

# Nonlinear Tax Incidence and Optimal Taxation in General Equilibrium\*

Dominik Sachs      Aleh Tsyvinski      Nicolas Werquin<sup>†</sup>  
LMU Munich                  Yale                  Toulouse

December 18, 2017

## Abstract

We study the incidence and optimal design of nonlinear income taxes in a Mirrleesian economy with a continuum of endogenous wages. We characterize in closed form the incidence of any nonlinear tax reform on individual variables (labor supplies, wages, utilities) and aggregate variables (government revenue, social welfare) by showing that this problem can be formalized as an integral equation. The general-equilibrium effects of tax reforms are driven by the interaction between the existing marginal tax rates and the complementarities between skills in production. We derive a simple formula for optimal taxes and extend two classical results: closed-form expression for the top tax rate and the U-shape of marginal tax rates. We further expose our results quantitatively using production functions that allow for distance-dependent elasticities of substitution between skills.

---

\*We thank Laurence Ales, Costas Arkolakis, Andy Atkeson, Alan Auerbach, Felix Bierbrauer, Carlos Da Costa, Cécile Gaubert, Austan Goolsbee, Piero Gottardi, Nathan Hendren, James Hines, Claus Kreiner, Tim Lee, Arash Nekoei, Michael Peters, Emmanuel Saez, Florian Scheuer, Hakan Selin, Chris Sleet, Stefanie Stantcheva, Philip Ushchev, Gianluca Violante, Ivan Werning, Sevin Yeltekin and Floris Zoutman for insightful comments and discussions.

<sup>†</sup>Toulouse School of Economics, University of Toulouse Capitole, Toulouse, France.

# Introduction

We study the incidence and the optimal design of nonlinear income taxes in a general equilibrium [Mirrlees \(1971\)](#) economy. There is a continuum of skills that are imperfectly substitutable in production. The aggregate production function uses as inputs the labor effort of all skills. The wage, or marginal product, of each skill type is endogenous. Specifically, it is decreasing in the aggregate labor effort of its own skill if the marginal productivity of labor is decreasing, and increasing in the aggregate labor effort of those skills that are complementary in production. Agents choose their labor supply optimally given their wage and the tax schedule.

We connect two classical strands of the public finance literature that have so far been somewhat disconnected: the tax incidence literature ([Harberger, 1962](#); [Kotlikoff and Summers, 1987](#); [Fullerton and Metcalf, 2002](#)), and the literature on optimal nonlinear income taxation ([Mirrlees, 1971](#); [Stiglitz, 1982](#); [Diamond, 1998](#); [Saez, 2001](#)). The objective of the tax incidence analysis is to characterize the first-order effects of locally reforming a given, potentially suboptimal, tax system on the distribution of individual wages, labor supplies, and utilities, as well as on government revenue and social welfare in partial and general equilibrium. We provide closed-form analytical formulas for the incidence of any tax reform in our environment with arbitrarily nonlinear taxes and a continuum of endogenous wages. A characterization of optimal taxes in general equilibrium is then obtained by imposing that no tax reform has a positive impact on social welfare.

We start by focusing on the incidence of general tax reforms in a model where the utility function is quasilinear – we generalize this assumption later. When wages are exogenous, the effects of a tax change on the labor supply of a given agent can be easily derived as a function of the elasticity of labor supply of that agent ([Saez, 2001](#)). The key difficulty in general equilibrium is that this in turn impacts the wage, and thus the labor supply, of every other individual. This further affects the wage distribution, which influences labor supply decisions, and so on. Solving for the fixed point in the labor supply adjustment of each agent is the key step in the tax incidence analysis and the primary technical challenge of our paper.

We show that this a priori complex problem of deriving the effects of an arbitrary tax reform on individual labor supply can be mathematically formalized as solving an integral equation. The tools of the theory of integral equations allow us to derive an analytical solution to this problem for a general production function, which

furthermore has a clear economic interpretation. Specifically, this solution can be represented a series; its first term is the partial-equilibrium impact of the reform, and each of its subsequent terms captures a successive round of cross-wage feedback effects in general equilibrium. These are expressed in terms of meaningful elasticities for any arbitrary production function, i.e., allowing for any pattern of complementarities between skills in production. Once we have characterized the incidence of tax reforms on labor supply, it is straightforward to derive the incidence on individual wages and indirect utilities. We show in particular that in general equilibrium, if all the skills are complements in production, an increase in the marginal tax rate at a given income level, conditional on an absolute tax rise, raises the welfare of agents earning that income, and reduces everyone else's welfare.

Next, we analyze the aggregate incidence of tax reforms on government revenue and social welfare. We derive a general formula that shows that, in response to an increase in the marginal tax rate at a given income level, the standard tax incidence formula obtained in the model with exogenous wages is modified to include a general-equilibrium term that depends on the covariation between (i) the shape of the schedule of marginal tax rates, weighted by the labor supply elasticities, in the initial economy, and (ii) the pattern of complementarities in production with the skill where the tax rate has been perturbed. The optimal tax schedule is immediately obtained as a by-product of this formula by equating to zero the impact of any tax reform on social welfare.

We derive further implications of this general formula by focusing on specific functional forms for the initial economy's tax schedule and the production function. First, we show that if the initial tax schedule is linear and the labor supply elasticity is the same for all agents, then the general-equilibrium forces have no impact on aggregate government revenue in addition to those already obtained assuming exogenous wages. To understand this result, suppose that the government raises the marginal tax rate at a given income level. This disincentivizes the labor supply of the agents who initially earn that income, which in turn raises their own wage, since the marginal product of labor is decreasing, and lowers the wage of the skills that are complementary in production. By Euler's homogeneous function theorem, the impact of these wage adjustments on aggregate income is equal to zero if the production function has constant returns to scale. If moreover the labor supply elasticity is the same for all agents, the aggregate income change is also zero after the adjustment of labor supply due to these wage changes. Since the marginal tax rate is originally the same for

the whole population, the impact on government revenue is also equal to zero. Next, we analyze the more general case where the marginal tax rates are monotonic (say, increasing) with income, so that the initial tax schedule is progressive, and assume in addition that the production function has a constant elasticity of substitution (CES). In this case, the benefits of reforming the tax schedule in the direction of a higher progressivity are larger in general equilibrium than the conventional formula assuming exogenous wages would predict. This is because an increase in the marginal tax rate on high incomes leads to an increase in their wage that raises government revenue by a larger amount (since their marginal tax rate is initially higher) than the same tax hike implemented at lower income levels. In other words, starting from a progressive tax code, the general equilibrium forces raise the revenue gains from increasing further the progressivity of the tax schedule.

Our numerical simulations quantify and generalize these insights. For a CES production function with an elasticity of substitution calibrated to the U.S. economy, we find that the efficiency loss from raising the marginal tax rate on top incomes is lower once the general equilibrium forces are taken into account, compared to the values we would obtain by applying the formula derived assuming exogenous wages. We then apply our formulas to another commonly used production function in the literature, namely, the Translog technology.<sup>1</sup> Specifically, we introduce a sophisticated functional form for the parameters that formalizes and makes operational the idea that workers with closer productivities are more substitutable. Our main insights are qualitatively similar, but slightly reinforced, for this “distance-dependent” specification of technology.

We then consider various generalizations of our baseline model, and show that the techniques and intuitions we derived in our simple environment carry over to more sophisticated frameworks with no additional technical difficulties. We first generalize our results to general individual preferences that allow for income effects. Second, we allow for several sectors or education levels in the economy, where there is a continuum of skills within each group (and as a consequence, overlapping wage distributions). Third, we allow for both intensive margin (hours) and extensive margin (participation) choices of labor supply. Fourth, we allow for non-constant returns to scale. For each of these extensions, we derive closed-form tax incidence formulas using the same methodological tools as in our baseline environment.

---

<sup>1</sup>See [Bucci and Ushchev \(2014\)](#) for a careful study of various production functions with a continuum of inputs and constant or variable elasticities of substitution.

Next, we derive the implications of our analysis regarding the optimal (social welfare-maximizing) tax schedule. Recall that our tax incidence analysis immediately delivers a general characterization of optimal taxes. In the spirit of [Piketty \(1997\)](#), [Saez \(2001\)](#), [Chetty \(2009\)](#), we aim to obtain an optimum formula that depends on a parsimonious number of parameters which can be estimated empirically. To do so, we specialize our production function to have a constant elasticity of substitution (CES) between any pairs of types. This allows us to derive particularly sharp and transparent theoretical insights and quantitative results.

There are two key differences between our optimal tax formula and that typically derived in the literature assuming exogenous wages ([Diamond, 1998](#); [Saez, 2001](#)). First, because of the decreasing marginal productivity of labor, the relevant labor supply elasticity is smaller, implying lower disincentive effects of raising the marginal taxes, and higher optimal rates. This is because a higher tax rate reduces labor supply, which in turn raises the wage, and hence the labor supply, of these agents. Second, marginal tax rates should be lower (resp., higher) for agents whose welfare is valued less (resp., more) than average. This is because an increase in the marginal tax rate of a given skill type increases her wage at the expense of all other types. This term generalizes the insight obtained by [Stiglitz \(1982\)](#) in a model with two skills to the workhorse model of income taxation.

We finally extend two of the most influential results from the Mirrleesian literature to our framework with endogenous wages. The first is the optimal top tax rate formula of [Saez \(2001\)](#). We derive a particularly simple closed-form generalization of this result in terms of one additional sufficient statistic, namely, the (finite) elasticity of substitution between skills. The second result is the familiar U-shaped pattern of optimal marginal tax rates first obtained by [Diamond \(1998\)](#). The general equilibrium forces not only confirm this pattern, but make it even more pronounced, with a stronger dip in the bulk of the income distribution. We provide an economic intuition for this more pronounced U-shape that is based on the same economic reasoning than our results on the incidence of tax reforms. Thus, besides the study of the incidence of reforms of the current tax system, our tax reform approach also complements the mechanism-design approach to optimal taxation in general equilibrium by providing a clearer economic understanding of the optimality conditions. Numerical simulations confirm these insights. Moreover, when the production function is Translog and the elasticity of substitution is distance-dependent, the optimal tax schedule is very close to the Cobb-Douglas limit.

**Related Literature.** This paper is related to the literature on tax incidence (see, e.g., [Harberger \(1962\)](#) and [Shoven and Whalley \(1984\)](#) for the seminal papers, [Hines \(2009\)](#) for emphasizing the importance of general equilibrium in taxation, and [Kotlikoff and Summers \(1987\)](#) and [Fullerton and Metcalf \(2002\)](#)) for comprehensive surveys. Our paper extends this framework to an economy with a continuum of (labor) inputs with arbitrary nonlinear tax schedules, i.e., we study tax incidence in the workhorse model of optimal nonlinear labor income taxation of ([Mirrlees, 1971](#); [Diamond, 1998](#)).

The optimal taxation problem in general equilibrium with arbitrary nonlinear tax instruments has originally been studied by [Stiglitz \(1982\)](#) in a model with two types. The key result of [Stiglitz \(1982\)](#) is that at the optimum tax system, general equilibrium forces lead to a more regressive tax schedule. In the recent optimal taxation literature, there are two strands that relate to our work. First, a series of important contributions by [Scheuer \(2014\)](#); [Rothschild and Scheuer \(2013, 2014\)](#); [Scheuer and Werning \(2017\)](#), [Chen and Rothschild \(2015\)](#), [Ales, Kurnaz, and Sleet \(2015\)](#), [Ales and Sleet \(2016\)](#), and [Ales, Bellofatto, and Wang \(2017\)](#) form the modern analysis of optimal nonlinear taxes in general equilibrium.<sup>2</sup> Specifically, [Rothschild and Scheuer \(2013, 2014\)](#) generalize [Stiglitz \(1982\)](#) to a setting with  $N$  sectors and a continuum of (infinitely substitutable) skills in each sector, leading to a multidimensional screening problem. [Ales, Kurnaz, and Sleet \(2015\)](#) and [Ales and Sleet \(2016\)](#) microfound the production function by incorporating an assignment model into the Mirrlees framework and study the implications of technological change and CEO-firm matching for optimal taxation. Our baseline model is simpler than those of [Rothschild and Scheuer \(2013, 2014\)](#) and [Ales, Kurnaz, and Sleet \(2015\)](#). In particular, different types earn different wages (there is no overlap in the wage distributions of different types, as opposed to the framework of [Rothschild and Scheuer \(2013, 2014\)](#)), and the production function is exogenous (in contrast to [Ales, Kurnaz, and Sleet \(2015\)](#)).<sup>3</sup> The general distinction is that these papers focus on optimal taxation by applying the methods of mechanism design, whereas our use of the variational approach and integral equa-

---

<sup>2</sup>[Rothstein \(2010\)](#) studies the desirability of EITC-type tax reforms in a model with heterogeneous labor inputs and nonlinear taxation. He only considers own-wage effects, however, and no cross-wage effects. Further he treats intensive margin labor supply responses as occurring along linearized budget constraints.

<sup>3</sup>Finally, our setting is distinct from those of [Scheuer and Werning \(2016, 2017\)](#), whose modeling of the technology is such that the general equilibrium effects cancel out at the optimum tax schedule, so that the formula of [Mirrlees \(1971\)](#) extends to their general production functions. We discuss in detail the difference between our framework and theirs in [Appendix A.2.4](#).

tions allow us to study more generally the incidence of reforming in any direction an arbitrary tax system – as we show, the (possibly suboptimal) tax system to which the reform is applied is a crucial determinant of the direction and size of the general equilibrium effects. Moreover, for optimal income taxes, our setting and methods allow us to get sharper and novel characterizations: transparent optimum tax formula, closed-form expression for the top tax rate, generalization of the U-shape of the optimal marginal tax rates. Note finally that we also analyze an extension of our baseline framework to production functions that allow for overlapping wage distribution in Section 4.2.

Our modeling of the production function is motivated by an empirical literature that estimates the impact of immigration on the native wage distribution and groups workers according to their position in the wage distribution (Card (1990), Borjas et al. (1997), Dustmann, Frattini, and Preston (2013)). The empirical literature on immigration is a useful benchmark because it studies the impact of labor supply shocks of certain skills on relative wages, which is exactly the channel we want to analyze in our tax setting (except that in our model the labor supply shocks are induced by tax reforms). An alternative in the immigration literature is to group workers by education levels (Borjas, 2003; Card, 2009). We fully extend our analysis and results to a production function with different education groups in Section 4.2.

Our study of tax incidence is based on a variational, or “tax reform” approach, originally pioneered by Piketty (1997), Saez (2001, 2002), and extended to several other contexts by, e.g., Kleven, Kreiner, and Saez (2009) and Golosov, Tsyvinski, and Werquin (2014). In this paper we extend this approach to the general equilibrium framework with endogenous wages. We derive a parsimonious and intuitive extension of the Diamond (1998) formula for optimal marginal tax rates in terms of sufficient statistics, and show that the general-equilibrium correction to the optimum is U-shaped. We then derive a closed-form expression for the optimal top tax rate and thereby extend that of Saez (2001) to endogenous wages.<sup>4</sup>

Finally, our paper is related to the literature that characterizes optimal government policy, within restricted classes of nonlinear tax schedules, in general equilibrium extensions of the continuous-type Mirrleesian framework. Heathcote, Storesletten,

---

<sup>4</sup>Our generalization of the optimal top tax rate to the case of endogenous wages is related to Piketty, Saez, and Stantcheva (2014), who extend the Saez (2001) top tax formula to a setting with a compensation bargaining channel using a variational approach. More generally, Rothschild and Scheuer (2016) study optimal taxation in the presence of rent-seeking. In this paper we abstract from such considerations and assume that individuals are paid their marginal productivity.

and Violante (2016) study optimal tax progressivity in a model where agents face idiosyncratic risk and can invest in their skills. Itskhoki (2008) and Antras, de Gortari, and Itskhoki (2016) characterize the impact of distortionary redistribution of the gains from trade in an open economy. Their production functions are CES with a continuum of skills and restrict the tax schedule to be of the CRP functional form. On the one hand, our model is simpler than their framework as we study a static and closed economy with exogenous skills. On the other hand, for most of our theoretical analysis we do not restrict ourselves to a particular functional form for taxes nor the production function. Our papers share, however, one important goal: to derive simple closed form expressions for the effects of tax reforms in general-equilibrium Mirrleesian environments.

This paper is organized as follows. Section 1 describes our framework and defines the key structural elasticity variables. In Section 2 we analyze the tax incidence problem with a continuum of wages and nonlinear income taxes, focusing on the impact of tax reforms on individual variables (labor supply, wages, utilities). In Section 2 we derive the incidence of taxes on aggregate variables (government revenue, social welfare). In Section 4 we explore various generalizations of our baseline environment. Finally, in Section 5 we derive optimal taxes in general equilibrium. The proofs of all the formulas and results of this paper are gathered in the Appendix.

## 1 The baseline environment

### 1.1 Preferences, technology and equilibrium

#### Individual behavior

In the simplest version of our model, individuals have a quasilinear utility function over consumption  $c$  and labor supply  $l$  given by  $c - v(l)$ , where the disutility of labor  $v : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  is twice continuously differentiable, strictly increasing, and strictly convex.

