Δ denotes differences between consecutive quantiles, i.e. $\Delta \eta_{AR}(q) = \eta_{AR}(q) - \eta_{AR}(q - 10)$ and $\Delta \ln \bar{w}(q) = \ln \bar{w}(q) - \ln \bar{w}(q - 10)$. Note that since [Acemoglu and Restrepo's \(2017b\)](#) estimates are not available for $q = 0$ and 100, we omit the bottom and top deciles from our approximation; this is equivalent to setting $\Delta \eta_{AR}(10) = \Delta \eta_{AR}(100) = 0$.

Ignoring differences in labor supply elasticities across quantiles and approximating changes in log-wages by changes in log-earnings, we can finally compute the optimal tax on robots as

$$t_m^* \simeq \frac{\varepsilon_H}{\varepsilon_M + 1} \frac{1 - s_m}{s_m} \sum_{q \in \mathcal{D} / \{10\}} \tau(q) \bar{s}(q) \frac{\Delta \ln \omega(q)}{\Delta \ln y_m^*}, \quad (\text{E.10})$$

²²Figure 13 in [Acemoglu and Restrepo \(2017b\)](#) reports the reduced-form coefficient obtained from regressing changes in wages on the exogenous exposure to robots (measured as adoption of industrial robots in nine other European economies from 1993 to 2007). To go from the semi-elasticity with respect to exogenous exposure to robots to the semi-elasticity with respect to US exposure, we divide these coefficients by the first-stage coefficient, equal to 2, obtained from regressing US exposure on exogenous exposure.

Variables	Values	Source
Panel A: Baseline		
$\varepsilon_H, \varepsilon_M$	0.5	Chetty (2012)
s_m	0.4%	Graetz and Michaels (2018)
$\{\tau(q)\}$	$\{0, 0.07, 0.07, 0.07, 0.14, 0.14, 0.14, 0.21, 0.27\}$	Guner et al. (2014)
$\{\bar{s}(q)\}$	$\{0, 0.02, 0.03, 0.04, 0.06, 0.08, 0.1, 0.12, 0.16\}$	World Wealth and Income Database
$\{\frac{\Delta \ln \omega(q)}{\Delta \ln y_m^*} \times 100\}$	$\{0.16, 0.82, 0.22, 0.20, 0.21, 0.73, 1.30, 0.06\}$	Acemoglu and Restrepo (2017b)
Panel B: Inelastic Marshallian labor supply		
ε_M	0	n.a.
Panel C: Effective US marginal tax rates		
$\{\tau(q)\}$	$\{0, 0.07, 0.07, 0.08, 0.12, 0.12, 0.15, 0.17, 0.21\}$	Guner et al. (2014)
Panel D: Fitted wage semi-elasticity		
$\{\frac{\Delta \ln \omega(q)}{\Delta \ln y_m^*} \times 100\}$	$\{0.26, 0.43, 0.40, 0.38, 0.40, 0.49, 0.64, 0.75\}$	Acemoglu and Restrepo (2017b)
Panel E: Fitted relative wage elasticity		
$\{\frac{\Delta \ln \omega(q)}{\Delta \ln y_m^*} \times 100\}$	$\{0.45, 0.23, 0.25, 0.40, 0.59, 0.72, 0.68, 0.37\}$	Acemoglu and Restrepo (2017b)

Table E.1: Quantitative Implementation of Optimal Tax on Robots

Notes: Panels A-E report the values used to compute the optimal tax on robots in Columns (1)-(5), respectively, of Table 1. All variables are set to their baseline values unless stated otherwise.

with s_m , $\bar{s}(q)$, and $\Delta \ln \omega(q) / \Delta \ln y_m^*$ defined as

$$s_m \equiv \frac{p_m^* y_m^*}{p_m^* y_m^* + \bar{x}}$$

$$\bar{s}(q) \equiv \frac{\bar{x}(q)}{\bar{x}}$$

$$\frac{\Delta \ln \omega(q)}{\Delta \ln y_m^*} \equiv \frac{y_m^* \Delta \eta_{AR}(q)}{\Delta \ln \bar{s}(q)},$$

where $\bar{x} \equiv \sum_{q \in \mathcal{D} \cup \{100\}} \bar{x}(q)$ denotes total labor earnings. Table E.1 reports the value of all the variables entering equation (E.10). As mentioned in Section 5.2, our baseline computation (Panel A) uses: Hicksian labor supply elasticities from Chetty (2012); spending on robots from Graetz and Michaels (2018); statutory marginal tax rates from Guner et al. (2014); and shares of labor earnings by deciles in the United States from the World Wealth and Income Database. The elasticity of relative wages $\Delta \ln \omega(q) / \Delta \ln y_m^*$ for each decile is computed using $\frac{\Delta \ln \omega(q)}{\Delta \ln y_m^*} \equiv \frac{y_m^* \Delta \eta_{AR}(q)}{\Delta \ln \bar{s}(q)}$, with $y_m^* \simeq 1.2$ and $\Delta \eta_{AR}(q)$ from Acemoglu and Restrepo (2017b) and $\Delta \ln \bar{s}(q)$ from the World Wealth and Income Database.

Our sensitivity exercises first set the Marshallian labor supply elasticity to zero (in Panel B) and

then set the marginal income tax rates to their effective values, also from [Guner et al. \(2014\)](#) (in Panel C). For the last two panels (Panels D and E), we use fitted values of the wage semi-elasticity and relative wage elasticity, respectively. Specifically, for Panel D, we estimate via OLS,

$$\eta_{AR}(q) = \sum_{k=0}^3 \beta_{\text{Panel D},k} q^k + \text{error}_{\text{Panel D}}(q),$$

and use $\hat{\eta}_{AR}(q) = \sum_{k=0}^3 \hat{\beta}_{\text{Panel D},k} q^k$ to compute $(\frac{\Delta \ln \omega(q)}{\Delta \ln y_m^*})_{\text{Panel D}} = \frac{y_m^* \Delta \hat{\eta}_{AR}(q)}{\Delta \ln \bar{s}(q)}$. For Panel E, in turn, we estimate via OLS,

$$\frac{\Delta \ln \omega(q)}{\Delta \ln y_m^*} = \sum_{k=0}^3 \beta_{\text{Panel E},k} q^k + \text{error}_{\text{Panel E}}(q)$$

and directly use $(\frac{\Delta \ln \omega(q)}{\Delta \ln y_m^*})_{\text{Panel E}} = \sum_{k=0}^3 \hat{\beta}_{\text{Panel E},k} q^k$ to compute the optimal tax on robots.

E.3 Quantitative Implementation of Optimal Tax on Chinese Imports

To implement the formula in Proposition 3 in the case of Chinese imports, we follow the exact same approach as in Section (5). The only difference is that we now use estimates of $\eta_{CLP}(q)$ from [Chetverikov et al. \(2016\)](#) as the empirical counterpart of $(\frac{\delta \ln \bar{w}(q)}{\delta y_m^*})|_{\delta G^*=0}$ for households at the q -th quantile of the income distribution. These estimates are available across quintiles $q \in \mathcal{Q} \equiv \{5, 10, \dots, 95\}$. Following the same steps as before, we obtain the following counterpart of equation (E.10),

$$t_m^* \simeq \frac{\varepsilon_H}{\varepsilon_M + 1} \frac{1 - s_m}{s_m} \sum_{q \in \mathcal{Q}/\{5\}} \tau(q) \bar{s}(q) \frac{\Delta \ln \omega(q)}{\Delta \ln y_m^*}. \quad (\text{E.11})$$

with the share $s_m \equiv p_m^* y_m^* / (p_m^* y_m^* + \bar{x})$ now a function of Chinese imports, $p_m^* y_m^*$, and the elasticity $\Delta \ln \omega(q) / \Delta \ln y_m^*$ now defined as $y_m^* \Delta \eta_{CLP}(q) / \Delta \ln \bar{s}(q)$, with $y_m^* \simeq 2.2$, which corresponds to the ratio of the value of US imports from China (\$330 billion) to the number of US workers (153 million) in 2007, expressed in thousands of US dollars per worker. The values of all variables entering equation (E.11) are reported in Table E.2.

Variables	Values	Source
Panel A: Baseline		
$\varepsilon_H, \varepsilon_M$	0.5	Chetty (2012)
s_m	3.7%	Autor et al. (2013)
$\{\tau(q)\}$	{0, 0, 0.07, 0.07, 0.07, 0.07, 0.07, 0.07, 0.07, 0.14, 0.14, 0.14, 0.14, 0.14, 0.21, 0.21, 0.21, 0.27, 0.28}	Guner et al. (2014)
$\{\bar{s}(q)\}$	{0, 0, 0.01, 0.01, 0.01, 0.02, 0.02, 0.02, 0.03, 0.03, 0.04, 0.04, 0.05, 0.05, 0.06, 0.07, 0.08, 0.09, 0.11}	World Wealth and Income Database
$\left\{\frac{\Delta \ln \omega(q)}{\Delta \ln y_m^*} \times 100\right\}$	{-0.27, 0.04, 1.56, 1.57, 1.39, -0.02, 4.93, -0.55, 1.12, -0.68, -1.62, 3.46, 0.49, -2.94, -2.05, 1.13, 6.33, -4.12}	Chetverikov et al. (2016)
Panel B: Inelastic Marshallian labor supply		
ε_M	0	n.a.
Panel C: Effective US marginal tax rates		
$\{\tau(q)\}$	{0, 0, 0.07, 0.07, 0.07, 0.07, 0.08, 0.08, 0.08, 0.12, 0.12, 0.12, 0.15, 0.15, 0.18, 0.18, 0.20, 0.21, 0.22}	Guner et al. (2014)
Panel D: Fitted wage semi-elasticity		
$\left\{\frac{\Delta \ln \omega(q)}{\Delta \ln y_m^*} \times 100\right\}$	{0.16, 0.60, 0.88, 0.94, 0.97, 1.07, 1.02, 1.08, 0.97, 0.93, 0.78, 0.64, 0.44, 0.19, -0.06, -0.30, -0.45, -0.45}	Chetverikov et al. (2016)
Panel E: Fitted relative wage elasticity		
$\left\{\frac{\Delta \ln \omega(q)}{\Delta \ln y_m^*} \times 100\right\}$	{-0.03, 0.55, 0.96, 1.22, 1.34, 1.35, 1.27, 1.12, 0.92, 0.69, 0.45, 0.22, 0.21, -0.12, -0.20, -0.17, -0.03, 0.24}	Chetverikov et al. (2016)

Table E.2: Quantitative Implementation of Optimal Tax on Chinese Imports

Notes: Panels A-E report values used to compute the optimal tax on Chinese imports in Columns (1)-(5), respectively, of Table 1. All variables are set to their baseline values unless stated otherwise.

F Section 6

F.1 Preliminaries

Government Problem. Suppose, as in Section 6, that skill heterogeneity is one-dimensional, $\theta \in [0, 1]$, so that we can write social welfare as a function of the utility of each skill type, $\mathcal{W}(\{U(\theta)\})$. Using the revelation principle, we can express the government problem as

$$\max_{U, n, p} \mathcal{W}(\{U(\theta)\})$$

subject to

$$\theta \in \operatorname{argmax}_{\tilde{\theta}} u \left(C(n(\tilde{\theta}), U(\tilde{\theta})), n(\tilde{\theta}) \frac{w(p, n; \tilde{\theta})}{w(p, n; \theta)} \right),$$

$$G^*(c(p, n, U) - y(p, n); \phi) \leq 0,$$

where $C(n(\theta), U(\theta))$ is the aggregate consumption required to achieve utility $U(\theta)$ given labor supply $n(\theta)$, that is the solution to $u(C, n(\theta)) = U(\theta)$, and $c(p, n, U) = \int c(p, C(n(\theta), U(\theta))) dF(\theta)$ is the total demand for goods conditional on prices, p , labor supply, $n \equiv \{n(\theta)\}$, and utility levels, $U \equiv \{U(\theta)\}$, with $c(p, C)$ the solution to $\min_{\tilde{c}(\theta)} \{p \cdot \tilde{c} | v(\tilde{c}) \geq C\}$.

The envelope condition associated with the Incentive Compatibility constraint gives

$$U'(\theta) = -u_n(C(n(\theta), U(\theta)), n(\theta)) n(\theta) \omega(p, n; \theta)$$

with $\omega(p, n; \theta) \equiv \frac{w_\theta(p, n; \theta)}{w(p, n; \theta)}$. For piecewise differentiable allocations, the envelope condition and monotonicity of the mapping from wages, $w(p, n; \theta)$, to before-tax earnings, $w(p, n; \theta)n(\theta)$ is equivalent to incentive compatibility. We will focus on cases where $w(p, n; \theta)$ is increasing in θ , which for a given allocation can be interpreted as a normalization or ordering of θ . Under the previous conditions, we can rearrange our planning problem as

$$\max_{U, n, p} \mathcal{W}(\{U(\theta)\})$$

subject to

$$U'(\theta) = -u_n(C(n(\theta), U(\theta)), n(\theta)) n(\theta) \omega(p, n; \theta),$$

$$G^*(c(p, n, U) - y(p, n); \phi) \leq 0.$$

Under the functional-form assumptions of Section 6 this simplifies further into

$$\max_{U, n, p_m} \int U(\theta) d\Lambda(\theta) \tag{F.1a}$$

subject to

$$U'(\theta) = h'(n(\theta)) n(\theta) \omega(p_m; \theta), \tag{F.1b}$$

$$\phi \left[\int (U(\theta) + h(n(\theta))) dF(\theta) - y_f(p_m, n) \right] - y_m(p_m, n) \leq 0, \tag{F.1c}$$

with $h(n(\theta)) \equiv \frac{(n(\theta))^{1+1/\varepsilon}}{1+1/\varepsilon}$ and $\Lambda(\theta) = \lambda + \theta(1 - \lambda)$.