There is a continuum of skills  $\theta \in \Theta = [\underline{\theta}, \bar{\theta}] \subset \mathbb{R}_+$ , distributed according to the continuous p.d.f.  $f(\cdot)$  and c.d.f.  $F(\cdot)$ . An individual with skill  $\theta$  earns a wage  $w(\theta)$  that she takes as given. She chooses her labor supply  $l(\theta)$  and earns taxable income  $y(\theta) = w(\theta)l(\theta)$ . Her consumption is equal to  $y(\theta) - T(y(\theta))$ , where  $T : \mathbb{R}_+ \rightarrow \mathbb{R}$

is a twice continuously differentiable income tax schedule. The optimal labor supply choice  $l(\theta)$  satisfies the first-order condition of the utility-maximization problem:<sup>5</sup>

$$v'(l(\theta)) = [1 - T'(w(\theta)l(\theta))]w(\theta). \quad (1)$$

We denote by  $\mathcal{U}(\theta)$  the agent's indirect utility and by  $L(\theta) \equiv l(\theta)f(\theta)$  the total amount of labor supplied by individuals of type  $\theta$ .

**Remark.** Without loss of generality, we can assume that wages  $w(\theta)$  are strictly increasing in the types  $\theta$ . In other words, one can interpret each  $\theta$  as a given skill involved in production, and order these skills by their wage, given the tax schedule  $T$ . In particular, we can normalize  $\Theta = [0, 1]$  with a distribution  $f(\theta)$  that is uniform, so that  $\theta$  indexes the agent's percentile in the wage distribution.<sup>6</sup> We show in Appendix A.2.3 that, by the Spence-Mirrlees condition, the pre-tax income function  $\theta \mapsto y(\theta)$  is then also strictly increasing. There is therefore a one-to-one map between skills  $\theta$  (or wages  $w(\theta)$ ) and pre-tax incomes  $y(\theta)$ .<sup>7</sup> We denote by  $f_Y(y(\theta)) = (y'(\theta))^{-1}f(\theta)$  the density of incomes and by  $F_Y(\cdot)$  the corresponding c.d.f. in the economy.

## Production and wages

There is a continuum of mass 1 of identical firms that produce output using the labor of all skills  $\theta$ . We represent the continuum of labor inputs  $\mathcal{L} = \{L(\theta)\}_{\theta \in \Theta}$  as a finite, non-negative measure on the compact metric space  $(\Theta, \mathcal{B}(\Theta))$ .<sup>8</sup> We then define the production function as  $\mathcal{F}(\mathcal{L}) = \mathcal{F}(\{L(\theta)\}_{\theta \in \Theta})$  and write the representative firm's

---

<sup>5</sup>Note that the dependence of labor supply on the initial tax schedule  $T$  is left implicit for simplicity. Whenever necessary, we denote the solution to (1) by  $l(\theta; T)$ .

<sup>6</sup>In Section 4.2, we relax the assumption that all agents assigned to a given skill  $\theta$  earn the same wage  $w(\theta)$ .

<sup>7</sup>If the tax system changes, generally, the ordering of wages may change (see Section 3.4 for details). Our analysis does not require that the initial ordering remains unaffected by the tax reforms we consider.

<sup>8</sup>Thus, for any Borelian set  $B \in \mathcal{B}(\Theta)$  (e.g., an interval in  $\Theta$ ),  $\mathcal{L}(B)$  is the total amount of labor supplied by individuals with productivity  $\theta \in B$ . This construction follows Hart (1979) and Fradera (1986).

profit-maximization problem given the wage schedule  $\{w(\theta)\}_{\theta \in \Theta}$  as<sup>9</sup>

$$\max_{\mathcal{L}} \left[ \mathcal{F}(\mathcal{L}) - \int_{\Theta} w(\theta) L(\theta) d\theta \right].$$

We assume that the production function  $\mathcal{F}$  has constant returns to scale. In equilibrium, firms earn no profits and the wage  $w(\theta)$  is equal to the marginal productivity of type- $\theta$  labor, that is,<sup>10</sup>

$$w(\theta) = \frac{\partial}{\partial L(\theta)} \mathcal{F}(\mathcal{L}). \quad (2)$$

A commonly used production function, which we use for some of our results, is the CES technology.

**Example 1. (CES technology.)** The production function has a constant elasticity of substitution (CES) if

$$\mathcal{F}(\mathcal{L}) = \left[ \int_{\Theta} a(\theta) (L(\theta))^{\frac{\sigma-1}{\sigma}} d\theta \right]^{\frac{\sigma}{\sigma-1}}, \quad (3)$$

for some constant  $\sigma \in [0, \infty)$  and parameters  $a(\theta) \in \mathbb{R}_+$ . The wage schedule is given by  $w(\theta) = a(\theta) (L(\theta) / \mathcal{F}(\mathcal{L}))^{-1/\sigma}$ . The cases  $\sigma = 1$  and  $\sigma = 0$  correspond respectively to the Cobb-Douglas and Leontieff production functions, and  $\sigma = \infty$  implies that wages are exogenous.

## Government

Government revenue is given by

$$\mathcal{R} = \int_{\Theta} T(y(\theta)) f(\theta) d\theta. \quad (4)$$

---

<sup>9</sup>Note that an alternative interpretation of our framework is that different types of workers produce different types of goods that are imperfect substitutes in household consumption. See, e.g., [Acemoglu and Autor \(2011\)](#).

<sup>10</sup>Equation (2) ignores the technical difficulties involved with the fact that  $\mathcal{F}$  has continuum of arguments. Formally,  $w(\theta)$  is equal to the Gateaux derivative of the production function when the labor effort schedule  $\mathcal{L}$  is perturbed in the direction of the Dirac measure at  $\theta$ , that is,  $w(\theta) = \lim_{\mu \rightarrow 0} \frac{1}{\mu} \{ \mathcal{F}(\mathcal{L} + \mu \delta_{\theta}) - \mathcal{F}(\mathcal{L}) \}$ .

We denote the local rate of progressivity of the tax schedule  $T$  by

$$p(y) \equiv -\frac{\partial \ln [1 - T'(y)]}{\partial \ln y} = \frac{yT''(y)}{1 - T'(y)}.$$

It is equal to (minus) the elasticity of the retention rate  $1 - T'(y)$  with respect to income  $y$ .

**Example 2. (CRP tax schedule.)** The tax schedule has a constant rate of progressivity (CRP) if  $T(y) = y - \frac{1-m}{1-p}y^{1-p}$ , for  $p < 1$ .<sup>11</sup> This tax schedule is linear (resp., progressive, regressive), i.e., the marginal tax rates  $T'(y)$  and the average tax rates  $T(y)/y$  are constant (resp., increasing, decreasing), if  $p = 0$  (resp.,  $p > 0$ ,  $p < 0$ ).

The government evaluates social welfare by means of a Bergson-Samuelson concave social welfare function  $G : \mathbb{R} \rightarrow \mathbb{R}$ . Denote by  $\lambda$  the marginal value of public funds.<sup>12</sup> We then define social welfare, expressed in monetary units, by:

$$\mathcal{W} \equiv \mathcal{R} + \frac{1}{\lambda} \int_{\Theta} G[y(\theta) - T(y(\theta)) - v(l(\theta))] f(\theta) d\theta. \quad (5)$$

We denote by  $g(\theta)$ , or equivalently  $g(y(\theta))$ , the social marginal welfare weight<sup>13</sup> associated with individuals of type  $\theta$  as

$$g(\theta) = \frac{1}{\lambda} G'[y(\theta) - T(y(\theta)) - v(l(\theta))].$$

The weight  $g(\theta)$  is the social value of giving one additional unit of consumption to individuals with type  $\theta$ , relative to distributing it uniformly in the whole population.

## 1.2 Elasticity concepts

We now introduce notations for the elasticities that we use in our incidence and optimal tax formulas. All of them are structural parameters that are known in closed-form.

<sup>11</sup>See, e.g., [Musgrave and Thin \(1948\)](#); [Bénabou \(2002\)](#); [Heathcote, Storesletten, and Violante \(2016\)](#).

<sup>12</sup>The marginal value of public funds  $\lambda$  is determined by imposing that all of the perturbations of the tax system that we consider are revenue-neutral, that is, by redistributing (or taxing) lump-sum any excess revenue. Thus  $\lambda$  is the social value of distributing an additional unit of revenue uniformly in the entire population. In the optimal taxation problem that we study in Sections 3.4 and 5,  $\lambda$  is naturally the Lagrange multiplier on the government budget constraint.

<sup>13</sup>See, e.g., [Saez and Stantcheva \(2016\)](#).

## Labor supply elasticities

We define the elasticity of labor supply of skill type  $\theta$  with respect to the retention rate  $r(\theta) \equiv 1 - T'(y(\theta))$  as<sup>14</sup>

$$\varepsilon_r^S(\theta) = \frac{\partial \ln l(\theta)}{\partial \ln r(\theta)} = \frac{e(\theta)}{1 + p(y(\theta))e(\theta)}, \quad (6)$$

where  $e(\theta) = \frac{v'(l(\theta))}{l(\theta)v''(l(\theta))}$  and the superscript  $S$  stands for (labor) supply. This elasticity differs from the more standard variable  $e(\theta)$  as it accounts for the fact that if the tax schedule is nonlinear, a change in individual labor supply  $l(\theta)$  induces endogenously a change in the marginal tax rate  $T'(y(\theta))$  (given by the rate of progressivity  $p(y(\theta))$  of the tax schedule), and hence a further labor supply adjustment  $e(\theta)$ . Solving for this fixed point leads to the correction term  $p(y(\theta))e(\theta)$  in the denominator of (6).<sup>15</sup> For further details see Appendix A.1.

We also define the elasticity of labor supply of type  $\theta$  with respect to the wage  $w(\theta)$  as

$$\varepsilon_w^S(\theta) = \frac{\partial \ln l(\theta)}{\partial \ln w(\theta)} = (1 - p(y(\theta))) \varepsilon_r^S(\theta). \quad (7)$$

This elasticity differs from (6) because a wage change affects  $(1 - T'(y(\theta)))w(\theta)$  both directly as in the case of an exogenous perturbation in the retention rate, and indirectly through its effect on the marginal tax rate  $T'(w(\theta)l(\theta))$  if the tax schedule is nonlinear. The latter is accounted for by the correction  $p(y(\theta))\varepsilon_r^S(\theta)$  in (7).

## Cross-wage and own-wage elasticities

We define the structural cross-elasticity of the wage of type  $\theta'$ ,  $w(\theta')$ , with respect to the labor supply of type  $\theta \neq \theta'$ ,  $L(\theta)$ , as:<sup>16,17</sup>

$$\gamma(\theta', \theta) = \frac{\partial \ln w(\theta')}{\partial \ln L(\theta)} = \frac{L(\theta) \mathcal{F}_{\theta', \theta}''}{\mathcal{F}_{\theta'}'}, \quad (8)$$

<sup>14</sup>Since there is a one-to-one map between types  $\theta$  and incomes  $y(\theta)$ , we can write interchangeably  $\varepsilon_r^S(\theta)$  or  $\varepsilon_r^S(y(\theta))$ , and similarly for the elasticities  $\varepsilon_w^S(\theta)$  and  $\alpha(\theta)$  defined below.

<sup>15</sup>See also [Jacquet and Lehmann \(2017\)](#) and [Scheuer and Werning \(2017\)](#).

<sup>16</sup>We assume that  $\theta' \mapsto \gamma(\theta', \theta)$  is a continuous map on  $\Theta \setminus \{\theta\}$ .

<sup>17</sup>The natural change of variables between types  $\theta$  and incomes  $y(\theta)$  for the cross-wage elasticities reads  $\gamma(y(\theta_1), y(\theta_2)) = (y'(\theta_2))^{-1} \gamma(\theta_1, \theta_2)$ . See Appendix A.2 for details.

where  $\mathcal{F}'_{\theta'}$  and  $\mathcal{F}''_{\theta',\theta}$  denote the first and second partial derivatives of the production function  $\mathcal{F}$  with respect to the labor inputs of types  $\theta'$  and  $\theta$ . The structural cross-wage elasticity between two skills  $(\theta', \theta)$  with  $\theta' \neq \theta$  is non-zero if they are imperfect substitutes. Although this is not necessary for our theoretical analysis,<sup>18</sup> we usually assume for simplicity in our discussions below that these elasticities are positive (as is always the case, for instance, if the production function is CES), i.e., that any two skills  $\theta' \neq \theta$  are Edgeworth complements in production.

The function  $\theta' \mapsto \frac{\partial \ln w(\theta')}{\partial \ln L(\theta)}$  is generally discontinuous at  $\theta' = \theta$ , i.e., when we consider the impact of the labor of type- $\theta$  agents on their own wage. We call own-wage elasticity  $\alpha(\theta)$  the difference between  $\frac{\partial \ln w(\theta)}{\partial \ln L(\theta)}$  and  $\lim_{\theta' \rightarrow \theta} \gamma(\theta', \theta)$ . By subtracting from  $\frac{\partial \ln w(\theta)}{\partial \ln L(\theta)}$  the complementarity  $\lim_{\theta' \rightarrow \theta} \gamma(\theta', \theta)$  between the skill  $\theta$  and its neighboring skills  $\theta' \approx \theta$ , the variable  $\alpha(\theta)$  captures the impact of the labor effort  $L(\theta)$  on the wage  $w(\theta)$  arising purely from the fact that the marginal productivity of skill  $\theta$  is a non-constant (say, decreasing) function of the aggregate labor of its own type. Formally, we define

$$\alpha(\theta) = - \left[ \frac{\partial \ln w(\theta)}{\partial \ln L(\theta)} - \lim_{\theta' \rightarrow \theta} \frac{\partial \ln w(\theta')}{\partial \ln L(\theta)} \right] = - \left[ \frac{L(\theta) \mathcal{F}''_{\theta,\theta}}{\mathcal{F}'_{\theta}} - \lim_{\theta' \rightarrow \theta} \frac{L(\theta) \mathcal{F}''_{\theta',\theta}}{\mathcal{F}'_{\theta'}} \right]. \quad (9)$$

This elasticity is positive if the marginal productivity of the labor input  $L(\theta)$  is decreasing. Although this is not required for our theoretical analysis, we assume for simplicity in our discussions below that this is the case for all skills  $\theta$ .<sup>19</sup>

**Example. (CES technology.)** In the case of a CES production function, the cross-wage elasticities are given by  $\gamma(\theta', \theta) = \frac{1}{\sigma} a(\theta) (L(\theta) / \mathcal{F}(\mathcal{L}))^{\frac{\sigma-1}{\sigma}}$  and the own-wage elasticities are given by  $\alpha(\theta) = 1/\sigma$  for all  $\theta', \theta$ . Note that  $\alpha(\theta) > 0$  is constant and that  $\gamma(\theta', \theta) > 0$  does not depend on  $\theta'$ , implying that a change in the labor supply of type  $\theta$  has the same effect, in percentage terms, on the wage of every type  $\theta' \neq \theta$ . Denoting by  $\sigma(\theta', \theta) = - \left[ \frac{\partial \ln(w(\theta')/w(\theta))}{\partial \ln(L(\theta')/L(\theta))} \right]^{-1}$  the elasticity of substitution between any two labor inputs, we have  $\sigma(\theta', \theta) = \sigma$  for all  $(\theta', \theta) \in \Theta^2$ .

---

<sup>18</sup>For instance, a Translog production function can allow for negative cross-wage elasticities. Negative cross-input effects arise, for instance, in the capital-skill complementarity literature (see [Krusell et al. \(2000\)](#)).

<sup>19</sup>The variable  $\alpha(\theta)$  can also be interpreted as the inverse of the partial-equilibrium elasticity  $1/\varepsilon_w^D(\theta)$  of labor demand of type  $\theta$  with respect to the wage  $w(\theta)$ .

## Elasticities of equilibrium labor

We define the elasticities of labor of type  $\theta$  in partial equilibrium, i.e., keeping the prices  $w(\theta')$  and quantities  $L(\theta')$  constant in all other “markets”  $\theta' \neq \theta$ , by

$$\varepsilon_r(\theta) = \frac{\varepsilon_r^S(\theta)}{1 + \alpha(\theta)\varepsilon_w^S(\theta)}, \text{ and } \varepsilon_w(\theta) = \frac{\varepsilon_w^S(\theta)}{1 + \alpha(\theta)\varepsilon_w^S(\theta)}. \quad (10)$$

Intuitively, a percentage increase in the labor supply of type- $\theta$  individuals by  $\varepsilon_r^S(\theta)$  or  $\varepsilon_w^S(\theta)$ , caused by an increase in their retention rate or their wage, lowers their own wage (due to the decreasing marginal productivity) by  $\alpha(\theta)$ , which in turn dampens the initial increase in their labor supply by  $\alpha(\theta)\varepsilon_w^S(\theta)$ . Solving for the fixed point leads to expressions (10).<sup>20</sup>

## 2 Tax incidence

Consider a given initial, potentially suboptimal, tax schedule  $T$ , e.g., the U.S. tax code. In this section we derive closed-form formulas for the first-order effects of arbitrary local perturbations (“tax reforms”) of this tax schedule on individual labor supplies, wages and indirect utilities.