Lagrangian. The Lagrangian associated with the planner's problem (F.1) is given by

$$\begin{aligned} \mathcal{L} = & \int U(\theta)d\Lambda(\theta) + \int \mu(\theta) (U'(\theta) - h'(n(\theta))n(\theta)\omega(p_m;\theta)) d\theta \\ & - \gamma[\phi[\int (U(\theta) + h(n(\theta)))dF(\theta) - y_f(p_m, n)] - y_m(p_m, n)]. \end{aligned}$$

Integrating by parts, we get

$$\begin{aligned} \mathcal{L} = & \int U(\theta)d\Lambda(\theta) - \int \mu'(\theta)U(\theta)d\theta + U(1)\mu(1) - U(0)\mu(0) \\ & - \int \mu(\theta)h'(n(\theta))n(\theta)\omega(p_m;\theta)d\theta \\ & - \gamma[\phi[\int (U(\theta) + h(n(\theta)))dF(\theta) - y_f(p_m, n)] - y_m(p_m, n)]. \end{aligned}$$

Since $U(1)$ and $U(0)$ are free we must have

$$\mu(0) = \mu(1) = 0.$$

First-order Conditions: $U(\theta)$. The first-order condition with respect to $U(\theta)$ leads to

$$\lambda(\theta) - \mu'(\theta) - \gamma\phi f(\theta) = 0.$$

Since $\mu(0) = 0$, integrating between 0 and θ , we get

$$\mu(\theta) = \Lambda(\theta) - \gamma\phi F(\theta).$$

Since $\mu(1) = 0$, we must also have

$$1 - \gamma\phi = 0.$$

Combining the two previous observations, we obtain

$$\frac{\mu(\theta)}{\gamma\phi} = \Lambda(\theta) - F(\theta). \quad (\text{F.2})$$

First-order Conditions: $n(\theta)$. The first-order condition with respect to $n(\theta)$ is given by

$$\begin{aligned} & \gamma\phi[y_{f,n(\theta)}(p_m, n) - h'(n(\theta)) + \frac{1}{\phi}y_{m,n(\theta)}(p_m, n)]f(\theta) \\ & = \mu(\theta)[h''(n(\theta))n(\theta) + h'(n(\theta))]\omega(p_m;\theta), \end{aligned} \quad (\text{F.3})$$

with $y_{f,n(\theta)} \equiv \partial y_f / \partial n(\theta)$ and $y_{m,n(\theta)} \equiv \partial y_m / \partial n(\theta)$. The first-order conditions of old technology firms, new technology firms, and households imply

$$\begin{aligned} y_{f,n(\theta)}(p_m, n) + p_m y_{m,n(\theta)}(p_m, n) &= w(\theta), \\ p_m^* &= 1/\phi, \\ h'(n(\theta)) &= w(\theta)(1 - \tau(\theta)). \end{aligned}$$

Thus, we can rearrange equation (F.3) into

$$\gamma\phi[w(\theta)\tau(\theta) + (p_m^* - p_m)y_{m,n(\theta)}]f(\theta) = \mu(\theta)h'(n(\theta))\left[\frac{\varepsilon(\theta) + 1}{\varepsilon(\theta)}\right]\omega(p_m; \theta), \quad (\text{F.4})$$

with $\varepsilon(\theta) \equiv d \ln(n(\theta)) / d \ln h'(n(\theta))$.

First-order Conditions: p_m . The first-order condition with respect to p_m is given by

$$\gamma\phi(p_m^* - p_m)y_{m,p_m} = \int \mu(\theta)h'(n(\theta))n(\theta)\omega_{p_m}(p_m; \theta)d\theta, \quad (\text{F.5})$$

with $y_{m,p_m} \equiv \partial y_m / \partial p_m$ and $\omega_{p_m} \equiv \partial \omega / \partial p_m$.

Optimal Tax on Machines Using equation (F.4) to substitute for $\mu(\theta)h'(n(\theta))/\gamma$ in equation (F.5) and noting that $\partial \ln y_m(p_m, n(\theta); \theta) / \partial \ln n(\theta) = 1$,²³ we obtain

$$p_m - p_m^* = \frac{\int \frac{\varepsilon(\theta)}{\varepsilon(\theta)+1} \cdot \frac{d \ln \omega(\theta)}{d \ln |y_m|} \cdot \tau(\theta) \cdot x(\theta) dF(\theta)}{|y_m| \left[1 - \int \frac{\varepsilon(\theta)}{\varepsilon(\theta)+1} \cdot \frac{d \ln \omega(\theta)}{d \ln |y_m|} \cdot \frac{y_m(\theta)}{|y_m|} dF(\theta) \right]}'$$

with $\frac{d \ln \omega(\theta)}{d \ln |y_m|} = \frac{\partial \ln p_m}{\partial \ln |y_m(p_m, n)|} \frac{d \ln \omega(p_m; \theta)}{d \ln p_m}$. Using $y_m^* = |y_m|$ and $t_m^* = p_m / p_m^* - 1$, this implies

$$t_m^* = \frac{\int \frac{\varepsilon(\theta)}{\varepsilon(\theta)+1} \frac{d \ln \omega(\theta)}{d \ln y_m^*} \tau(\theta) x(\theta) dF(\theta)}{p_m^* y_m^* \left[1 - \int \frac{\varepsilon(\theta)}{\varepsilon(\theta)+1} \frac{d \ln \omega(\theta)}{d \ln y_m^*} \frac{y_m(\theta)}{y_m^*} dF(\theta) \right]}. \quad (\text{F.6})$$

Optimal Income Tax. Equations (F.2) and (F.4) imply

$$[w(\theta)\tau(\theta) + (p_m^* - p_m)y_{m,n(\theta)}] = \frac{[\Lambda(\theta) - F(\theta)]}{f(\theta)} h'(n(\theta)) \left[\frac{\varepsilon(\theta) + 1}{\varepsilon(\theta)} \right] \omega(p_m; \theta).$$