### 2.1 Incidence of tax reforms on labor supply

As in the case of exogenous wages (Saez, 2001), analyzing the incidence of tax reforms relies crucially on solving for each individual’s change in labor supply in terms of behavioral elasticities. This problem is, however, much more involved in general equilibrium. In the setting with exogenous wages, in the absence of income effects, a change in the tax rate of a given individual, say  $\theta$ , induces only a change in the labor effort of that agent (measured by the elasticity (6)). In the general equilibrium setting, instead, this labor supply response of type  $\theta$  affects the wage, and hence the labor supply, of every other skill  $\theta' \neq \theta$ . This in turn feeds back into the wage of  $\theta$ , which further impacts labor supplies, and so on. Representing the total effect of

---

<sup>20</sup>Interpreting  $\alpha(\theta) = 1/\varepsilon_w^D(\theta)$  as the inverse of the labor demand elasticity with respect to the wage, we can write  $\frac{1}{\varepsilon_w(\theta)} = \frac{1}{\varepsilon_w^S(\theta)} + \frac{1}{\varepsilon_w^D(\theta)}$ . As expected from the Ramsey tax literature (see, e.g., Stiglitz (2015)), this sum of inverse elasticities of labor supply and labor demand is an important variable for our tax incidence analysis.

this infinite sequence caused by arbitrarily non-linear tax reforms is thus a priori a complex task.<sup>21</sup>

The key step towards the general characterization of the economic incidence of taxes, and our first main theoretical contribution, consists of showing that this problem can be mathematically formulated as solving an integral equation (Lemma 1).<sup>22</sup> We can thus apply the tools and results of the theory of integral equations to solve in closed-form for the labor supply adjustments in general equilibrium (Proposition 1). The incidence on wages and utilities is then straightforward to derive (Sections 2.2 and 2.3).

Formally, consider an arbitrary non-linear reform of the initial tax schedule  $T(\cdot)$ . This tax reform can be represented by a continuously differentiable function  $\tau(\cdot)$  on  $\mathbb{R}_+$ , so that the perturbed tax schedule is  $T(\cdot) + \mu\tau(\cdot)$ , where  $\mu \in \mathbb{R}$  parametrizes the size of the reform. Our aim is to compute the first-order effect of this perturbation on individual labor supply (i.e., the solution to the first-order condition (1)), when the magnitude of the tax change is small, i.e., as  $\mu \rightarrow 0$ . This is formally expressed by the Gateaux derivative of the labor supply functional  $T \mapsto l(\theta; T)$  in the direction  $\tau$ , that is,<sup>23</sup>

$$dl(\theta) \equiv \lim_{\mu \rightarrow 0} \frac{1}{\mu} [l(\theta; T + \mu\tau) - l(\theta; T)], \text{ and } \hat{l}(\theta) \equiv \frac{dl(\theta)}{l(\theta; T)}.$$

The variable  $dl(\theta)$  (resp.,  $\hat{l}(\theta)$ ) gives the absolute (resp., percentage) change in the labor supply of type  $\theta$  in response to the tax reform  $\tau$ , taking into account all of the general equilibrium effects induced by the endogeneity of wages. We define analogously the absolute changes in individual wages  $dw(\theta)$ , indirect utilities  $du(\theta)$ , government revenue  $d\mathcal{R}$  and social welfare  $d\mathcal{W}$ , and the corresponding percentage changes  $\hat{w}(\theta)$ ,  $\hat{u}(\theta)$ ,  $\hat{\mathcal{R}}$ ,  $\hat{\mathcal{W}}$ .

---

<sup>21</sup>We can always define, for each specific tax reform one might consider implementing, a “policy elasticity” (as in, e.g., [Hendren \(2015\)](#), [Piketty and Saez \(2013\)](#)), equal to each individual’s total labor supply response to the corresponding reform. However the key challenge of the incidence problem consists of expressing this total labor supply response in terms of the structural elasticity parameters introduced in Section 1.2.

<sup>22</sup>The general theory of linear integral equations is exposed in, e.g., [Tricomi \(1985\)](#), [Kress \(2014\)](#), and, as a concise introduction, in [Zemyan \(2012\)](#). Moreover, closed-form solutions can be derived in many cases (see [Polyanin and Manzhirov \(2008\)](#)). Finally, numerical techniques are widely available and can be easily implemented (see, e.g., [Press \(2007\)](#) and Section 2.6 in [Zemyan \(2012\)](#)), leading to straightforward quantitative evaluations of the incidence of arbitrary tax reforms (see Section 3.5).

<sup>23</sup>The notations  $dl(\theta)$  and  $\hat{l}(\theta)$  ignore for simplicity the dependence of these derivatives on the initial tax schedule  $T$  and on the tax reform  $\tau$ .

**Lemma 1.** *The incidence of a tax reform  $\tau$  of the initial tax schedule  $T$  on individual labor supplies,  $\hat{l}(\cdot)$ , is the solution to the integral equation*

$$\hat{l}(\theta) = -\varepsilon_r(\theta) \frac{\tau'(y(\theta))}{1-T'(y(\theta))} + \varepsilon_w(\theta) \int_{\Theta} \gamma(\theta, \theta') \hat{l}(\theta') d\theta', \quad (11)$$

for all  $\theta \in \Theta$ .

*Proof.* See Appendix B.1.1. □

Formula (11) is a linear Fredholm integral equation of the second kind with kernel  $\varepsilon_w(\theta) \gamma(\theta, \theta')$ . Its unknown, which appears under the integral sign, is the function  $\theta \mapsto \hat{l}(\theta)$ . We start by providing the interpretation of this equation.

Due to the reform, the retention rate  $r(\theta) = 1 - T'(y(\theta))$  of individual  $\theta$  changes, in percentage terms, by  $\hat{r}(\theta) = -\frac{\tau'(y(\theta))}{1-T'(y(\theta))}$ . By construction of the elasticity (10), this tax reform induces a direct percentage change in labor effort  $l(\theta)$  equal to  $\varepsilon_r(\theta) \times \hat{r}(\theta)$ . This is the expression we would obtain in partial equilibrium, i.e., in the absence of cross-wage effects. It resembles the expression one obtains assuming fixed wages<sup>24</sup> with one difference: since the marginal product of labor is decreasing, the initial labor supply change (say, decrease) due to the tax reform causes an own-wage increase, which in turn tends to raise labor supply and dampen the initial elasticity response. That is, the relevant elasticity is now  $\varepsilon_r(\theta)$  rather than  $\varepsilon_r^S(\theta)$ , with  $\varepsilon_r(\theta) < \varepsilon_r^S(\theta)$ .

In general equilibrium, the labor supply of type  $\theta$  is also impacted indirectly by the change in all other individuals' labor supplies, due to the skill complementarities in production. Specifically, the change in labor supply of each type  $\theta'$ ,  $\hat{l}(\theta')$ , triggers a change in the wage of type  $\theta$  equal to  $\gamma(\theta, \theta') \hat{l}(\theta')$ , and thus a further adjustment in her labor supply equal to  $\varepsilon_w(\theta) \gamma(\theta, \theta') \hat{l}(\theta')$ . Summing these effects over skills  $\theta' \in \Theta$  leads to formula (11).

We now characterize the solution to the integral equation (11). At this point it is also noteworthy that (11) can easily be solved numerically.

**Proposition 1.** *Assume that the condition  $\int_{\Theta^2} |\varepsilon_w(\theta) \gamma(\theta, \theta')|^2 d\theta d\theta' < 1$  holds.<sup>25</sup>*

<sup>24</sup>See, e.g., p. 217 in Saez (2001).

<sup>25</sup>This technical condition ensures that the infinite series (13) converges. We provide below sufficient conditions on primitives such that this convergence is ensured. In more general cases it can be easily verified numerically (see Section 3.5; all of our numerical simulations satisfied this restriction). Finally, when it is not satisfied, we can more generally express the solution to (11) with a representation similar to (12) but with a more complex resolvent (see Section 2.4 in Zemyan (2012)).

The unique solution to the integral equation (11) is given by

$$\hat{l}(\theta) = -\varepsilon_r(\theta) \frac{\tau'(y(\theta))}{1-T'(y(\theta))} - \varepsilon_w(\theta) \int_{\Theta} \Gamma(\theta, \theta') \varepsilon_r(\theta') \frac{\tau'(y(\theta'))}{1-T'(y(\theta'))} d\theta', \quad (12)$$

where for all  $(\theta, \theta') \in \Theta^2$ , the resolvent  $\Gamma(\theta, \theta')$  is defined by

$$\Gamma(\theta, \theta') \equiv \sum_{n=1}^{\infty} \Gamma_n(\theta, \theta'), \quad (13)$$

with  $\Gamma_1(\theta, \theta') = \gamma(\theta, \theta')$  and for all  $n \geq 2$ ,

$$\Gamma_n(\theta, \theta') = \int_{\Theta} \Gamma_{n-1}(\theta, \theta'') \varepsilon_w(\theta'') \gamma(\theta'', \theta') d\theta''.$$

We call  $\Gamma(\theta, \theta')$  the general-equilibrium (GE) cross-wage elasticity (as opposed to the structural cross-wage elasticity  $\gamma(\theta, \theta')$ ).

*Proof.* See Appendix B.1.2. □

The mathematical representation (12) of the solution to the integral equation (11) has a clear economic interpretation. The first term on the right hand side of (12),  $-\varepsilon_r(\theta) \frac{\tau'(y(\theta))}{1-T'(y(\theta))}$ , is the partial-equilibrium effect of the reform on labor supply  $l(\theta)$ , as already described in equation (11). The second (integral) term accounts for the cross-wage effects in general equilibrium. Note that this integral term has the same structure (and interpretation) as in formula (11), except that: (i) the unknown labor supply changes  $\hat{l}(\theta')$ , that had to be solved for, are now replaced by their partial-equilibrium values  $-\varepsilon_r(\theta') \frac{\tau'(y(\theta'))}{1-T'(y(\theta'))}$ ; and (ii) the structural cross-wage elasticity  $\gamma(\theta, \theta')$  is replaced by the GE cross-wage elasticity  $\Gamma(\theta, \theta')$ . As we describe in the next paragraph, this elasticity, defined by the series (13), expresses the total effect of the labor supply of type  $\theta'$  on the wage of type  $\theta$ , i.e., it accounts for the infinite sequence of general equilibrium adjustments induced by the complementarities in production.

We now interpret the definition (13) of the GE cross-wage elasticity  $\Gamma(\theta, \theta')$ . The first iterated kernel ( $n = 1$ ) in the series (13) is simply  $\Gamma_1(\theta, \theta') = \gamma(\theta, \theta')$ . It thus accounts for the impact of the labor supply of type  $\theta'$  on the wage of type  $\theta$  through direct cross-wage effects. The second iterated kernel ( $n = 2$ ) in (13) accounts for the impact of the labor supply of  $\theta'$  on the wage of  $\theta$ , indirectly through the behavior of

third parties  $\theta''$ . This term reads

$$\Gamma_2(\theta, \theta') = \int_{\Theta} \gamma(\theta, \theta'') \varepsilon_w(\theta'') \gamma(\theta'', \theta') d\theta''. \quad (14)$$

For any  $\theta'$ , a percentage change in the labor supply of  $\theta'$  induces a percentage change in the wage of any other type  $\theta''$  by  $\gamma(\theta'', \theta')$  (by definition (8)), and hence a percentage change in the labor supply of  $\theta''$  given by  $\varepsilon_w(\theta'') \gamma(\theta'', \theta')$  (by definition (10)). This in turn affects the wage of type  $\theta$  by the amount  $\gamma(\theta, \theta'') \varepsilon_w(\theta'') \gamma(\theta'', \theta')$ . Summing over all intermediate types  $\theta''$  leads to expression (14). An inductive reasoning shows similarly that the terms  $n \geq 3$  in the resolvent series (13) account for the impact of the labor supply of  $\theta'$  on the wage of  $\theta$  through  $n$  successive stages of cross-wage effects, e.g., for  $n = 3$ ,  $\theta' \rightarrow \theta'' \rightarrow \theta''' \rightarrow \theta$ .

**Example. (CES technology.)** Suppose that the production function is CES. In this case, the GE cross-wage elasticities are given by:

$$\Gamma(\theta, \theta') = \frac{\gamma(\theta, \theta')}{1 - \frac{1}{\sigma_{\mathbb{E}y}} \int_{\mathbb{R}_+} y \varepsilon_w(y) f_Y(y) dy}. \quad (15)$$

That is, the total impact  $\Gamma(\theta, \theta')$  of a change in the labor supply of type  $\theta'$  on the wage of type  $\theta$  is proportional to the direct (structural) effect  $\gamma(\theta, \theta')$ . This is because each round of cross-wage general equilibrium effects, i.e., each term in the resolvent series (13), is a fraction of the first round. This in turn follows from the fact that, with a CES technology, the cross-wage elasticity  $\gamma(\theta, \theta')$  depends only on  $\theta'$  and is independent of  $\theta$ , that is, a change in the aggregate labor supply of type  $\theta'$  induces the same percentage adjustment in the wage of every skill  $\theta \neq \theta'$ . Mathematically, the kernel  $\varepsilon_w(\theta) \gamma(\theta, \theta')$  of the integral equation (11) is then multiplicatively separable between  $\theta$  and  $\theta'$ , which makes it straightforward to solve (see Appendix B.1.3 for details).

### Sufficient conditions on primitives ensuring convergence of the resolvent

Suppose that the production function is CES with parameter  $\sigma > 0$ , that the initial tax schedule is CRP with parameter  $p < 1$ , and that the disutility of labor is isoelastic with parameter  $e > 0$ . We show in Appendix A.2.1 that we have in this case  $\frac{1}{\sigma_{\mathbb{E}y}} \mathbb{E}[y \varepsilon_w(y)] < 1$  so that, by formula (15),  $\Gamma(\theta, \theta') < \infty$ . The convergence of the resolvent series (13) is thus satisfied.

## 2.2 Incidence of tax reforms on wages

Once the labor supply response  $\hat{l}(\theta)$  is characterized in closed-form (Proposition 1), we can easily derive the incidence of an arbitrary tax reform  $\tau$  on individual wages. We show in Appendix B.2.1 that, for all  $\theta \in \Theta$ ,

$$\hat{w}(\theta) = \frac{1}{\varepsilon_w^S(\theta)} \left[ \varepsilon_r^S(\theta) \frac{\tau'(y(\theta))}{1 - T'(y(\theta))} + \hat{l}(\theta) \right]. \quad (16)$$

This equation expresses the changes in individual wages due to the tax reform  $\tau$ , as a function of the labor supply changes given by (12). Its interpretation is straightforward. Multiplying both sides of (16) by  $\varepsilon_w^S(\theta)$  simply gives the adjustment of type- $\theta$  labor supply,  $\hat{l}(\theta)$ , as the sum of its response in the case of exogenous wages,  $-\varepsilon_r^S \frac{\tau'}{1-T'}$ , and the general equilibrium effect induced by the wage change,  $\varepsilon_w^S \times \hat{w}$ .

## 2.3 Incidence of tax reforms on individual welfare

Finally, we can easily derive the incidence of an arbitrary tax reform  $\tau$  of the initial tax schedule  $T$  on individual indirect utilities. We show in Appendix B.2.2 that

$$du(\theta) = -\tau(y(\theta)) + (1 - T'(y(\theta)))y(\theta)\hat{w}(\theta), \quad (17)$$

for all  $\theta \in \Theta$ . The first term on the right hand side of equation (17),  $-\tau(y(\theta))$ , is due to the fact that a higher tax payment makes the individual poorer and hence reduces her utility. The second term accounts for the change in net income due to the wage adjustment  $\hat{w}(\theta)$ .

If wages were exogenous (i.e.,  $\hat{w}(\theta) = 0$  in (17)), the welfare of agent  $\theta$  would respond one-for-one to the change in her total tax bill,  $\tau(y(\theta))$ . In particular, the change in the marginal tax rate that the reform induces,  $\tau'(y(\theta))$ , would not affect her utility. This is a direct consequence of the envelope theorem: the marginal tax rate affects utility only to the extent that it leads to adjustments in labor supply (equation (1)); but labor supply is initially chosen optimally, hence a change in the marginal tax rate has only a second-order effect on welfare (conditional on a given absolute tax change).

In general equilibrium, however, this is no longer true, because labor supply changes also imply movements in wages, which have first-order effects on welfare. Therefore the change in the marginal tax rate impacts individual utilities, even con-

ditional on a given absolute tax change. Specifically, we show:

**Corollary 1.** *Assume that the cross-wage elasticities satisfy  $\gamma(\theta', \theta) \geq 0$  for all  $\theta', \theta$ . For a given absolute tax change  $\tau(y(\theta))$  at income  $y(\theta)$ , an increase in the marginal tax rate  $\tau'(y(\theta)) > 0$  raises the utility of agents with type  $\theta$  and lowers that of all other agents, i.e.,  $du(\theta) > 0$ , and  $du(\theta') < 0$  for all  $\theta' \neq \theta$ .*

*Proof.* See Appendix B.2.3. □

Intuitively, a higher marginal tax rate for individuals of type  $\theta$  makes them work less, because of the standard substitution effect, and earn more per hour worked, because of the decreasing marginal product of labor, which makes them better off. If the cross-wage elasticities are positive (as is the case, for instance, if the production function is CES), then the wages, and hence utilities, of all other types go down as a consequence of the lower labor effort of type  $\theta$ .<sup>26</sup> As we show in Section 5, this insight generalizes that of Stiglitz (1982) and is a crucial determinant of the structure of optimal marginal tax rates.