Using again the fact that $\partial \ln y_m(p_m, n(\theta); \theta) / \partial \ln n(\theta) = 1$ and $h'(n(\theta)) = w(\theta)(1 - \tau(\theta))$, from the first-order condition of the household's utility maximization problem, this leads to

$$\tau(\theta) = \tau^*(\theta) - \frac{p_m - p_m^*}{p_m} \frac{p_m y_m(\theta)}{x(\theta)} (1 - \tau^*(\theta)), \quad (\text{F.7})$$

with

$$\tau^*(\theta) \equiv \frac{1}{1 + \frac{\varepsilon(\theta)}{\varepsilon(\theta)+1} \cdot \frac{f(\theta)}{(\Lambda(\theta) - F(\theta))\omega(p_m; \theta)}}.$$

²³Recall that $y_m(p_m, n(\theta); \theta)$ is implicitly defined as the solution to $p_m = \partial g(y_m(\theta), n(\theta); \theta) / \partial y_m(\theta)$. Since $g(\cdot, \cdot; \theta)$ is homogeneous of degree one, this is equivalent to $p_m = \partial g(y_m(\theta) / n(\theta), 1; \theta) / \partial y_m(\theta)$. Differentiating, we therefore get $\partial \ln y_m(p_m, n(\theta); \theta) / \partial \ln n(\theta) = 1$.

Equation (16). Combining equations (F.6) and (F.7), we obtain

$$\frac{t_m^*}{1 + t_m^*} = \frac{\int \frac{\varepsilon(\theta)}{\varepsilon(\theta)+1} \cdot \frac{d \ln \omega(\theta)}{d \ln y_m^*} \cdot \tau^*(\theta) \cdot x(\theta) dF(\theta)}{p_m y_m^* \left[1 - \int \frac{\varepsilon(\theta)}{\varepsilon(\theta)+1} \cdot \frac{d \ln \omega(\theta)}{d \ln y_m^*} \cdot \tau^*(\theta) \cdot \frac{y_m(\theta)}{y_m^*} dF(\theta) \right]} \quad (\text{F.8})$$

In the parametric example of Section 6, we have assumed

$$\varepsilon(\theta) = \varepsilon \text{ for all } \theta, \quad (\text{F.9})$$

$$\Lambda(\theta) = \lambda + \theta(1 - \lambda) \text{ for all } \theta, \quad (\text{F.10})$$

$$f(\theta) = 1 \text{ for all } \theta, \quad (\text{F.11})$$

$$F(\theta) = \theta \text{ for all } \theta. \quad (\text{F.12})$$

We therefore immediately get

$$\tau^*(\theta, p_m) = \frac{1}{1 + \frac{\varepsilon}{\varepsilon+1} \cdot \frac{1}{\lambda(1-\theta)\omega(p_m; \theta)}}. \quad (\text{F.13})$$

In Section 6, we have also established that

$$w(p_m; \theta) = (1 - \theta)^{-1/\gamma(p_m)},$$

which implies

$$\omega(p_m; \theta) = \frac{1}{\gamma(p_m)} \cdot \frac{1}{1 - \theta}.$$

Substituting into equation (F.13), we therefore get

$$\tau^*(\theta) = \frac{1}{1 + \frac{\varepsilon}{\varepsilon+1} \frac{\gamma(p_m)}{\lambda}} \equiv \tau^*. \quad (\text{F.14})$$

In Section (6), we have also established that

$$\frac{d \ln \omega(p_m; \theta)}{d \ln p_m} = -\beta \gamma(p_m),$$

which implies

$$\frac{d \ln \omega(\theta)}{d \ln |y_m|} = -\beta \gamma(p_m) \frac{\partial \ln p_m}{\partial \ln |y_m(p_m, n)|} = \frac{d \ln \omega}{d \ln y_m^*}. \quad (\text{F.15})$$

Combining equations (F.8), (F.9), (F.14) and (F.15), we obtain

$$\frac{t_m^*}{1 + t_m^*} = \frac{\frac{\varepsilon}{\varepsilon+1} \frac{d \ln \omega}{d \ln y_m^*} \tau^* \int x(\theta) dF(\theta)}{1 - \frac{\varepsilon}{\varepsilon+1} \frac{d \ln \omega}{d \ln y_m^*} \tau^* p_m y_m^*}, \quad (\text{F.16})$$

Letting $s_m \equiv \frac{p_m y_m^*}{\int x(\theta) dF(\theta) + p_m y_m^*}$, this leads to equation (16).

F.2 Proof of Proposition 4

From equation (16), we know that

$$\frac{t_m^*}{1+t_m^*} = \frac{\frac{\varepsilon}{\varepsilon+1} \frac{d \ln \omega}{d \ln y_m^*} \tau^*}{1 - \frac{\varepsilon}{\varepsilon+1} \frac{d \ln \omega}{d \ln y_m^*} \tau^*} \frac{1-s_m}{s_m},$$

with

$$\begin{aligned} \frac{d \ln \omega}{d \ln |y_m|} &= -\beta \gamma(p_m) \frac{\partial \ln p_m}{\partial \ln |y_m(p_m, n)|}, \\ \tau^* &= \frac{1}{1 + \frac{\varepsilon}{\varepsilon+1} \frac{\gamma(p_m)}{\lambda}}, \\ s_m &= \frac{p_m |y_m(p_m, n)|}{\int x(\theta) dF(\theta) + p_m |y_m(p_m, n)|}. \end{aligned}$$

This expression can be rearranged as

$$\frac{t_m^*}{1+t_m^*} = \frac{\Phi}{\rho - \Phi} \frac{1-s_m}{s_m} \quad (\text{F.17})$$

with

$$\begin{aligned} \Phi &= -\frac{\lambda \varepsilon \beta \gamma(p_m)}{(\varepsilon+1)\lambda + \varepsilon \gamma(p_m)}, \\ \rho &= \frac{\partial \ln |y_m(p_m, n)|}{\partial \ln p_m}. \end{aligned} \quad (\text{F.18})$$

We first demonstrate that Φ , s_m , and ρ can be expressed as functions of t_m^* and ϕ .