### 3 Aggregate effects of tax reforms

Having derived the change in the equilibrium amount of labor (12) and the change in wages (16) in response to a tax reform  $\tau$ , the incidence on government revenue  $\mathcal{R}$ , defined in (4), directly follows:

$$d\mathcal{R} = \int_{\mathbb{R}_+} \tau(y) f_Y(y) dy + \int_{\mathbb{R}_+} T'(y) [\hat{l}(y) + \hat{w}(y)] y f_Y(y) dy, \quad (18)$$

where  $\hat{l}(y) \equiv \hat{l}(\theta_y)$  and  $\hat{w}(y) \equiv \hat{w}(\theta_y)$  are the labor supply and the wage changes of agents with income  $y$  (and type  $\theta_y$ ). The first term on the right hand side of (18) is the mechanical effect of the tax reform  $\tau(\cdot)$ , i.e., the change in government revenue if the individual behavior and her wage remained constant. The second term is the behavioral effect of the reform. The labor supply and wage adjustments  $\hat{l}(y)$  and  $\hat{w}(y)$  both induce a change in government revenue proportional to the marginal tax

---

<sup>26</sup>Note that an increase in the marginal tax rate at income  $y(\theta)$  implies in particular that individuals with skill  $\theta' > \theta$  are made worse off for two separate reasons: (i) their total tax bill is now mechanically higher, since the marginal tax rate on income  $y(\theta)$  has increased; (ii) their wage is lower, since the labor supply of agents  $\theta$  is distorted downward.

rate  $T'(y)$ . Summing these effects over all individuals using the density of incomes  $f_Y(y)$  yields (18).

In this section we derive the economic implications of formula (18). Sections 3.1 and 3.2 contain useful preliminary steps. Section 3.3 contains our main results. Section 3.4 extends the results of this section to a general social welfare objective and derives as a by-product a characterization of the optimal tax schedule.

## Elementary tax reforms

From now on, we focus on a specific class of “elementary” tax reforms, represented by the step function  $\tau(y) = (1 - F_Y(y^*))^{-1} \mathbb{I}_{\{y \geq y^*\}}$  for a given income level  $y^*$ .<sup>27</sup> That is, the total tax liability increases by the constant amount  $(1 - F_Y(y^*))^{-1}$  above income  $y^*$ , and the marginal tax rates are perturbed by the Dirac delta function at  $y^*$ ,  $\tau'(y) = (1 - F_Y(y^*))^{-1} \delta_{y^*}(y)$ . Intuitively, this reform consists of raising the marginal tax rate at only one income level  $y^* \in \mathbb{R}_+$ , which implies a uniform lump-sum increase in the total tax payment of agents with income  $y > y^*$ . The normalization by  $(1 - F_Y(y^*))^{-1}$  implies that the statutory increase in government revenue due to the reform (i.e., the first term in the right hand side of (18), which ignores the agents’ behavioral responses) is equal to \$1.<sup>28</sup> We denote by  $d\mathcal{R}(y^*)$  the total effect on government revenue (18) of the elementary tax reform at income  $y^*$ .

Finally, note that any other tax reform can be expressed as a weighted sum of such income-specific perturbations.<sup>29</sup> Specifically, the effect of an arbitrary tax reform  $\tau$  on government revenue is given by

$$d\mathcal{R}(\tau) = \int_{\mathbb{R}_+} d\mathcal{R}(y^*) (1 - F_Y(y^*)) \tau'(y^*) dy^*. \quad (19)$$

The focus on the elementary tax reforms at arbitrary income levels is thus without

---

<sup>27</sup>Note that the function  $\mathbb{I}_{\{y \geq y^*\}}$  is not differentiable. We show in Appendix C.1 that we can nevertheless use our theory to analyze this reform by applying (12) to a sequence of smooth perturbations  $\{\tau'_n(y)\}_{n \geq 1}$  that converges to the Dirac delta function  $\delta_{y^*}(y)$ . This notation simplifies the exposition and is made only for convenience. All of our formulas can be easily written for any smooth tax reform  $\tau$  rather than the step functions (see formula (19)).

<sup>28</sup>Heuristically, consider a perturbation that raises the marginal tax rate by  $dT'$  on a small income interval  $[y^* - dy, y^*]$ , so that the total tax payment above income  $y^*$  raises by the amount  $dT' \times dy$  equal to, say, \$1. This class of tax reforms has been introduced by Saez (2001). Then shrink the size of the income interval on which the tax rate is increased, i.e.  $dy \rightarrow 0$ , while keeping the increase in the tax payment above  $y^*$  fixed at \$1. The limit of the marginal tax rate increase  $dT'$  is the Dirac measure at  $y^*$ , and the change in the total tax bill converges to its c.d.f., the step function  $\mathbb{I}_{\{y \geq y^*\}}$ .

<sup>29</sup>See Golosov, Tsyvinski, and Werquin (2014) for details.

loss of generality.

### 3.1 Comparison to the exogenous-wage benchmark

Most of the taxation literature assumes exogenous wages. In this case, the incidence on government revenue is given by expression (18) with  $\hat{w}(y) \equiv \hat{w}_{\text{ex}}(y) = 0$  and  $\hat{l}(y) \equiv \hat{l}_{\text{ex}}(y) = -\varepsilon_r^S(y) \frac{\tau'(y)}{1-T'(y)}$ . Applying this formula to the elementary tax reform at income  $y^*$  leads to

$$d\mathcal{R}_{\text{ex}}(y^*) = 1 - \varepsilon_r^S(y^*) \frac{T'(y^*)}{1-T'(y^*)} \frac{y^* f_Y(y^*)}{1-F_Y(y^*)}. \quad (20)$$

Equation (20) expresses the impact of an increase in the marginal tax rate at income  $y^*$  as the sum of a mechanical increase in government revenue, which is normalized to \$1 by construction, and a behavioral revenue loss equal to the product of: (i) the endogenous reduction in the labor income of agent  $y^*$ ,  $\frac{y^*}{1-T'(y^*)} \varepsilon_r^S(y^*)$ ; (ii) the share  $T'(y^*)$  of this income change that accrues to the government; and (iii) the hazard rate of the income distribution,  $\frac{f_Y(y^*)}{1-F_Y(y^*)}$ . The hazard rate is a cost-benefit ratio that measures the fraction  $f_Y(y^*)$  of agents whose labor supply is distorted by the reform, relative to the fraction  $1 - F_Y(y^*)$  of agents whose tax bill increases lump-sum.

By contrast, in the general equilibrium environment, the labor supply responses  $\hat{l}(y)$  in formula (18) are equal to the sum of those obtained in the model with exogenous wages,  $\hat{l}_{\text{ex}}(y)$ , and those induced by the wage adjustments,  $\varepsilon_w^S(y) \hat{w}(y)$  (see formula (16)). We can therefore rewrite (18), for the elementary tax reform at income  $y^*$ , as

$$d\mathcal{R}(y^*) = d\mathcal{R}_{\text{ex}}(y^*) + \int_{\mathbb{R}_+} [T'(y) (1 + \varepsilon_w^S(y))] \hat{w}(y) y f_Y(y) dy, \quad (21)$$

where  $d\mathcal{R}_{\text{ex}}(y^*)$  is given by (20). In this expression, the term  $T'(y) (1 + \varepsilon_w^S(y))$  accounts for the effects of a unit change in the wage  $\hat{w}(y)$  on government revenue, both directly (via the term  $T'(y)$ ) and through the labor supply responses it induces (via the term  $T'(y) \varepsilon_w^S(y)$ ). The integral term in equation (21) therefore isolates the effects of the endogeneity of wages on government revenue, and thus allows for a clear comparison with the benchmark formula (20). The goal of the remainder of this section is to analyze this novel term.

### 3.2 Relationship between the own- and cross-wage elasticities

Before deriving an expression for the aggregate effects of tax reforms in general equilibrium, we start by stating the following lemma, which provides two versions of Euler's homogeneous function theorem in our economy.

**Lemma 2.** *The following relationship between the own-wage elasticity and the structural cross-wage elasticities is satisfied: for all  $y^*$ ,*

$$-\alpha(y^*) + \int_{\mathbb{R}_+} \tilde{\gamma}(y, y^*) y f_Y(y) dy = 0, \quad (22)$$

where  $\tilde{\gamma}(y, y^*) = \frac{\gamma(y, y^*)}{y^* f_Y(y^*)}$ . Equivalently, this can be expressed as a relationship between the own-wage elasticity and the GE cross-wage elasticities: for all  $y^*$ ,

$$-\alpha(y^*) + \int_{\mathbb{R}_+} \tilde{\Gamma}(y, y^*) y f_Y(y) dy = 0, \quad (23)$$

where  $\tilde{\Gamma}(y, y^*) = (1 + \alpha(y) \varepsilon_w^S(y))^{-1} \frac{\Gamma(y, y^*)}{y^* f_Y(y^*)}$ .

*Proof.* See Appendix A.2. Note that if the production function is CES, then  $\tilde{\gamma}(y, y^*) = 1/(\sigma \mathbb{E}[y])$  is constant and, if in addition the disutility of labor is isoelastic and the initial tax schedule is CRP, then  $\tilde{\Gamma}(y, y^*) = \tilde{\gamma}(y, y^*) = 1/(\sigma \mathbb{E}[y])$ .  $\square$

To interpret these equations, consider a one percent increase in the labor supply of agents with income  $y^*$ , induced for instance by a lower marginal tax rate. Their own wage decreases by  $\alpha(y^*)$  percent, while the wages of agents  $y \neq y^*$  increase by  $\gamma(y, y^*)$  percent. Because the production function has constant returns to scale, summing these effects has no impact on aggregate income: keeping labor supplies fixed, the income losses  $y^* \alpha(y^*)$  of agents with skill  $\theta^*$  are exactly compensated in the aggregate by the income gains  $y \gamma(y, y^*)$  of the other types  $\theta \neq \theta^*$ . This leads to equation (22). Now, these wage changes induce labor supply changes, which in turn affect wages through further rounds of own- and cross-effects in general equilibrium. Since Euler's homogeneous function theorem applies at every stage, the effect on aggregate income of all these wage adjustments is again equal to zero, that is,

$$\int_{\mathbb{R}_+} l(y) \times dw(y) f_Y(y) dy = \int_{\mathbb{R}_+} y \times \hat{w}(y) f_Y(y) dy = 0. \quad (24)$$

This equation implies that keeping labor supplies fixed at their initial level, the reshuffling of wages have distributional effects but keeps the economy's aggregate output constant. Using formulas (12) and (16) then leads to equation (23).

Suppose now that the disutility of labor is isoelastic, and that the initial tax schedule (before the tax reform) is linear. These assumptions imply that the elasticity  $\varepsilon_w^S(y)$  and the marginal tax rate  $T'(y)$  are constant. Constrasting equations (24) (which captures the change in aggregate output coming only from the wage changes) and (21) (which depends on the change in aggregate output coming from both the labor supply and the wage changes) immediately leads to the following corollary:

**Corollary 2.** *Suppose that the disutility of labor is isoelastic and that the initial tax schedule is linear. Then the incidence of an arbitrary nonlinear tax reform on government revenue is identical to that obtained assuming exogenous wages, i.e., for all  $y^*$ ,*

$$d\mathcal{R}(y^*) = d\mathcal{R}_{ex}(y^*),$$

where  $d\mathcal{R}_{ex}(y^*)$  is given by (20).

Note that this result holds for any production function that has constant returns to scale. Intuitively, constant returns to scale and the constant elasticity of labor supply imply that the income gain of skill  $\theta^*$  is exactly compensated in the aggregate by the income losses of the other types  $\theta \neq \theta^*$ . Since initially all income levels pay the same marginal tax rate, the government's tax revenue gain coming from the higher income of skill  $\theta^*$  is thus exactly compensated by the tax revenue losses coming from the rest of the population. Therefore the general-equilibrium contribution to the incidence of any tax reform on government budget, i.e. the integral term in (21), is equal to zero.

In the next section, we analyze the incidence of tax reforms on government revenue in the general case where the initial tax schedule is arbitrarily nonlinear.

### 3.3 Effects of tax reforms on government revenue

The following proposition expresses the impact of tax reforms on government revenue in general equilibrium and compares it to the expression (20) obtained in the model with exogenous wages.

**Proposition 2.** *The incidence of the elementary tax reform at income  $y^*$  on govern-*

ment revenue is given by

$$d\mathcal{R}(y^*) = d\mathcal{R}_{ex}(y^*) + \varepsilon_r(y^*) \frac{\Omega(y^*)}{1 - T'(y^*)} \frac{y^* f_Y(y^*)}{1 - F_Y(y^*)}, \quad (25)$$

where the variable  $\Omega(y^*)$  is defined by

$$\Omega(y^*) = \int_{\mathbb{R}_+} \left[ T'(y^*) (1 + \varepsilon_w^S(y^*)) - T'(y) (1 + \varepsilon_w^S(y)) \right] \tilde{\Gamma}(y, y^*) y f_Y(y) dy. \quad (26)$$

*Proof.* See Appendix C.2. □

Before deriving the economic implications of equation (25), it is useful to first sketch its proof. The direct effect of the elementary tax reform at income level  $y^*$  is to cause a decrease in the labor supply of type  $\theta^*$  by  $-\frac{y^* \varepsilon_r(y^*)}{1 - T'(y^*)}$ . Using Proposition 1, we can show that this implies (i) a positive own-wage effect on skill  $\theta^*$  equal to  $-\frac{y^* \varepsilon_r(y^*)}{1 - T'(y^*)} \times (-\alpha(y^*)) > 0$ , and (ii) a negative cross-wage effect on each skill  $\theta \in \Theta$  (or income  $y$ ) equal to  $-\frac{y^* \varepsilon_r(y^*)}{1 - T'(y^*)} \times \frac{\Gamma(y, y^*)}{1 + \alpha(y) \varepsilon_w^S(y)} < 0$ . In turn, a wage adjustment  $\hat{w}(y)$  at income  $y$  impacts government revenue by  $T'(y) (1 + \varepsilon_w^S(y)) \times \hat{w}(y)$ . Hence the general-equilibrium contribution to the incidence of the tax reform on government revenue is given by

$$\Omega(y^*) = T'(y^*) (1 + \varepsilon_w^S(y^*)) \alpha(y^*) - \int_{\mathbb{R}_+} T'(y) (1 + \varepsilon_w^S(y)) \tilde{\Gamma}(y, y^*) y f_Y(y) dy. \quad (27)$$

Euler's theorem (23) then leads to expression (26).

Suppose first that the disutility of labor is isoelastic and that the tax schedule is CRP (progressive or regressive), so that the labor supply elasticity  $\varepsilon_w^S(y)$  is constant. Moreover, assume in addition that the production function is CES, so that  $\alpha(y^*)$  and  $\tilde{\Gamma}(y, y^*)$  are also constant. We then have

$$\begin{aligned} \Omega(y^*) &\propto T'(y^*) \alpha(y^*) - \int_{\mathbb{R}_+} T'(y) \tilde{\Gamma}(y, y^*) y f_Y(y) dy \\ &= \frac{1}{\sigma} \left[ T'(y^*) - \int_{\mathbb{R}_+} T'(y) \frac{y}{\mathbb{E}y} f_Y(y) dy \right]. \end{aligned}$$

Suppose that the marginal tax rates are increasing in the initial economy, i.e., the rate of progressivity is  $p > 0$ . Then, in response to the tax reform, the government revenue gain from the higher income of agents  $\theta^*$ , which is proportional to  $T'(y^*)$ , is

increasing in  $y^*$ , whereas the tax revenue loss from the rest of the population, which is proportional to  $-\mathbb{E}[T'(y)y]$ , is independent of  $y^*$ . That is, the larger the income  $y^*$ , the higher the marginal tax rate at  $y^*$  relative to the (income-weighted) average marginal tax rate in the economy. Therefore, starting from a progressive tax schedule, the revenue gains from raising the marginal tax rates at the top and lowering them at the bottom, i.e., from raising further the progressivity of the tax schedule, are higher in general equilibrium ( $\Omega(y^*) > 0$ ) than the standard partial-equilibrium formula (20) would predict.

More generally, for an arbitrary production function, equation (27) implies that the general equilibrium contribution  $\Omega(y^*)$  is positive (resp., negative) if the marginal tax rate at  $y^*$  is larger (resp., smaller) than a weighted-average marginal tax rate in the economy, where the weights are now given by the modified cross-wage elasticities  $\tilde{\Gamma}(y, y^*)$ . Therefore, the efficiency loss from raising the tax rate at income  $y^*$  depends, in addition to the standard variables obtained in partial equilibrium, on the co-variation between the initial marginal tax rates  $T'(y)$  in the economy, and the complementarities  $\Gamma(y, y^*)$  with agent  $y^*$  in production. This is intuitive: the income of every individual is affected by the tax reform; the fiscal consequences depend on the size of these income changes (captured by the schedule of GE cross-wage elasticities) times the share of these changes that accrues to the government (captured by the schedule of marginal tax rates in the economy).

We conclude this section by formally stating the result we obtained above in the special case of a CES production function.

**Corollary 3.** *Suppose that the disutility of labor is isoelastic, that the initial tax schedule is CRP, and that the production function is CES. The incidence of the elementary tax reform at income  $y^*$  on government revenue can then be written as*

$$d\mathcal{R}(y^*) = d\mathcal{R}_{ex}(y^*) + \phi \varepsilon_r^S \frac{T'(y^*) - \bar{T}'}{1 - T'(y^*)} \frac{y^* f_Y(y^*)}{1 - F_Y(y^*)}, \quad (28)$$

where  $d\mathcal{R}_{ex}(y^*)$  is given by (20),  $\phi = \frac{1+\varepsilon_w}{\sigma+\varepsilon_w}$ ,<sup>30</sup> and  $\bar{T}' = \mathbb{E}[yT'(y)]/\mathbb{E}y$ . Thus the general-equilibrium effect on incidence is positive (resp., negative) if the marginal tax

---

<sup>30</sup>Recall that if the production function is CES, each round of cross-wage effects in general equilibrium is a constant fraction of the first round, so that the GE elasticity  $\Gamma(y, y^*)$  is proportional to the structural (first-round) elasticity  $\gamma(y, y^*)$  (equation (15)). In equation (28), the terms  $\varepsilon_r^S \frac{T'(y^*) - \bar{T}'}{1 - T'(y^*)} \frac{y^* f_Y(y^*)}{1 - F_Y(y^*)}$  are those that would be obtained by ignoring the full general-equilibrium adjustment of wages and labor supply, and instead focusing only on the first round of wage adjustments;  $\phi$  is then the discount factor that accounts for the second, third, etc. rounds of general equilibrium

rate at  $y^*$  is larger (resp., smaller) than the income-weighted average marginal tax rate in the economy.

*Proof.* See Appendix C.3. □

### 3.4 Social welfare and optimal tax schedule

The analysis of Section 3.3 can be easily extended to compute the incidence of tax reforms on social welfare, replacing  $d\mathcal{R}(y^*)$  with  $d\mathcal{W}(y^*)$  in equation (25).<sup>31</sup> The technical details are gathered in Appendix C.4. The first difference is that the exogenous-wage term  $d\mathcal{R}_{\text{ex}}(y^*)$  in the right hand side of (25) is replaced by  $d\mathcal{W}_{\text{ex}}(y^*) = d\mathcal{R}_{\text{ex}}(y^*) - \bar{g}(y^*)$ , where  $-\bar{g}(y^*) \equiv -\mathbb{E}[g(y) | y \geq y^*]$  is the total welfare loss from a higher marginal tax rate at income  $y^*$ . The second difference is that the variable  $T'(y) (1 + \varepsilon_w^S(y))$  in equation (26), which measures the total impact of a wage adjustment  $\hat{w}(y)$  on the government budget, is now replaced by the more general expression

$$\psi(y) = (1 + \varepsilon_w^S(y)) T'(y) + g(y) (1 - T'(y)). \quad (29)$$

The second term on the right hand side comes from the fact that the share  $1 - T'(y)$  of the income gain due to the wage adjustment  $\hat{w}(y)$  is kept by the individual; this in turn raises social welfare in proportion to the welfare weight  $g(y)$ .

The optimum tax schedule maximizes social welfare (5) subject to the constraint that government tax revenue (4) is non-negative. By imposing that the welfare effects of any tax reform of the initial tax schedule  $T$  are equal to zero, our tax incidence analysis immediately delivers a characterization of the optimum tax rates. In the model with exogenous wages (Diamond, 1998), the optimum schedule  $T'_{\text{ex}}(\cdot)$  is characterized by

$$\frac{T'_{\text{ex}}(y^*)}{1 - T'_{\text{ex}}(y^*)} = \frac{1}{\varepsilon_r^S(y^*)} (1 - \bar{g}(y^*)) \frac{1 - F_Y(y^*)}{y^* f_Y(y^*)}.$$

The tax rate at income  $y^*$ ,  $T'_{\text{ex}}(y^*)$ , is decreasing in the labor supply elasticity, the average social marginal welfare weight above income  $y^*$ , and the hazard rate of the income distribution. In the general-equilibrium model, we obtain instead:

---

wage adjustments.

<sup>31</sup>Note that if the government's welfare objective is Rawlsian and some agents always earn zero income, so that the social marginal welfare weights  $g(y)$  are equal to zero for all  $y > 0$ , the incidence of tax reforms on social welfare is the same as that on government revenue analyzed in Section 3.3.

**Corollary 4.** *The welfare-maximizing tax schedule  $T$  satisfies: for all  $y^* \in \mathbb{R}_+$ ,*<sup>32</sup>

$$\frac{T'(y^*)}{1 - T'_{ex}(y^*)} = \frac{T'_{ex}(y^*)}{1 - T'_{ex}(y^*)} + \tilde{\Omega}(y^*), \quad (30)$$

where  $\tilde{\Omega}(y^*)$  is the general-equilibrium contribution to optimal taxes:

$$\tilde{\Omega}(y^*) = \frac{\varepsilon_r(y^*)}{\varepsilon_r^S(y^*)} \int_{\mathbb{R}_+} [\psi(y^*) - \psi(y)] \tilde{\Gamma}(y, y^*) y f_Y(y) dy,$$

where  $\psi(\cdot)$  is defined by (29) and  $\tilde{\Gamma}(y, y^*)$  is as defined in Lemma 2.

*Proof.* See Appendix C.4. □

We show in Appendix C.4. that the optimal tax formula (30) can be easily expressed as an integral equation in  $T'(\cdot)$ . It can then be solved using the same techniques as those we used to derive the incidence of tax reforms on labor supply in Section 2.1. Moreover, the kernel of this integral equation becomes multiplicatively separable when the production function is CES, which leads to a straightforward solution and a particularly simple optimal tax formula. We devote Section 5 to a detailed analysis of this case.

### 3.5 Numerical exploration

In this section we calibrate our model to the U.S. economy and evaluate quantitatively the effects on government revenue of the elementary tax reforms at each income level, given by formula (25). We assume that the disutility of labor  $v(l)$  is isoelastic with parameter  $e = 0.33$  (Chetty, 2012),<sup>33</sup> and that the U.S. tax schedule is CRP with parameters  $p = 0.151$  and  $\tau = -3$  (Heathcote, Storesletten, and Violante, 2016). To match the U.S. yearly earnings distribution, we assume that  $f_Y(\cdot)$  is log-normal with mean 10 and variance 0.95 up to income  $y = \$150,000$ , above which we append a

---

<sup>32</sup>Note that in equation (30), the variables  $\varepsilon_r^S(y^*)$ ,  $\bar{g}(y^*)$ , and  $\frac{1-F_Y(y^*)}{y^* f_Y(y^*)}$  that appear in  $T'_{ex}(y^*)$  are evaluated in an economy where the general-equilibrium optimum tax schedule  $T$  (and not the exogenous-wage optimum  $T'_{ex}$ ) is implemented.

<sup>33</sup>In Appendix F.2, we discuss the connection between our model and the empirical literature that estimates the elasticity of taxable income (see, e.g., Saez, Slemrod, and Giertz (2012) for a survey). We show that the estimate for the taxable income elasticity at income  $y(\theta)$  maps in our model to the variable  $\varepsilon_r(\theta)$  defined in (10).

Pareto distribution with coefficient  $\pi = 1.5$ , i.e.,  $\lim_{\bar{y} \rightarrow \infty} \mathbb{E}[y|y \geq \bar{y}] / \bar{y} = \frac{\pi}{\pi-1} = 3$  (Diamond and Saez, 2011). As in Saez (2001), we obtain the distribution of wages  $w(\theta)$  from the earnings distribution and the individual first-order conditions (1). After choosing values for the elasticities of substitution, we can infer the remaining parameters of the production function. See Appendix F.1 for a more detailed description of the calibration procedure. We first study the case of a CES production function and then extend our results to a Translog production function, for which the elasticities of substitution are distance-dependent.

### CES production function

We first assume that the production function is CES and illustrate numerically the analytical result of Corollary 3. We choose an elasticity of substitution  $\sigma \in \{0.6; 3.1\}$ . The value  $\sigma = 0.6$  is taken from Dustmann, Frattini, and Preston (2013) who study the impact of immigration along the U.K. wage distribution and, as in our framework, group workers according to their position in the wage distribution.<sup>34,35</sup> The value  $\sigma = 3.1$  is taken from Heathcote, Storesletten, and Violante (2016), who structurally estimate this CES parameter for the U.S. by targeting cross-sectional moments of the joint equilibrium distribution of wages, hours, and consumption. There is no clear consensus in the empirical literature on how responsive relative wages are to changes in labor supply, and therefore on the appropriate value of  $\sigma$ ,<sup>36</sup> our two values are on the lower and higher sides of the typical empirical estimates.

Our results for the CES specification are illustrated in Figure 1. We plot the government revenue impact of elementary tax reforms at each income level in the model with exogenous wages (i.e., equation (20), illustrated by the red bold curve) and in general equilibrium (i.e., equation (28), illustrated by the black dashed curve),

---

<sup>34</sup>This literature is a useful benchmark because it studies the impact on relative wages of labor supply shocks of certain skills, which is exactly the channel we want to analyze in our tax setting (except that for us the labor supply shocks are caused by tax reforms rather than immigration inflows).

<sup>35</sup>Card (1990) and Borjas et al. (1997) also measure the skill type by the relative wage position when studying the impact of immigration on native wages. The setting of Dustmann, Frattini, and Preston (2013) fits our setting particularly well because they group workers into fine groups: 20 groups that contain 5% of the workforce respectively. In Appendix F.3, we formally show that the elasticity of substitution estimated in a framework with discrete earnings groups (e.g., percentiles or quartiles) can be used to calibrate a CES production function with a continuum of types.

<sup>36</sup>See, e.g., the debate on the impact of immigration on natives' wages (Peri and Yasenov, 2015; Borjas, 2015).

as a function of the income  $y(\theta)$  where the marginal tax rate is perturbed. A value of 0.7, say, at a given income  $y(\theta)$ , means that for each additional dollar of tax revenue mechanically levied by the tax reform at  $y(\theta)$ , the government effectively gains 70 cents, while 30 cents are lost through the behavioral responses of individuals; that is, the marginal excess burden of this tax reform is 30%. First, consider the red bold line: it has a U-shaped pattern which reflects of the shape of  $\frac{y^* f_Y(y^*)}{1-F_Y(y^*)}$  in (20). This is a well-known finding in the literature (Diamond, 1998; Saez, 2001).

The difference between the black dashed curve and the red bold curve captures the additional revenue effect due to the endogeneity of wages. In line with our analytical result of Corollary 3, we observe that this difference is positive for intermediate and high incomes (starting from about \$77,000, where the marginal tax rate equals its income-weighted average). Raising the marginal tax rates for these income levels is more desirable, in terms of government revenue, when the general equilibrium effects are taken into account, while the opposite holds for low income levels. The magnitude of the difference is substantial: the marginal excess burden from increasing the marginal tax rate on income \$200,000 is equal to 0.22 cents (resp., 0.30 cents) per dollar if  $\sigma = 0.6$  (resp.,  $\sigma = 3.1$ ) instead of 0.34 if  $\sigma = \infty$ , i.e., it is reduced by 35% (resp., 12%) due to the general equilibrium effects. Hence the model with exogenous wages significantly underestimates the revenue gains from increasing the progressivity of the tax code.

We explore the robustness of these results in Appendix F.4. We first consider other specifications of the U.S. tax schedule, in particular, we account for the phasing-out of transfers, as estimated by Guner, Rauh, and Ventura (2017), which implies high marginal tax rates at the bottom of the income distribution. Our main insight regarding the additional benefit of raising progressivity in general equilibrium is mitigated but not reversed (Figure 7). Second, we compute the effects of the elementary tax reforms on social welfare. As discussed in Corollary 1, the general equilibrium forces imply an increase (resp., decrease) in wages and utilities for individuals whose marginal tax rate increases (resp., for everyone else). This channel reduces the benefits of raising the progressivity of the tax schedule. Nevertheless our main result still holds if the social marginal welfare weights fall sufficiently fast with income (Figure 12).

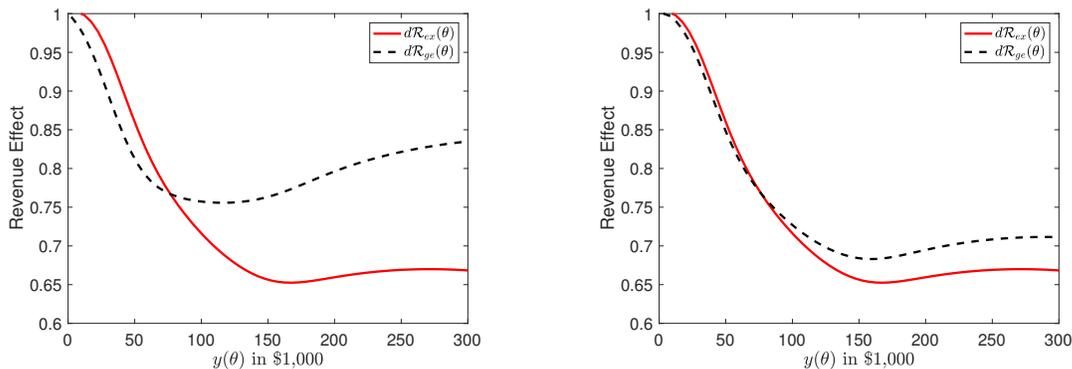


Figure 1: Revenue gains of elementary tax reforms at each income  $y(\theta)$ . Red bold lines: exogenous wages (equation (20)). Black dashed lines: CES technology with  $\sigma = 0.6$  (left panel) and  $\sigma = 3.1$  (right panel) (equation (28)).

### Translog production function

A criticism of the CES production function with a continuum of types is that high-skill workers (say) are equally substitutable with middle-skill workers as they are with low-skill workers. We therefore propose a more flexible parametrization of the production function that allows us to obtain distance-dependent elasticities of substitution, i.e., such that closer skill types are stronger substitutes.<sup>37</sup>

Specifically, in this paragraph we explore quantitatively the implications of the transcendental-logarithmic (Translog) production function. This specification can be used as a second-order local approximation to any production function (Christensen, Jorgenson, and Lau, 1973). With a continuum of labor inputs, its functional form is given by

$$\begin{aligned} \ln \mathcal{F}(\{L(\theta)\}_{\theta \in \Theta}) &= a_0 + \int_{\Theta} a(\theta) \ln L(\theta) d\theta + \dots \\ &\quad \frac{1}{2} \int_{\Theta} \beta(\theta) (\ln L(\theta))^2 d\theta + \frac{1}{2} \int_{\Theta \times \Theta} \chi(\theta, \theta') (\ln L(\theta)) (\ln L(\theta')) d\theta d\theta', \end{aligned} \quad (31)$$

where for all  $\theta, \theta'$ ,  $\int_{\Theta} a(\theta') d\theta' = 1$ ,  $\chi(\theta, \theta') = \chi(\theta', \theta)$ , and  $\beta(\theta) = -\int_{\Theta} \chi(\theta, \theta') d\theta'$ . These restrictions ensure that the technology has constant returns to scale. When  $\chi(\theta, \theta') = 0$  for all  $\theta, \theta'$ , the production function is Cobb-Douglas.

<sup>37</sup>Teulings (2005) obtains this distance-dependent property in an assignment model.

The elasticity of substitution between the labor of types  $\theta$  and  $\theta'$  is given by

$$\sigma(\theta, \theta') = \left[ 1 + \left( \frac{1}{\rho(\theta)} + \frac{1}{\rho(\theta')} \right) \chi(\theta, \theta') \right]^{-1},$$

where  $\rho(\theta) = \frac{w(\theta)L(\theta)}{\mathcal{F}(\mathcal{L})} = \frac{y(\theta)f(\theta)}{\mathbb{E}(y)}$  denotes the type- $\theta$  labor share of output.<sup>38</sup> To obtain distance-dependent elasticities, we propose the following specification of the exogenous parameters:

$$\chi(\theta, \theta') = \left( \frac{1}{\rho^c(\theta)} + \frac{1}{\rho^c(\theta')} \right)^{-1} \left[ c_1 - c_2 \exp \left( -\frac{1}{2s^2} (y^c(\theta) - y^c(\theta'))^2 \right) \right],$$

where  $c_1, c_2$  are constants, and where  $\rho^c(\theta)$  and  $y^c(\theta)$  are the current (i.e., empirically measured given the actual tax system) income share and income of type  $\theta$ . The local (i.e., such that  $\rho(\theta) = \rho^c(\theta)$  and  $y(\theta) = y^c(\theta)$ ) elasticity of substitution between workers in percentiles  $\theta$  and  $\theta'$  is then given by

$$\sigma(\theta, \theta') = \left\{ 1 + c_1 \left[ c_2 - \exp \left( -\frac{1}{2s^2} (y^c(\theta) - y^c(\theta'))^2 \right) \right] \right\}^{-1}. \quad (32)$$

The parameters  $c_1$  and  $c_2$  determine the values of the elasticity of substitution between types  $(\theta, \theta')$  with  $|y(\theta) - y(\theta')| \rightarrow \infty$  and  $\theta \approx \theta'$ , respectively. The parameter  $s$  specifies the rate at which  $\sigma(\theta, \theta')$  falls as  $\theta'$  moves away from  $\theta$ .