Using the fact that $p_m = (1+t_m^*)/\phi$, we can immediately rearrange equation (F.18) as

$$\Phi = -\frac{\lambda \varepsilon \beta \gamma((1+t_m^*)/\phi)}{(\varepsilon+1)\lambda + \varepsilon \gamma((1+t_m^*)/\phi)} \equiv \Phi(t_m^*, \phi). \quad (\text{F.19})$$

To express s_m as a function of t_m^* and ϕ , we further need to solve for the optimal labor supply of each agent, $n(\theta)$, which itself depends on the marginal income tax rates, $\tau(\theta)$. Together with equations (F.14), equation (F.7) implies

$$\tau(\theta) = \frac{(\varepsilon+1)\lambda - \frac{t_m^*}{1+t_m^*} \frac{p_m y_m(\theta)}{x(\theta)} \varepsilon \gamma(p_m)}{(\varepsilon+1)\lambda + \varepsilon \gamma(p_m)}.$$

From the first-order condition of the old technology firms, we know that

$$\frac{p_m y_m(\theta)}{x(\theta)} = -\beta \ln(1-\theta), \quad (\text{F.20})$$

which leads to

$$\tau(\theta) = \frac{(\varepsilon+1)\lambda + \frac{t_m^*}{1+t_m^*} \beta \varepsilon \gamma(p_m) \ln(1-\theta)}{(\varepsilon+1)\lambda + \varepsilon \gamma(p_m)}. \quad (\text{F.21})$$

The optimal labor supply is given by the agent's first-order condition

$$n(\theta) = ((1 - \tau(\theta))w(\theta))^\varepsilon. \quad (\text{F.22})$$

Combining equations (F.21) and (F.22) with the fact that $w(p_m; \theta) = (1 - \theta)^{-1/\gamma(p_m)}$, we get

$$n(\theta) = \left(\frac{\varepsilon\gamma(p_m)}{(\varepsilon + 1)\lambda + \varepsilon\gamma(p_m)} \right)^\varepsilon \left(1 - \frac{t_m^*}{1 + t_m^*} \beta \ln(1 - \theta) \right)^\varepsilon (1 - \theta)^{-\varepsilon/\gamma(p_m)},$$

and in turn,

$$\int w(\theta)n(\theta)d\theta = \left(\frac{\varepsilon\gamma(p_m)}{(\varepsilon + 1)\lambda + \varepsilon\gamma(p_m)} \right)^\varepsilon \int \left(1 - \frac{t_m^*}{1 + t_m^*} \beta \ln(1 - \theta) \right)^\varepsilon (1 - \theta)^{-\frac{1+\varepsilon}{\gamma(p_m)}} d\theta$$

Using equation (F.20), we further get

$$p_m y_m(\theta) = -\beta \ln(1 - \theta) \left(\frac{\varepsilon\gamma(p_m)}{(\varepsilon + 1)\lambda + \varepsilon\gamma(p_m)} \right)^\varepsilon \left(1 - \frac{t_m^*}{1 + t_m^*} \beta \ln(1 - \theta) \right)^\varepsilon (1 - \theta)^{-\frac{1+\varepsilon}{\gamma(p_m)}}, \quad (\text{F.23})$$

and in turn,

$$p_m |y_m| = -\beta \left(\frac{\varepsilon\gamma(p_m)}{(\varepsilon + 1)\lambda + \varepsilon\gamma(p_m)} \right)^\varepsilon \int \ln(1 - \theta) \left(1 - \frac{t_m^*}{1 + t_m^*} \beta \ln(1 - \theta) \right)^\varepsilon (1 - \theta)^{-\frac{1+\varepsilon}{\gamma(p_m)}} d\theta. \quad (\text{F.24})$$

The aggregate share of robots is therefore given by

$$s_m = \frac{\int \beta \ln(1 - \theta) \left(1 - \frac{t_m^*}{1 + t_m^*} \beta \ln(1 - \theta) \right)^\varepsilon (1 - \theta)^{-\frac{1+\varepsilon}{\gamma((1+t_m^*)/\phi)}} d\theta}{\int (\beta \ln(1 - \theta) - 1) \left(1 - \frac{t_m^*}{1 + t_m^*} \beta \ln(1 - \theta) \right)^\varepsilon (1 - \theta)^{-\frac{1+\varepsilon}{\gamma((1+t_m^*)/\phi)}} d\theta} \equiv s_m(t_m^*, \phi), \quad (\text{F.25})$$

where we have again used $p_m = (1 + t_m^*)/\phi$. Let us now turn to the elasticity ρ . From equation (F.20) and the fact that $w(p_m; \theta) = (1 - \theta)^{-1/\gamma(p_m)}$, we get

$$p_m y_m(\theta) = -\beta \ln(1 - \theta) n(\theta) (1 - \theta)^{-1/\gamma(p_m)}. \quad (\text{F.26})$$

Using the previous expression with the definition of $\rho \equiv \frac{\partial \ln |y_m(p_m, n)|}{\partial \ln p_m}$, we get

$$\rho = \int \frac{y(p_m, n(\theta); \theta)}{|y_m(p_m, n)|} \frac{d \ln w(p_m; \theta)}{d \ln p_m} d\theta - 1.$$

Combining the previous expressions with equations (F.23), (F.24), and using the fact that $\frac{d \ln w(p_m; \theta)}{d \ln p_m} = \beta \ln(1 - \theta)$ and $p_m = (1 + t_m^*)/\phi$, we get

$$\rho = \frac{\int (\beta \ln(1 - \theta) - 1) \ln(1 - \theta) \left(1 - \frac{t_m^*}{1 + t_m^*} \beta \ln(1 - \theta) \right)^\varepsilon (1 - \theta)^{-\frac{1+\varepsilon}{\gamma((1+t_m^*)/\phi)}} d\theta}{\int \ln(1 - \theta) \left(1 - \frac{t_m^*}{1 + t_m^*} \beta \ln(1 - \theta) \right)^\varepsilon (1 - \theta)^{-\frac{1+\varepsilon}{\gamma((1+t_m^*)/\phi)}} d\theta} \equiv \rho(t_m^*, \phi). \quad (\text{F.27})$$

At this point, we have established that the three statistics in equation (F.17) can be expressed as $\Phi(t_m^*, \phi)$, $\rho(t_m^*, \phi)$, and $s_m(t_m^*, \phi)$. We can therefore rearrange equation (F.17) as

$$H(t_m^*, \Phi(t_m^*, \phi), \rho(t_m^*, \phi), s_m(t_m^*, \phi)) = 0,$$

with

$$H(t_m^*, \Phi, \rho, s_r) \equiv \frac{\Phi}{\rho - \Phi} \cdot \frac{1 - s_m}{s_m} - \frac{t_m^*}{1 + t_m^*}.$$

By the Implicit Function Theorem, we have

$$\frac{dt_m^*}{d\phi} = -\frac{dH/d\phi}{dH/dt_m^*}. \quad (\text{F.28})$$

Since the tax on robots is chosen to maximize welfare, the second derivative of the government's value function, expressed as a function of t_m^* only, must be negative. Noting that H corresponds to its first derivative—which is equal to zero at the optimal tax—we therefore obtain

$$dH/dt_m^* < 0. \quad (\text{F.29})$$

Since $\gamma(\cdot)$ is a strictly increasing function, equation (F.19) implies

$$\frac{\partial \Phi(t_m^*, \phi)}{\partial \phi} > 0. \quad (\text{F.30})$$

To establish the monotonicity of s_m and ρ with respect to ϕ , it is convenient to introduce the following function:

$$d(t_m^*, \phi, \zeta; \theta) = (1 - \beta \frac{t_m^*}{1 + t_m^*} \ln(1 - \theta))^\epsilon (1 - \theta)^{-\frac{1+\epsilon}{\gamma((1+t_m^*)/\phi)}} (\ln(1 - \theta))^{-\zeta}.$$

By construction, d is log-supermodular in (ϕ, ζ, θ) . Since log-supermodularity is preserved by integration, the following function,

$$D(\phi, \zeta) = \int d(t_m^*, \phi, \zeta; \theta) d\theta,$$

is also log-supermodular. It follows that

$$\begin{aligned} \frac{D(\phi, \zeta = 0)}{D(\phi, \zeta = -1)} &= \frac{\int (1 - \beta \frac{t_m^*}{1 + t_m^*} \ln(1 - \theta))^\epsilon (1 - \theta)^{-\frac{1+\epsilon}{\gamma((1+t_m^*)/\phi)}} d\theta}{\int (\ln(1 - \theta)) (1 - \beta \frac{t_m^*}{1 + t_m^*} \ln(1 - \theta))^\epsilon (1 - \theta)^{-\frac{1+\epsilon}{\gamma((1+t_m^*)/\phi)}} d\theta} \text{ is increasing in } \phi, \\ \frac{D(\phi, \zeta = -2)}{D(\phi, \zeta = -1)} &= \frac{\int (\ln(1 - \theta))^2 (1 - \beta \frac{t_m^*}{1 + t_m^*} \ln(1 - \theta))^\epsilon (1 - \theta)^{-\frac{1+\epsilon}{\gamma((1+t_m^*)/\phi)}} d\theta}{\int (\ln(1 - \theta)) (1 - \beta \frac{t_m^*}{1 + t_m^*} \ln(1 - \theta))^\epsilon (1 - \theta)^{-\frac{1+\epsilon}{\gamma((1+t_m^*)/\phi)}} d\theta} \text{ is decreasing in } \phi. \end{aligned}$$

Noting that

$$\begin{aligned} s_m &= \frac{1}{1 - \frac{1}{\beta} \frac{D(\phi, \zeta = 0)}{D(\phi, \zeta = -1)}}, \\ \rho &= \beta \frac{D(\phi, \zeta = -2)}{D(\phi, \zeta = -1)} - 1, \end{aligned}$$

we obtain that

$$\frac{\partial s_m(t_m^*, \phi)}{\partial \phi} > 0, \quad (\text{F.31})$$

$$\frac{\partial \rho(t_m^*, \phi)}{\partial \phi} < 0. \quad (\text{F.32})$$

Since $\frac{\partial H}{\partial \Phi} < 0$, $\frac{\partial H}{\partial s_m} < 0$, and $\frac{\partial H}{\partial \rho} > 0$, inequalities (F.30)-(F.32) imply

$$\frac{dH}{d\phi} = \frac{\partial H}{\partial \Phi} \frac{\partial \Phi}{\partial \phi} + \frac{\partial H}{\partial s_m} \frac{\partial s_m}{\partial \phi} + \frac{\partial H}{\partial \rho} \frac{\partial \rho}{\partial \phi} < 0.$$

Combining this observation with equation (F.28) and (F.29), we conclude that $dt_m^*/d\phi < 0$.

F.3 Proof of Proposition 5

Optimal Tax on Machines (t_m^*). Consider the optimal tax on machines, t_m^* . In the proof of Proposition 4, we have already established that

$$\frac{t_m^*}{1 + t_m^*} = \frac{\Phi}{\rho - \Phi} \frac{1 - s_m}{s_m}.$$

with

$$\begin{aligned} \Phi &= -\frac{\lambda \varepsilon \beta \gamma ((1 + t_m^*)/\phi)}{(\varepsilon + 1)\lambda + \varepsilon \gamma ((1 + t_m^*)/\phi)}, \\ s_m &= \frac{\int \beta \ln(1 - \theta) (1 - \frac{t_m^*}{1 + t_m^*} \beta \ln(1 - \theta))^\varepsilon (1 - \theta)^{-\frac{1 + \varepsilon}{\gamma((1 + t_m^*)/\phi)}} d\theta}{\int (\beta \ln(1 - \theta) - 1) (1 - \frac{t_m^*}{1 + t_m^*} \beta \ln(1 - \theta))^\varepsilon (1 - \theta)^{-\frac{1 + \varepsilon}{\gamma((1 + t_m^*)/\phi)}} d\theta}, \\ \rho &= \frac{\int (\beta \ln(1 - \theta) - 1) \ln(1 - \theta) (1 - \frac{t_m^*}{1 + t_m^*} \beta \ln(1 - \theta))^\varepsilon (1 - \theta)^{-\frac{1 + \varepsilon}{\gamma((1 + t_m^*)/\phi)}} d\theta}{\int \ln(1 - \theta) (1 - \frac{t_m^*}{1 + t_m^*} \beta \ln(1 - \theta))^\varepsilon (1 - \theta)^{-\frac{1 + \varepsilon}{\gamma((1 + t_m^*)/\phi)}} d\theta}. \end{aligned}$$

It follows that Φ is strictly decreasing in λ , whereas s_m and ρ are independent of λ . Invoking the Implicit Function Theorem in the exact same way as we did in the proof of Proposition 4, we therefore get: $dt_m^*/d\lambda > 0$.