The left panel of Figure 2 shows the elasticity of substitution  $\sigma(\theta, \theta')$  as a function of  $\theta$  for such a specification, where  $\sigma(\theta, \theta')$  varies between 0.5 and 10. We let  $\theta \in \Theta = [0, 1]$  be the agent's percentile in the income distribution. We choose two values for  $\theta'$ : the type that earns the median income (\$33,500) and the type at the 95th percentile of the income distribution (\$126,500), i.e.,  $\theta' = 0.5$  (red bold line) and  $\theta' = 0.95$  (black dashed line). This illustrates how substitutable is the labor supply of a given skill type, measured by its income level  $y(\theta)$  on the x-axis, with the skills at the median and the 95th percentile. By construction, the elasticity of substitution equals 10 as  $\theta \rightarrow \theta'$ , then decreases with the distance  $|\theta - \theta'|$ , and converges to a value of 0.5 as  $\theta \rightarrow 1$ . As a comparison, we also plot the elasticity of substitution for a Cobb-Douglas production function, which is equal to 1 for any pair of types  $(\theta, \theta')$ . In Appendix F.4 we illustrate the cross-wage elasticities  $\gamma(\theta, \theta')$  and also explore alternative Translog specifications.

---

<sup>38</sup>We derive all of our results about the Translog production function in Appendix A.2.2.

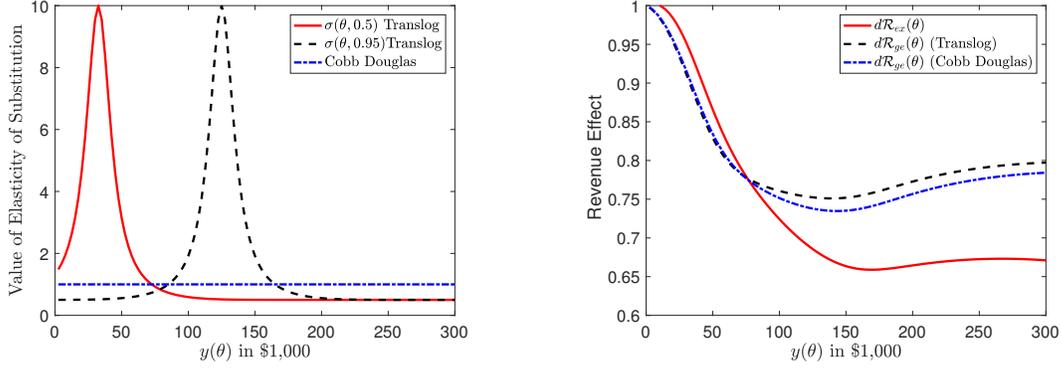


Figure 2: Left panel: Red bold (resp., black dashed) line: elasticity of substitution between types with income  $y(\theta)$  and the 50th (resp., 95th) percentile, for the Translog specification (32). Blue dashed-dotted lines: Cobb-Douglas specification. Right panel: Black dashed line (resp., red bold line, blue dashed-dotted line): Revenue gains of elementary tax reforms at income  $y(\theta)$  for the Translog specification (32) (resp., for exogenous wages, Cobb-Douglas production).

The right panel of Figure 2 plots the incidence on government revenue of the elementary tax reforms at each income  $y(\theta)$  (equation 25) for the Translog specification (32) (black dashed curve) and compares them to the Cobb-Douglas technology (blue dashed-dotted curve). The general-equilibrium contribution with distance-dependence is also positive for high incomes and of slightly larger magnitude.

## 4 Generalizations

In this section we show how the methodology and the results of Section 1 can be generalized to alternative and more sophisticated environments. We only briefly describe these extensions here; all of the details and proofs are gathered in Appendix D.

### 4.1 Income effects

In this section we extend the model of Section 1 to a general utility function over consumption and labor supply  $U(c, l)$ , where  $U_c, U_{cc} > 0$  and  $U_l, U_{ll} < 0$ . This specification allows for arbitrary substitution and income effects. The definitions of the structural cross- and own-wage elasticities  $\gamma(\theta, \theta')$  and  $\alpha(\theta)$  are identical to (8) and (9). We define the compensated (Hicksian) elasticity of labor supply with respect to the retention rate  $\varepsilon_r^S(\theta)$  and the income effect  $\varepsilon_R^S(\theta)$  as in (6), except that  $e(\theta)$  in the numerator is replaced by the standard expressions for, respectively,

$e_r^c(\theta) = \frac{\partial \ln l(\theta)}{\partial \ln r(\theta)} \Big|_u$  and  $e_R(\theta) = r(\theta) w(\theta) \frac{\partial l(\theta)}{\partial R}$ , given by (55) and (56) in Appendix D.1.1.<sup>39</sup> The labor supply elasticity with respect to the wage  $\varepsilon_w^S(\theta)$  is now defined by  $\varepsilon_w^S(\theta) = (1 - p(y(\theta))) \varepsilon_r^S(\theta) + \varepsilon_R^S(\theta)$ . The other elasticities are then defined exactly as in Section 1; in particular, the GE cross-wage elasticity  $\Gamma(\theta, \theta')$  is defined by (13), using the generalized expression for  $\varepsilon_w(\theta)$ .

With these general preferences, the incidence of an arbitrary tax reform  $\tau$  on individual labor supply is given by the following formula, which generalizes (12):

$$\hat{l}(\theta) = \hat{l}_{\text{pe}}(\theta) + \varepsilon_w(\theta) \int_{\Theta} \Gamma(\theta, \theta') \hat{l}_{\text{pe}}(\theta') d\theta', \quad (33)$$

where  $\varepsilon_w(\theta)$ , and  $\Gamma(\theta, \theta')$  are replaced by their generalized definitions, and where

$$\hat{l}_{\text{pe}}(\theta) \equiv -\varepsilon_r(\theta) \frac{\tau'(y(\theta))}{1 - T'(y(\theta))} + \varepsilon_R(\theta) \frac{\tau(y(\theta))}{(1 - T'(y(\theta))) y(\theta)}.$$

The interpretation of this formula is identical to that of (12), except that the partial-equilibrium impact of the reform  $\hat{l}_{\text{pe}}(\theta)$  is modified: in addition to the substitution effect already described in the quasilinear model, labor supply now also rises by an amount proportional to  $\varepsilon_R(\theta)$  due to an income effect induced by the higher total tax payment  $\tau(y(\theta))$  of agent  $\theta$ .<sup>40</sup> The (closed-form) incidence on wages, utilities and government revenue are then derived identically to the corresponding formulas in Section 2. Adding income effects to the basic framework therefore does not add any difficulty to our tax incidence analysis.

We can also generalize the result of Corollary 3 characterizing the incidence of tax reforms on government revenue. Suppose in addition to the assumptions of Corollary 3 that the utility function has the form  $U(c, l) = \frac{c^{1-\eta}}{1-\eta} - \frac{l^{1+\frac{1}{\varepsilon}}}{1+\frac{1}{\varepsilon}}$ . The revenue effect of the elementary tax reform at income  $y^*$  is then given by

$$\begin{aligned} d\mathcal{R}(y^*) &= d\mathcal{R}_{\text{ex}}(y^*) \\ &+ \phi \varepsilon_r^S \frac{T'(y^*) - \bar{T}'}{1 - T'(y^*)} \frac{y^* f_Y(y^*)}{1 - F_Y(y^*)} - \phi \varepsilon_r^S \eta \mathbb{E} \left[ \frac{T'(y) - \bar{T}'}{1 - T'(y)} \Big| y > y^* \right], \end{aligned} \quad (34)$$

where  $\bar{T}' = \mathbb{E}[yT'(y)]/\mathbb{E}y$  is the income-weighted average marginal tax rate in the

<sup>39</sup>See, e.g., p. 208 in Saez (2001).

<sup>40</sup>The partial-equilibrium formula for  $\hat{l}_{\text{pe}}(\theta)$  is identical to that derived in models with exogenous wages by Saez (2001) and Golosov, Tsyvinski, and Werquin (2014) (Proposition 1), except that that now the elasticities  $\varepsilon_r(\theta)$  and  $\varepsilon_R(\theta)$  take into account the own-wage effects  $\alpha(\theta)$ .

economy and where  $\phi$  is as defined in Corollary 3. Note that for  $\eta = 0$ , this formula reduces to equation (28). If  $\eta > 0$  and the baseline tax schedule is progressive, then the second general-equilibrium contribution has the opposite sign of the first. If top incomes are Pareto distributed and the baseline tax schedule is CRP, we derive in Appendix D.1.2 a necessary and sufficient condition on the progressivity parameter  $p$ , the Pareto coefficient and the curvature of the utility function  $\eta$  such that the first general-equilibrium term in (34) is larger than the second as  $y \rightarrow \infty$ . For empirically plausible values of the income effect parameter, the magnitude of the general-equilibrium contribution to government revenue incidence obtained in Section 3.5 is reduced by a third (and hence keeps the same direction).

## 4.2 Several sectors or education levels

The aggregate production function  $\mathcal{F}$  assumed in Section 1 takes as inputs the labor supply of each one-dimensional skill type  $\theta \in \Theta$ . In this framework, the skill  $\theta$  of an agent can be interpreted as her percentile in the wage distribution  $\{w(\theta)\}_{\theta \in \Theta}$ . Suppose now that the population is divided into  $N$  groups (e.g., sectors, education levels, etc.). Each group  $i$  is composed of a continuum of agents indexed by their skill  $\theta \in \Theta$  who earn wage  $w_i(\theta)$ . The assignment of each individual to a given group  $i$  is exogenous. Note that the wage distributions  $\{w_i(\theta)\}_{\theta \in \Theta}$  and  $\{w_j(\theta)\}_{\theta \in \Theta}$  of different groups  $i \neq j$  overlap.

The aggregate production function is now defined by

$$\mathcal{F} \left( \{L_i(\theta)\}_{(\theta,i) \in \Theta \times \{1, \dots, N\}} \right), \quad (35)$$

where  $L_i(\theta)$  is the aggregate labor supply of the agents of type  $\theta$  who work in sector  $i$ . Extending our analysis to this general production function is straightforward; the technical details are gathered in Appendix D.2.1. We define the wage, labor supply, and income of type  $\theta$  in sector  $i$  by  $w_i(\theta)$ ,  $l_i(\theta)$ , and  $y_i(\theta)$  respectively. We define as in Section 1.2, for each pair of skills  $(\theta, \theta')$  and groups  $(i, j)$ , the cross-wage elasticity  $\gamma_{i,j}(\theta, \theta')$ , the own-wage elasticity  $\alpha_i(\theta)$ , and the partial-equilibrium labor elasticities  $\varepsilon_{r,i}(\theta)$ ,  $\varepsilon_{w,i}(\theta)$ . A change of variables then allows us to define, for each income-group pair  $(y, i)$ , the wage  $w_{y,i}$  of the agents who earn income  $y$  in group  $i$ , and the  $N \times 1$  vector  $\mathbf{w}_y = (w_{y,i})_{i=1, \dots, N}$ . We define analogously the vectors  $\mathbf{l}_y$ ,  $\hat{\mathbf{l}}_y$ ,  $\boldsymbol{\varepsilon}_r(y)$ ,  $\boldsymbol{\varepsilon}_w(y)$ , and the  $N \times N$  matrices  $\boldsymbol{\gamma}(y, y')$  and  $\boldsymbol{\Gamma}(y, y')$ . We show that the result of Lemma 1

is now replaced by a system of linear integral equations, which can be solved using analogous steps as those leading to Proposition 1. We obtain that the incidence of an arbitrary tax reform  $\tau$  on individual labor supplies is given in closed-form by

$$\hat{l}_y = -\varepsilon_r(y) \frac{\tau'(y)}{1-T'(y)} - \int_{\mathbb{R}_+} \text{Diag}(\varepsilon_w(y)) \mathbf{\Gamma}(y, y') \varepsilon_r(y') \frac{\tau'(y')}{1-T'(y')} dy'. \quad (36)$$

The interpretation of this formula is identical to that of (12), with the only difference that the incidence of tax reforms now naturally depends on a larger number of (cross-sector) elasticities.

### Example: High School and College labor

A special case of the general production function (35) is the so-called canonical model (Acemoglu and Autor, 2011), where individuals are categorized according to their level of education (high school vs. college); this model has been studied empirically by Katz and Murphy (1992) and Card and Lemieux (2001). Let  $\mathcal{L}_H = \int_{\Theta} l_H(\theta) \theta g_H(\theta) d\theta$  and  $\mathcal{L}_C = \int_{\Theta} l_C(\theta) \theta g_C(\theta) d\theta$  denote the aggregate labor inputs in efficiency units in sectors  $i = H, C$ , respectively, where sector  $H$  is composed of the high school-educated workers (with density  $g_H$  over types  $\theta \in \Theta$ ), and sector  $C$  is composed of the college-educated workers (with density  $g_C$  over types  $\theta \in \Theta$ ). Suppose finally that the aggregate production function is given by a CES aggregator of  $\mathcal{L}_H$  and  $\mathcal{L}_C$ , i.e.,

$$\mathcal{F} = \left[ \mathcal{L}_H^{\frac{\sigma-1}{\sigma}} + \mathcal{L}_C^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}.$$

In this model, there is an infinite elasticity of substitution between workers within each education level, and a finite and constant elasticity of substitution  $\sigma$  across the two groups. The wage of an individual with type  $\theta$  in the education group  $i \in \{H, C\}$  is given by  $w_i(\theta) = \partial \mathcal{F} / \partial L_i(\theta)$ . In particular, we have  $\frac{w_i(\theta')}{w_i(\theta)} = \frac{\theta'}{\theta}$  for any two types  $(\theta, \theta')$ , so that the relative wages within each group  $i$  are given by the ratio of the corresponding exogenous skills.

As we show in Appendix D.2.2, Corollary 3 can be easily extended to this environment. Suppose that the disutility of labor is isoelastic and that the initial tax schedule is CRP, so that the labor supply elasticities  $\varepsilon_r^S$  and  $\varepsilon_w^S$  are constant. We

obtain in this case

$$d\mathcal{R}(y^*) = d\mathcal{R}_{\text{ex}}(y^*) + \phi \varepsilon_r^S [s_C(y^*) - s_C] \frac{\bar{T}'_C - \bar{T}'_H}{1 - T'(y^*)} \frac{y^* f_Y(y^*)}{1 - F_Y(y^*)}, \quad (37)$$

where  $d\mathcal{R}^{\text{ex}}(y^*)$  is given by (20),  $\bar{T}'_i = \int T'(y) y f_{Y,i}(y) dy$  is the income-weighted average marginal tax rate in education group  $i = H, C$ ,  $s_C(y^*)$  is the share of individuals earning  $y^*$  that are college-educated,  $s_C$  is the share of aggregate income accruing to college-educated workers, and  $\phi$  is defined as in Corollary 3.

Comparing equation (37) to equation (28) reveals two differences arising from the alternative modeling of the production. The first is that the general-equilibrium effect now depends on the difference between the average marginal tax rates in the two education groups (or sectors), rather than on the difference between the marginal tax rate at income  $y^*$  vs. in the population as a whole. This difference becomes clear if we interpret our production function in Section 1 as one that treats each skill  $\theta$  as a distinct sector. Second, the general-equilibrium contribution features an additional term that captures the difference between the share of education group  $C$  at income level  $y^*$  and the overall share of income accruing to group  $C$ . This term is positive if college educated labor is over-represented at income level  $y^*$ . This is because in this case an increase in the marginal tax rate at  $y^*$  raises wages in sector 1 and lowers them in sector 2. Note that this new term is bounded above by 1, and is equal to 1 if sector 1 is composed of all of the agents with type  $\theta^*$  (and only them), as is the case in Section 1.

### 4.3 Intensive and extensive margins

In this section we extend the model of Section 1 to an environment where individuals choose their labor supply both on the intensive margin (hours  $l$  conditional on participating in the labor force) and on the extensive margin (participation decision). Heterogeneity is two-dimensional: individuals are indexed by their skill type  $\theta \in [\underline{\theta}, \bar{\theta}]$  and by their cost of working  $\kappa \in \mathbb{R}_+$ . The utility function is given by  $U(c, l) = c - v(l) - \kappa \mathbb{I}_{\{l > 0\}}$ , where  $\mathbb{I}_{\{l > 0\}}$  is an indicator function equal to 1 if the agent is employed (i.e.,  $l > 0$ ). The intensive margin choice of labor effort  $l(\theta)$  depends on the marginal tax rate  $T'(y(\theta))$ , while the extensive margin choice consists of a participation threshold  $\kappa^*(\theta)$  that depends on the average tax rate relative to transfers,  $T(y(\theta)) - T(0)$ . The version of this model with exogenous wages has been

analyzed by [Saez \(2002\)](#).