Constrained Optimal Tax on Machines (t_m^c). Let us first characterize the constrained optimal tax on machines, t_m^c . Suppose that the income tax schedule is exogenously set at $T = T^c$, with T^c a linear tax schedule with constant marginal tax rate $\tau^c \in [0, 1]$. Proposition (1) implies

$$\frac{t_m^c}{1 + t_m^c} = [\lambda(1 - \tau^c) - \tau^c \varepsilon] \int \frac{w(\theta) n(\theta)}{p_m y_m^*} \frac{\delta \ln w(\theta)}{\delta \ln y_m^*} \Big|_{\delta T=0} d\theta. \quad (\text{F.33})$$

Next we compute $w(\theta) n(\theta)$, $p_m y_m^*$, and $\frac{\delta \ln w(\theta)}{\delta \ln y_m^*} \Big|_{\delta T=0}$, all evaluated at $t_m = t_m^c$. As already noted in Section 6.1, the equilibrium wage schedule is given by

$$w(p_m; \theta) = (1 - \theta)^{-1/\gamma(p_m)}; \quad (\text{F.34})$$

and as already argued in the proof of Proposition 4, the labor supply of households satisfies (F.22),

$$n(\theta) = ((1 - \tau^c)w(\theta))^\varepsilon.$$

Combining these two expressions, earnings can be expressed as

$$w(\theta)n(\theta) = (1 - \tau^c)^\varepsilon(1 - \theta)^{-(1+\varepsilon)/\gamma(p_m)}. \quad (\text{F.35})$$

As already argued in the proof of Proposition 4, the firm's demand for machines also satisfies (F.20),

$$\frac{p_m y_m(\theta)}{w(\theta)n(\theta)} = -\beta \ln(1 - \theta),$$

which further implies

$$p_m y_m^* = -\beta(1 - \tau^c)^\varepsilon \int (1 - \theta)^{-(1+\varepsilon)/\gamma(p_m)} \ln(1 - \theta) d\theta = \frac{\beta(1 - \tau^c)^\varepsilon}{[1 - (1 + \varepsilon)/\gamma(p_m)]^2}. \quad (\text{F.36})$$

Finally, note that $\frac{\delta \ln w(\theta)}{\delta \ln y_m^*} \Big|_{\delta T=0}$ can be expressed as

$$\frac{\delta \ln w(\theta)}{\delta \ln y_m^*} \Big|_{\delta T=0} = \frac{d \ln w(p_m; \theta) / d \log p_m}{d \ln y_m^*(p_m, n(p_m); \theta) / d \log p_m}.$$

Equations (F.34) and (F.36) therefore imply

$$\frac{\delta \ln w(\theta)}{\delta \ln y_m^*} \Big|_{\delta T=0} = -\frac{[1 - (1 + \varepsilon)/\gamma(p_m)]\beta \ln(1 - \theta)}{2\beta(1 + \varepsilon) + 1 - (1 + \varepsilon)/\gamma(p_m)}. \quad (\text{F.37})$$

Combining equations (F.36), (F.36), (F.36), and (F.37), we obtain

$$\frac{t_m^c}{1 + t_m^c} = \frac{[\lambda(1 - \tau^c) - \tau^c \varepsilon][1 - (1 + \varepsilon)/\gamma(p_m)]}{2\beta(1 + \varepsilon) + 1 - (1 + \varepsilon)/\gamma(p_m)}. \quad (\text{F.38})$$

Following the same strategy as in the proof of Proposition 4, we can rearrange the previous expression as

$$H(t_m^c, \lambda) = 0,$$

with

$$H(t_m^c, \lambda) \equiv \frac{[\lambda(1 - \tau^c) - \tau^c \varepsilon][1 - \frac{1+\varepsilon}{\gamma((1+t_m^c)/\phi)}]}{2\beta(1 + \varepsilon) + 1 - \frac{1+\varepsilon}{\gamma((1+t_m^c)/\phi)}} - \frac{t_m^c}{1 + t_m^c}.$$

By the Implicit Function Theorem, we have

$$\frac{dt_m^c}{d\lambda} = -\frac{dH/d\lambda}{dH/dt_m^c}.$$

Since the tax on robots is chosen to maximize welfare, the second derivative of the government's value function, expressed as a function of t_m^c only, must be negative. Noting that H corresponds to its first derivative—which is equal to zero at the optimal tax—we therefore obtain $dH/dt_m^c < 0$. Since $dH/d\lambda > 0$, we conclude that $\frac{dt_m^c}{d\lambda} > 0$.

F.4 Numerical Example

Baseline Economy. In the numerical example of Section 6.3, we set $\varepsilon = 0.5$ and normalize $\phi = 1$. We then set α and β so that if T is a linear income tax schedule with $\tau^c = 27\%$ and there is no tax on machine $t_m = 0$, the share of machines in gross output and the elasticity of relative wage satisfy

$$s_m = 2.1\%, \quad (\text{F.39})$$

$$\frac{d \ln \omega}{d \ln y_m^*} = 0.5\%. \quad (\text{F.40})$$

By equations (F.35) and (F.36), we know that

$$s_m \equiv \frac{p_m y_m^*}{\int w(\theta) n(\theta) d\theta + p_m y_m^*} = \frac{\beta}{\beta + 1 - (1 + \varepsilon)/\gamma(p_m)}. \quad (\text{F.41})$$

By equations (F.15) and (F.27), evaluated with a zero tax on machines, we also know that

$$\frac{d \ln \omega}{d \ln y_m^*} = \frac{\beta \gamma(p_m) [1 - (1 + \varepsilon)/\gamma(p_m)]}{2\beta + 1 - (1 + \varepsilon)/\gamma(p_m)}. \quad (\text{F.42})$$

Using $\varepsilon = 0.5$, $\gamma(p_m) \equiv 1/(\alpha - \beta \ln p_m)$, and $p_m = 1$, we can rearrange equations (F.39)-(F.42) as

$$\begin{aligned} \frac{\beta}{\beta + 1 - 1.5\alpha} &= 2.1\%, \\ \frac{\beta[1 - 1.5\alpha]}{\alpha[2\beta + 1 - 1.5\alpha]} &= 0.5\%. \end{aligned}$$

The solution of the previous system gives $\alpha = 0.57$, and $\beta = 0.003$, as argued in Section 6.3.

Optimal Taxes using Propositions 2 and 3. In Section 6.3, we discuss, together with the optimal tax t_m^* and the constrained optimal tax t_m^c , the taxes on machines that would be obtained by using the formulas in Propositions 2 and 3 in the baseline economy: $(t_m)_{\delta \bar{U}=0} \simeq 2.28\%$ and $(t_m)_{\delta \bar{U}=0, \text{f.o.a}} \simeq 2.18\%$, respectively. We now formally describe how we derive these two tax rates.

Consider first $(t_m)_{\delta \bar{U}=0}$. From Proposition 2, we know that

$$\frac{(t_m)_{\delta \bar{U}=0}}{1 + (t_m)_{\delta \bar{U}=0}} = \int \tau^c \frac{(w(\theta) n(\theta))_{\text{baseline}}}{(p_m y_m^*)_{\text{baseline}}} \frac{\varepsilon}{\varepsilon + 1} \left(\frac{\delta \ln \omega(\theta)}{\delta \ln y_m^*} \Big|_{\delta \bar{U}=0} \right)_{\text{baseline}} d\theta. \quad (\text{F.43})$$

In the baseline economy, we also know that $\phi = 1$ and $t_m = 0$, which implies $p_m = 1$. Since the baseline economy also features a constant marginal income tax rate τ^c , equation (F.35) in Appendix F.4 implies

$$(w(\theta) n(\theta))_{\text{baseline}} = (1 - \tau^c)^\varepsilon (1 - \theta)^{-\alpha(1+\varepsilon)}, \quad (\text{F.44})$$

whereas equation F.36 implies

$$(p_m y_m^*)_{\text{baseline}} = \frac{\beta(1 - \tau^c)^\varepsilon}{[1 - \alpha(1 + \varepsilon)]^2}. \quad (\text{F.45})$$

To compute $\left(\frac{\delta \ln \omega(\theta)}{\delta \ln y_m^*} \Big|_{\delta \bar{U}=0}\right)_{\text{baseline}}$, note that

$$\left(\frac{\delta \ln \omega(\theta)}{\delta \ln y_m^*} \Big|_{\delta \bar{U}=0}\right)_{\text{baseline}} = \frac{(d \ln \omega(p_m; \theta) / d \ln p_m)_{\text{baseline}}}{(\delta \ln y_m^* / \delta \ln p_m \Big|_{\delta \bar{U}=0})_{\text{baseline}}}, \quad (\text{F.46})$$

$$\begin{aligned} \left(\frac{\delta \ln y_m^*}{\delta \ln p_m} \Big|_{\delta \bar{U}=0}\right)_{\text{baseline}} &= \left(\frac{\partial \ln |y_m(p_m, n)|}{\partial \ln p_m}\right)_{\text{baseline}} \\ &+ \int \left(\frac{\partial \ln |y_m(p_m, n)|}{\partial \ln n(\theta)}\right)_{\text{baseline}} \left(\frac{\delta \ln n(\theta)}{\delta \ln p_m} \Big|_{\delta \bar{U}=0}\right)_{\text{baseline}} d\theta, \end{aligned} \quad (\text{F.47})$$

$$\left(\frac{\delta \ln n(\theta)}{\delta \ln p_m} \Big|_{\delta \bar{U}=0}\right)_{\text{baseline}} = -\frac{\epsilon}{\epsilon + 1} \left(\frac{d \ln \omega(p_m; \theta)}{d \ln p_m}\right)_{\text{baseline}}, \quad (\text{F.48})$$

where the third equality has been established in Appendix D.3, with

$$\left(\frac{d \ln \omega(p_m; \theta)}{d \ln p_m}\right)_{\text{baseline}} = -\frac{\beta}{\alpha}, \quad (\text{F.49})$$

$$\left(\frac{\partial \ln |y_m(p_m, n)|}{\partial \ln p_m}\right)_{\text{baseline}} = \frac{\int (\beta \ln(1 - \theta) - 1) \ln(1 - \theta) (1 - \theta)^{-\alpha(1 + \epsilon)} d\theta}{\int \ln(1 - \theta) (1 - \theta)^{-\alpha(1 + \epsilon)} d\theta}, \quad (\text{F.50})$$

$$\left(\frac{\partial \log |y_m(p_m, n)|}{\partial \log n(\theta)}\right)_{\text{baseline}} = 1, \quad (\text{F.51})$$

where the first equality has been established in Section (6) and the second and third derive from equation (F.26) in Appendix F.2. Combining equations (F.43)-(F.51), we obtain

$$\frac{(t_m)_{\delta \bar{U}=0}}{1 + (t_m)_{\delta \bar{U}=0}} = \frac{\epsilon}{\epsilon + 1} \frac{\tau^c}{\alpha} \frac{[1 - \alpha(1 + \epsilon)]^2}{2\beta + 1 - \alpha(1 + \epsilon) + \beta\epsilon(1 - \frac{1}{\alpha(\epsilon + 1)})}.$$

Next consider $(t_m)_{\delta \bar{U}=0, \text{f.o.a.}}$. From Proposition 3, we know that

$$\frac{(t_m)_{\delta \bar{U}=0, \text{f.o.a.}}}{1 + (t_m)_{\delta \bar{U}=0, \text{f.o.a.}}} = \int \tau^c \frac{(w(\theta)n(\theta))_{\text{baseline}}}{(p_m y_m^*)_{\text{baseline}}} \frac{\epsilon}{\epsilon + 1} \left(\frac{\delta \ln \omega(\theta)}{\delta \ln y_m^*} \Big|_{\delta T=0}\right)_{\text{baseline}} d\theta, \quad (\text{F.52})$$

since the variation $\delta T = 0$ is a budget-balanced variation. To compute $\left(\frac{\delta \ln \omega(\theta)}{\delta \ln y_m^*} \Big|_{\delta T=0}\right)_{\text{baseline}}$, we again use

$$\left(\frac{\delta \ln \omega(\theta)}{\delta \ln y_m^*} \Big|_{\delta T=0}\right)_{\text{baseline}} = \frac{(d \ln \omega(p_m; \theta) / d \ln p_m)_{\text{baseline}}}{(\delta \ln y_m^* / \delta \ln p_m \Big|_{\delta T=0})_{\text{baseline}}},$$

where $(d \ln \omega(p_m; \theta) / d \ln p_m)_{\text{baseline}} = -\beta/\alpha$ and $(\delta \ln y_m^* / \delta \ln p_m \Big|_{\delta T=0})_{\text{baseline}}$ can be computed using (F.36) as in Appendix F.3. This leads to

$$\left(\frac{\delta \ln \omega(\theta)}{\delta \ln y_m^*} \Big|_{\delta T=0}\right)_{\text{baseline}} = \frac{\beta}{\alpha} \frac{1 - \alpha(1 + \epsilon)}{1 - \alpha(1 + \epsilon) + 2\beta(1 + \epsilon)}. \quad (\text{F.53})$$

Combining equations (F.44), (F.45), (F.52), and (F.53), we obtain

$$\frac{(t_m)_{\delta \bar{U}=0, \text{f.o.a.}}}{1 + (t_m)_{\delta \bar{U}=0, \text{f.o.a.}}} = \frac{\epsilon}{\epsilon + 1} \frac{\tau^c}{\alpha} \frac{[1 - \alpha(1 + \epsilon)]^2}{2\beta + 1 - \alpha(1 + \epsilon) + 2\beta\epsilon}.$$