The elasticities  $\eta_T^S(\theta)$  and  $\eta_w^S(\theta)$  of the participation rate at skill  $\theta$  with respect to the average tax rate and the wage are defined by the standard equations (71, 72) given in Appendix D.3. We then naturally define the partial-equilibrium elasticities  $\eta_T(\theta)$  and  $\eta_w(\theta)$  by (73). These elasticities are determined by the reservation density of agents with skill  $\theta$  who are close to indifference between participation and non-participation in the initial tax system. The GE cross-wage elasticity  $\Gamma(\theta, \theta')$  is still defined by (13), except that  $\varepsilon_w(\theta')$  is now replaced by  $(\varepsilon_w(\theta') + \eta_w(\theta'))$  in each term of the resolvent series. The incidence of an arbitrary tax reform  $\tau$  on the total labor supply  $L(\theta)$  of agents of skill  $\theta$  is given by the following formula, which generalizes Proposition 1:

$$\hat{L}(\theta) = \hat{L}_{\text{pe}}(\theta) + (\varepsilon_w(\theta) + \eta_w(\theta)) \int_{\Theta} \Gamma(\theta, \theta') \hat{L}_{\text{pe}}(\theta') d\theta', \quad (38)$$

where  $\varepsilon_w(\theta)$ , and  $\Gamma(\theta, \theta')$  are replaced by their generalized definitions, and where

$$\hat{L}_{\text{pe}}(\theta) \equiv -\varepsilon_r(\theta) \frac{\tau'(y(\theta))}{1 - T'(y(\theta))} - \eta_T(\theta) \frac{\tau(y(\theta))}{y(\theta) - T(y(\theta)) + T(0)}.$$

The interpretation of this formula is identical to that of (12), with two differences. First, the partial-equilibrium impact ( $\hat{L}^{\text{pe}}(\theta)$ ) is modified: in addition to the substitution effect already described in the quasilinear model, the tax reform now raises the tax payment of agents with skill  $\theta$  by  $\tau(y(\theta))$ , which lowers the total labor supply of that skill by an amount proportional to  $\eta_T(\theta)$ , by inducing those with a large fixed cost of working to drop out of the labor force. Second, the change in the wage of type  $\theta$  induces a decrease in total hours (from both intensive and extensive margin responses) given by  $(\varepsilon_w(\theta) + \eta_w(\theta))$  rather than simply  $\varepsilon_w(\theta)$ . From this formula, it is straightforward to obtain the incidence of any tax reform on individual labor supplies, wages, participation thresholds, participation rates, utilities, and government revenue. These are all derived in Appendix D.3. Adding participation decisions to the basic framework therefore does not add any difficulty to the incidence analysis.

## 4.4 Other extensions

In Appendix D.4, we derive the incidence of tax reforms for other extensions of our baseline model, namely, the cases where the production function has non-constant

returns to scale and capital is an input in production. These extensions can be easily analyzed using the techniques we introduced in Section 2.

## 5 A simple analysis of optimal taxes

Our tax incidence analysis of Section 2 delivered a general formula for the optimal tax schedule in general equilibrium (Corollary 4). In this section, we simplify and analyze it further by assuming that the production function is CES. This simplifying assumption allows us to generalize parsimoniously and transparently the optimal taxation formula of Diamond (1998), as all of the own- and cross-wage effects are then summarized by a single elasticity of substitution  $\sigma$ . This also allows us to generalize the standard public finance results (closed-form expression for the optimal top tax rate, U-shape of the marginal tax rates), and to immediately quantify the magnitude of the “Stiglitz effects” that arise in general equilibrium Stiglitz (1982) as a function of the degree of complementarity between skills in production. We finally provide robustness tests by computing the optimal tax schedule obtained for the distance-dependent Translog specification introduced in Section 3.5.

### 5.1 Optimal tax schedule with CES technology

Recall that formula (30) is an integral equation that can be solved to obtain the optimal tax rates.<sup>41</sup> In the case of a CES production function, this integral equation has a multiplicatively separable kernel and can thus be easily solved. We obtain the following proposition:

---

<sup>41</sup>Note that the income distribution is endogenous in the formula for optimal tax rates (39). As discussed in Section 1.1, there exists a monotone mapping between incomes and skills whenever the mapping between wages and skills is monotone – and this is w.l.o.g. because we can always relabel the type  $\theta$  as the agent’s position in the wage distribution. One caveat applies, however: as taxes change, the ordering of types may also change. Thus, expression (30) assumes that skills are ordered given the optimum tax schedule, rather than given the existing (suboptimal) tax code that may be used for the calibration in numerical simulations. Nevertheless, we show in Appendix A.2.3. that when the production function is CES (which we assume throughout Section 5), the ordering of types never changes, independently of the (possibly non-local) tax reform that is implemented, as long as the resulting tax function is differentiable. Moreover, in all of our numerical explorations for optimal income taxes using a Translog production function, the ordering of types is always the same for the optimal tax schedule and for the tax schedule to which the economy is calibrated.

**Proposition 3.** *Assume that the production function is CES with elasticity of substitution  $\sigma > 0$ . Then the optimal marginal tax rate at income  $y^*$  satisfies*

$$\frac{T'(y^*)}{1 - T'(y^*)} = \frac{1}{\varepsilon_r(y^*)} (1 - \bar{g}(y^*)) \frac{1 - F_Y(y^*)}{y^* f_Y(y^*)} + \frac{g(y^*) - 1}{\sigma}. \quad (39)$$

*Proof.* See Appendix E.1. □

The first term on the right hand side of (39) is analogous to the optimal tax formula obtained in the model with exogenous wages (Diamond, 1998; Saez, 2001): the marginal tax rate at income  $y^*$  is proportional to the inverse elasticity of labor  $\varepsilon_r(y^*)$  and to the hazard rate of the income distribution  $\frac{1 - F_Y(y^*)}{y^* f_Y(y^*)}$ , and is decreasing in the average marginal social welfare weight  $\bar{g}(y^*)$ . The only difference is that the relevant labor supply elasticity  $\varepsilon_r(y^*)$  accounts for the own-wage effects. Since  $\varepsilon_r(y^*) = \frac{\varepsilon_r^S(y^*)}{1 + \varepsilon_w^S(y^*)/\sigma} < \varepsilon_r^S(y^*)$ , this tends to raise optimal marginal tax rates, more so for lower values of the elasticity of substitution  $\sigma$ . Intuitively, increasing the marginal tax rate at  $y^*$  leads these agents to lower their labor supply, which raises their own wage and thus reduces the behavioral response.

The second term,  $(g(y^*) - 1)/\sigma$ , which works in the opposite direction, captures the fact that the wage and welfare of type  $\theta^*$  increase due to a higher marginal tax rate  $T'(y^*)$ , at the expense of the other individuals whose wage and welfare decrease (see Corollary 1). Suppose that the government values the welfare of individuals  $\theta^*$  less than average, i.e.,  $g(y^*) < 1$ .<sup>42</sup> This negative externality induced by the behavior of  $\theta^*$  implies that the cost of raising the marginal tax rate at  $y^*$  is higher than in partial equilibrium, and tends to lower the optimal tax rate. Conversely, the government gains by raising the optimal tax rates of individuals  $y^*$  whose welfare is valued more than average, i.e.,  $g(y^*) > 1$ . This induces these agents to work less and earn a higher wage, which makes them strictly better off, at the expense of the other individuals in the economy, whose wage decreases. This term is therefore a force for higher marginal tax rates at the bottom and lower marginal tax rates at the top if the government has a strictly concave (redistributive) social objective.

Formula (39) parsimoniously generalizes the optimal tax formula derived by Diamond (1998) and Saez (2001) assuming exogenous wages ( $\sigma = \infty$ ) to a general CES production function. It depends on one additional sufficient statistic, namely, the

---

<sup>42</sup>Since lump-sum transfers are available to the government, the average marginal social welfare weight in the economy is equal to 1.

elasticity of substitution between skills in production  $\sigma$ . This formula also extends the general-equilibrium analysis of [Stiglitz \(1982\)](#) in a model with two skills to the workhorse framework of taxation, i.e., with a continuum of types and arbitrary non-linear taxes. It allows us to go beyond the purely qualitative insights obtained in the two-type framework and make operational the theory of optimal tax design in general equilibrium. In particular, we show in the next two subsections how the key results obtained with exogenous wages, namely, the characterization of the optimal top tax rate and the U-shape of marginal tax rates, are affected.

## 5.2 Top tax rate

We now derive the implications of formula [\(39\)](#) for the optimal top tax rate. Let  $\pi > 1$  denote the Pareto coefficient of the tail of the income distribution in the data, that is,  $1 - F_Y(y) \underset{y \rightarrow \infty}{\sim} c/y^\pi$  for some constant  $c$ .

**Corollary 5.** *Assume that the production function is CES with elasticity of substitution  $\sigma > 0$  and that the disutility of labor is isoelastic with parameter  $e$ . Assume moreover that given the current tax schedule, incomes are Pareto distributed at the tail with coefficient  $\pi > 1$ , and that the top marginal tax rate that applies to these incomes is constant. Assume finally that the marginal social welfare weights at the top converge to  $\bar{g}$  (given the optimal tax schedule). Then the top tax rate of the optimal tax schedule satisfies*

$$\tau^* = \frac{1 - \bar{g}}{1 - \bar{g} + \pi \varepsilon_r \xi}, \quad (40)$$

with  $\varepsilon_r = \frac{e}{1+e/\sigma}$  and  $\xi = (1 - \pi \varepsilon_r / \sigma)^{-1}$ . In particular,  $\tau^*$  is strictly smaller than the optimal top tax rate in the model with exogenous wages ( $\sigma = \infty$ ).

*Proof.* See [Appendix E.2](#). The non-trivial part of the proof consists of showing that for a CES production function, if the income distribution has a Pareto tail in the data, then it has the same Pareto tail at the optimum, even though the wage distribution is endogenous.  $\square$

Formula [\(40\)](#) generalizes the familiar top tax rate result of [Saez \(2001\)](#) (in which  $\varepsilon_r = \varepsilon_r^S$  and  $\xi = 1$ ) to a CES production function. There is one new sufficient statistic, the elasticity of substitution between skills in production  $\sigma$ , that is no longer restricted to being infinite. This proposition implies a strictly lower top marginal tax rate than if wages were exogenous. Immediate calculations of the optimal top tax rate illustrate

the power of this formula. Suppose that  $\bar{g} = 0$ ,  $\pi = 2$ ,  $e = 0.5$ , and  $\sigma = 1.5$ .<sup>43</sup> We immediately obtain that the optimal tax rate on top incomes is equal to  $\tau^{*ex} = 50\%$  in the model with exogenous wages, and falls to  $\tau^* = 40\%$  once the general equilibrium effects are taken into account. Suppose instead that  $\pi = 1.5$  and  $e = 0.33$ , then we get  $\tau^{*ex} = 66\%$  and  $\tau^* = 64\%$ . In this case the trickle-down forces barely affect the optimum tax rate quantitatively. Figure 3 shows more comprehensive comparative statics.<sup>44</sup>

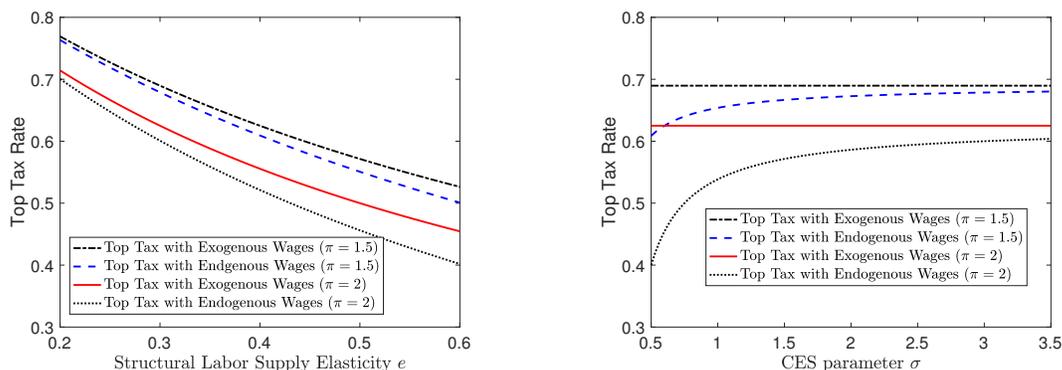


Figure 3: Optimal top tax rate as a function of the labor supply elasticity  $e$  (left panel,  $\sigma = 3.1$  fixed) and the elasticity of substitution  $\sigma$  (right panel,  $e = 0.33$  fixed) and for varying Pareto parameters  $\pi$ .

### 5.3 U-shape of optimal marginal tax rates

We finally analyze the impact of general equilibrium on the shape of optimal tax rates. Suppose for simplicity that the social planner is Rawlsian, i.e., it maximizes the lump-sum component of the tax schedule, so that  $g(y) = 0$  for all  $y > 0$ .<sup>45</sup> The partial-equilibrium equivalent of formula (39) for optimal taxes (for which the second term on the right hand side is equal to zero) generally implies a U-shaped pattern of marginal tax rates (Diamond, 1998; Saez, 2001) because the hazard rate  $\frac{1-F_Y(y)}{yf_Y(y)}$  is a U-shaped function of income  $y$ . If this is the case, then the additional term  $-1/\sigma < 0$  in (39) leads to a general equilibrium correction for  $T'(\cdot)$  that is also

<sup>43</sup>These values are meant to be only illustrative but they are in the range of those estimated in the empirical literature. See the calibration in Section 3.5.

<sup>44</sup>See also Green and Phillips (2015) who study quantitatively the size of the optimal top tax rate in a two-sector model.

<sup>45</sup>Thus, if the income of the lowest type is positive, we assume that there are some additional agents in the economy who are unable to work and whose consumption equals the demogrant.

U-shaped, because the optimal marginal tax rate  $T'(y^*)$  is increasing and concave in the right hand side of (39).<sup>46</sup> This suggests that the general equilibrium forces tend to reinforce the U-shape of the optimum tax schedule.

To formalize this intuition using our tax incidence analysis of Section 2, we start by defining a benchmark optimal tax schedule with exogenous wages, to which we can compare our general-equilibrium formula.

**Defining a benchmark with exogenous wages.** First, we define the marginal tax rates that a partial-equilibrium planner would set from Diamond (1998) using the same data to calibrate the model, and making the same assumptions about the utility function, but wrongly assuming that the wage distribution is exogenous:

$$\frac{T'_{\text{ex}}(y(\theta))}{1 - T'_{\text{ex}}(y(\theta))} = \left(1 + \frac{1}{\varepsilon_r^S(\hat{w}(\theta))}\right) \frac{1 - F_{\hat{W}}(\hat{w}(\theta))}{f_{\hat{W}}(\hat{w}(\theta))\hat{w}(\theta)}, \quad (41)$$

where  $\hat{w}(\theta)$  are the wages inferred from the data, i.e., obtained from the incomes observed empirically and the first-order conditions (1), and  $F_{\hat{W}}$  is the corresponding wage distribution. Formula (41) is the benchmark to which we compare our optimal policy results numerically, thus directly highlighting how our policy implications differ from those of Diamond (1998).

A government that would implement this tax formula, however, would then observe that the implied distribution of wages does change and is not consistent with the optimality of the tax schedule (41). To overcome this inconsistency, we consider a self-confirming policy equilibrium (SCPE)  $T'_{\text{scpe}}(y(\theta))$ , as originally proposed by Rothschild and Scheuer (2013, 2016), which is such that implementing the tax schedule (41) generates a wage distribution given which these tax rates are optimal – in other words, this construction solves for the fixed point between the wage distribution and the tax schedule. We use this concept as our exogenous-wage benchmark for our theoretical analysis below.

**Comparing optimal taxes to those obtained with exogenous wages.** We can apply our tax incidence result of Proposition 2 using the SCPE tax schedule as our initial tax schedule. This exercise gives the (first-order) welfare gains of reforming

---

<sup>46</sup>That is, if the function  $\frac{h(y)}{1+h(y)}$  with  $h(y) = \frac{1}{\varepsilon_r(y)} \frac{1-F_Y(y)}{y f_Y(y)}$  is U-shaped, it is easy to check that the general-equilibrium correction to marginal tax rates  $y \mapsto \frac{h(y)-\sigma^{-1}}{1+h(y)-\sigma^{-1}} - \frac{h(y)}{1+h(y)}$  is then also U-shaped.

this tax schedule at any income level, and hence the shape of the general-equilibrium correction to the optimal policy obtained assuming exogenous wages. We provide the corresponding incidence formula for a general CES technology in Appendix E.3, and focus on the simpler Cobb-Douglas case ( $\sigma = 1$ ) here.

**Corollary 6.** *Suppose that the production function is Cobb-Douglas ( $\sigma = 1$ ), that the initial tax schedule  $T = T_{scpe}$  is the SCPE, and that the disutility of labor is isoelastic with parameter  $e$ . The incidence of the elementary tax reform at income  $y^*$  on government revenue (or Rawlsian welfare) is given by*

$$d\mathcal{R}_{scpe}(y^*) = 1 - \zeta \frac{1}{T'(y^*)} \left[ p(y^*) + \frac{1}{e} \right], \quad (42)$$

where  $\zeta^{-1} \equiv \frac{1}{\bar{T}'} [\bar{p} + \frac{1}{e}] > 0$  is a constant that depends on the income-weighted averages of the marginal tax rate  $\bar{T}' = \mathbb{E}[\frac{y}{\mathbb{E}y} T'(y)]$  and of the local rate of progressivity  $\bar{p} = \mathbb{E}[\frac{y}{\mathbb{E}y} p(y)]$  in the initial economy.

*Proof.* See Appendix E.3. □

The map  $y^* \mapsto d\mathcal{R}_{scpe}(y^*)$  in (42) gives the shape of the general-equilibrium correction to the optimal tax schedule obtained assuming exogenous wages. Importantly, just as the result of Corollary 3, formula (42) shows that the general-equilibrium effects of the tax reform have a shape that is inherited from that of the initial tax schedule. In particular, if the marginal tax rates  $T'(y^*)$  of the SCPE are U-shaped as a function of income, the term  $-1/T'(y^*)$  in equation (42) leads to a general-equilibrium correction that is itself U-shaped. Note, however, that the additional term in general equilibrium depends also on the rate of progressivity  $p(y^*)$  of the initial (SCPE) tax schedule. Nevertheless, since  $|p(y^*)| < 1 \ll \frac{1}{e} \approx 3$  (Chetty (2012)), the shape of  $d\mathcal{R}(y^*)$  as a function of  $y^*$  is mostly driven by the term  $-1/T'(y^*)$ ; our numerical simulations below confirm this intuition.

## 5.4 Numerical simulations

Our calibration is similar to that in Section 3.5, see Appendix F.1 for details. Throughout this section we consider a Rawlsian social objective. In Appendix F.5.2 we simulate optimal taxes for concave social welfare functions  $G$ ; our results are similar.

**The role of the elasticity of substitution.** The left panel of Figure 4 plots the optimal marginal tax rates as a function of types<sup>47</sup> for two different values of the elasticity of substitution, and for the exogenous-wage planner defined in (41). The latter schedule has a familiar U-shape (Diamond, 1998; Saez, 2001). In line with our theoretical results of Section 5.1, the top tax rate is lower in general equilibrium and decreasing with  $\sigma$ . Moreover, the optimal marginal tax rates are reduced by an even larger amount at income levels close to the bottom of the U (around \$100,000), and are higher at low income levels (below \$40,000).<sup>48</sup> This confirms our findings of Corollary 6 and implies that the U-shape obtained for exogenous-wages is reinforced by the general equilibrium considerations.

**Translog production function.** In the right panel of Figure 4, we illustrate the optimal marginal tax rates in case of the Translog production function with distance-dependent elasticities of substitution, as calibrated in Section 3.5, and compare it to the optimal tax schedule in the case of a Cobb-Douglas production function; the graph shows that the policy implications are hardly altered, which justifies our focus on the case of a CES production for the theoretical analysis of this section. In Appendix F.5.3, we consider alternative Translog specifications and obtain similar conclusions.

---

<sup>47</sup>The scale on the horizontal axis on the left panel is measured in income; e.g., the value of the optimal marginal tax rate at the notch \$100,000 is that of a type  $\theta$  who earns an income  $y(\theta) = \$100,000$  in the calibration to the U.S data. The income that this type earns in the optimal allocation is generally different (see the right panel). In Appendix F.5.1, we also provide the optimal tax schedule as a function of incomes at the optimum.

<sup>48</sup>Since the exogenous-wage tax rates are already very high at those low income levels, the general equilibrium effects are quantitatively very small (at most 1.8 percentage points).

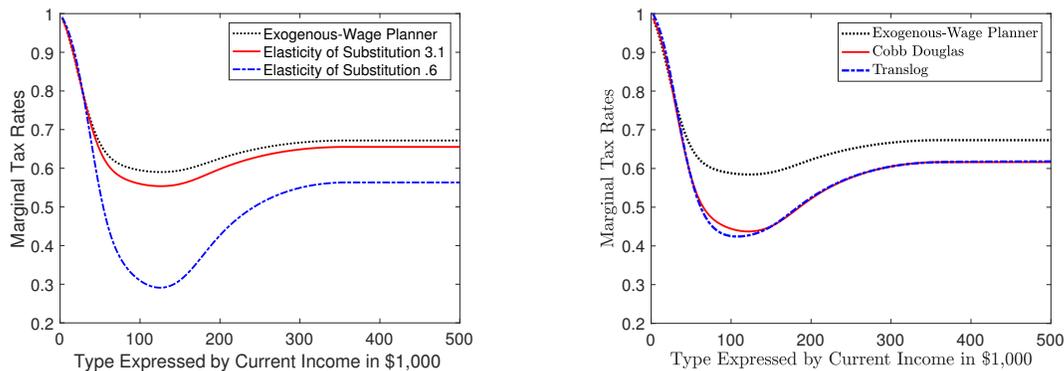


Figure 4: Optimal marginal tax rates as a function of types. Black dashed lines: exogenous wages. Left panel: CES technology with  $\sigma = 3.1$  (red bold line) and  $\sigma = 0.6$  (blue dashed-dotted line). Right panel: Translog technology (32) (blue dashed-dotted line) and Cobb-Douglas ( $\sigma = 1$ ) (red bold line).

**U-shape of the general-equilibrium effect.** Next, we plot in Figure 5 the shape of the general-equilibrium correction to the optimal taxes obtained in the model with exogenous wages. We do so by applying our incidence formula (25) using (41) (i.e., the black-dotted curve in Figure 4) as our initial tax schedule. Recall that Corollary 6 addresses the same question analytically using the SCPE as the exogenous-wage benchmark. The red bold line plots the effects of the tax reform according to the exogenous-wage planner (41). These effects are uniformly equal to zero by construction. The black dashed line shows that when the low-income marginal tax rates are high (as in the exogenous-wage optimum) rather than low (as in the CRP tax code assumed in Corollary 3), the general equilibrium forces call for lower tax rates for intermediate and high incomes, and higher marginal tax rates for low incomes. This graph implies that starting from the exogenous-wage optimum, the gains from perturbing the marginal tax rates are themselves U-shaped and negative, except at the very bottom of the income distribution, thus confirming our theoretical result of Corollary 6.

## Conclusion

In this paper we have developed a variational approach for the study of nonlinear tax reforms in general equilibrium. It can be used to study the incidence of reforming a given baseline tax schedule, e.g. the current U.S. tax code, as well as to characterize

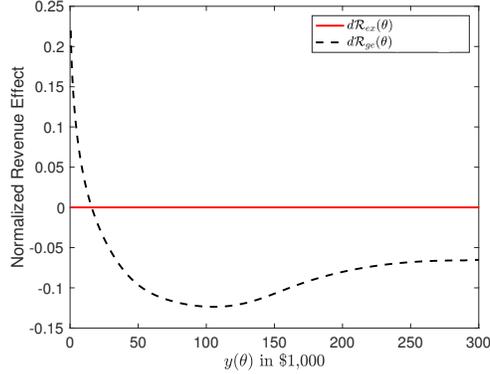


Figure 5: Tax incidence around the exogenous-wage optimum (41). Red bold line: model with exogenous wages. Black dashed line: CES production function ( $\sigma = 3.1$ ).

the optimal tax schedule. The key methodological tool that we brought into the analysis is the concept of integral equations. Our variational approach is powerful in that it allows to analyze sophisticated environments (with general individual preferences and choice variables, general production structures, etc.) with no additional technical difficulties as in the simpler baseline model.

## References

- ACEMOGLU, D. AND D. AUTOR (2011): “Skills, Tasks and Technologies: Implications for Employment and Earnings,” in *Handbook of Labor Economics Vol. 4*, ed. by O. Ashenfelter and D. Card, Elsevier, 1043–1171.
- ALES, L., A. A. BELLOFATTO, AND J. J. WANG (2017): “Taxing Atlas: Executive compensation, firm size, and their impact on optimal top income tax rates,” *Review of Economic Dynamics*, 26, 62–90.
- ALES, L., M. KURNAZ, AND C. SLEET (2015): “Technical change, wage inequality, and taxes,” *The American Economic Review*, 105, 3061–3101.
- ALES, L. AND C. SLEET (2016): “Taxing Top CEO Incomes,” *American Economic Review* (forthcoming).
- ANTRAS, P., A. DE GORTARI, AND O. ITSKHOKI (2016): “Globalization, Inequality and Welfare,” *Working Paper*.

- BÉNABOU, R. (2002): “Tax and Education Policy in a Heterogenous-Agent Economy: What of Levels of Redistribution Maximize Growth and Efficiency?” *Econometrica*, 70, 481–517.
- BORJAS, G. J. (2003): “The Labor Demand Curve Is Downward Sloping: Reexamining the Impact of Immigration on the Labor Market,” *The Quarterly Journal of Economics*, 118, 1335–1374.
- (2015): “The wage impact of the Marielitos: A reappraisal,” *ILR Review*.
- BORJAS, G. J., R. B. FREEMAN, L. F. KATZ, J. DINARDO, AND J. M. ABOWD (1997): “How much do immigration and trade affect labor market outcomes?” *Brookings papers on economic activity*, 1997, 1–90.
- BUCCI, A. AND P. USHCHEV (2014): “The Sources of Sharing Externalities: Specialization versus Competition,” *Working Paper*.
- CARD, D. (1990): “The impact of the Mariel boatlift on the Miami labor market,” *ILR Review*, 43, 245–257.
- (2009): “Immigration and Inequality,” *American Economic Review*, 99, 1–21.
- CARD, D. AND T. LEMIEUX (2001): “Can Falling Supply Explain the Rising Return to College for Younger Men? A Cohort-Based Analysis,” *Quarterly Journal of Economics*, 116, 705–746.
- CESARINI, D., E. LINDQVIST, M. J. NOTOWIDIGDO, AND R. OSTLING (2017): “The Effect of Wealth on Individual and Household Labor Supply: Evidence from Swedish Lotteries,” *American Economic Review* (*forthcoming*).
- CHEN, L. AND C. ROTHSCHILD (2015): “Screening with endogenous preferences,” *Working Paper*.
- CHETTY, R. (2009): “Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods,” *Annu. Rev. Econ.*, 1, 451–488.
- (2012): “Bounds on elasticities with optimization frictions: A synthesis of micro and macro evidence on labor supply,” *Econometrica*, 80, 969–1018.
- CHRISTENSEN, L. R., D. W. JORGENSON, AND L. J. LAU (1973): “Transcendental logarithmic production frontiers,” *The Review of Economics and Statistics*, 28–45.

- DIAMOND, P. A. (1998): “Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates,” *American Economic Review*, 88, 83–95.
- DIAMOND, P. A. AND J. A. MIRRLEES (1971): “Optimal taxation and public production II: Tax rules,” *The American Economic Review*, 61, 261–278.
- DIAMOND, P. A. AND E. SAEZ (2011): “The Case for a Progressive Tax: From Basic Research to Policy Recommendations,” *Journal of Economic Perspectives*, 25, 165–90.
- DUSTMANN, C., T. FRATTINI, AND I. P. PRESTON (2013): “The effect of immigration along the distribution of wages,” *The Review of Economic Studies*, 80, 145–173.
- FRADERA, I. (1986): “Perfect competition with product differentiation,” *International Economic Review*, 529–538.
- FULLERTON, D. AND G. E. METCALF (2002): “Tax incidence,” *Handbook of public economics*, 4, 1787–1872.
- GOLOSOV, M., A. TSYVINSKI, AND N. WERQUIN (2014): “A Variational Approach to the Analysis of Tax Systems,” *Working Paper*.
- GREEN, R. K. AND M. D. PHILLIPS (2015): “Demand for ‘The 1%’: Tax Incidence and Implications for Optimal Income Tax Rates,” *Working Paper*.
- GUNER, N., R. KAYGUSUZ, AND G. VENTURA (2014): “Income taxation of US households: Facts and parametric estimates,” *Review of Economic Dynamics*, 17, 559–581.
- GUNER, N., C. RAUH, AND G. VENTURA (2017): “Means-Tested Transfers in the US: Facts and Parametric Estimates,” *Working Paper*.
- HARBERGER, A. C. (1962): “The incidence of the corporation income tax,” *The Journal of Political Economy*, 215–240.
- HART, O. D. (1979): “Monopolistic competition in a large economy with differentiated commodities,” *The Review of Economic Studies*, 46, 1–30.

- HEATHCOTE, J., K. STORESLETTEN, AND G. L. VIOLANTE (2016): “Optimal tax progressivity: An analytical framework,” *Quarterly Journal of Economics* (forthcoming).
- HENDREN, N. (2015): “The Policy Elasticity,” in *Tax Policy and the Economy, Volume 30*, University of Chicago Press.
- HINES, J. R. (2009): “Peter Mieszkowski and the General Equilibrium Revolution in Public Finance,” in *Proceedings. Annual Conference on Taxation and Minutes of the Annual Meeting of the National Tax Association*, JSTOR, vol. 102, 213–216.
- IMBENS, G. W., D. B. RUBIN, AND B. I. SACERDOTE (2001): “Estimating the effect of unearned income on labor earnings, savings, and consumption: Evidence from a survey of lottery players,” *American Economic Review*, 778–794.
- ITSKHOKI, O. (2008): “Optimal redistribution in an open economy,” *Working Paper*.
- JACQUET, L. AND E. LEHMANN (2017): “Optimal Income Taxation with Composition Effects,” *IZA Discussion Paper No. 11019*.
- KATZ, L. F. AND K. M. MURPHY (1992): “Changes in Relative Wages, 1963-1987: Supply and Demand Factors,” *The Quarterly Journal of Economics*, 107, 35–78.
- KLEVEN, H., C. KREINER, AND E. SAEZ (2009): “The Optimal Income Taxation of Couples,” *Econometrica*, 77, 537–560.
- KOTLIKOFF, L. J. AND L. H. SUMMERS (1987): “Tax incidence,” *Handbook of public economics*, 2, 1043–1092.
- KRESS, R. (2014): *Linear integral equations (Third Edition)*, Springer.
- KRUSELL, P., L. E. OHANIAN, J.-V. RÍOS-RULL, AND G. L. VIOLANTE (2000): “Capital-skill complementarity and inequality: A macroeconomic analysis,” *Econometrica*, 68, 1029–1053.
- MIRRLEES, J. A. (1971): “An Exploration in the Theory of Optimum Income Taxation,” *The Review of Economic Studies*, 38, 175–208.
- MUSGRAVE, R. A. AND T. THIN (1948): “Income tax progression, 1929-48,” *Journal of political Economy*, 56, 498–514.

- PERI, G. AND V. YASENOV (2015): “The labor market effects of a refugee wave: Applying the synthetic control method to the Mariel boatlift,” Tech. rep., National Bureau of Economic Research.
- PIKETTY, T. (1997): “La redistribution fiscale face au chômage,” *Revue française d’économie*, 12, 157–201.
- PIKETTY, T. AND E. SAEZ (2013): “A Theory of Optimal Inheritance Taxation,” *Econometrica*, 81, 1851–1886.
- PIKETTY, T., E. SAEZ, AND S. STANTCHEVA (2014): “Optimal taxation of top labor incomes: A tale of three elasticities,” *American Economic Journal: Economic Policy*, 6, 230–271.
- POLYANIN, A. D. AND A. V. MANZHIROV (2008): *Handbook of integral equations*, CRC press.
- PRESS, W. H. (2007): *Numerical recipes 3rd edition: The art of scientific computing*, Cambridge university press.
- ROTHSCHILD, C. AND F. SCHEUER (2013): “Redistributive Taxation in the Roy Model,” *The Quarterly Journal of Economics*, 128, 623–668.
- (2014): “A theory of income taxation under multidimensional skill heterogeneity,” *NBER Working Paper No. 19822*.
- (2016): “Optimal Taxation with Rent-Seeking,” *Review of Economic Studies* (forthcoming).
- ROTHSTEIN, J. (2010): “Is the EITC as good as an NIT? Conditional cash transfers and tax incidence,” *American Economic Journal: Economic Policy*, 2, 177–208.
- SAEZ, E. (2001): “Using Elasticities to Derive Optimal Income Tax Rates,” *Review of Economic Studies*, 68, 205–229.
- (2002): “Optimal Income Transfer Programs: Intensive versus Extensive Labor Supply Responses,” *Quarterly Journal of Economics*, 117, 1039–1073.
- SAEZ, E., J. SLEMROD, AND S. H. GIERTZ (2012): “The elasticity of taxable income with respect to marginal tax rates: A critical review,” *Journal of economic literature*, 50, 3–50.

- SAEZ, E. AND S. STANTCHEVA (2016): “Generalized Social Marginal Welfare Weights for Optimal Tax Theory,” *American Economic Review*, 106, 24–45.
- SCHEUER, F. (2014): “Entrepreneurial Taxation with Endogenous Entry,” *American Economic Journal: Economic Policy*, 6, 126–163.
- SCHEUER, F. AND I. WERNING (2016): “Mirrlees meets Diamond-Mirrlees,” *NBER Working Paper No. 22076*.
- (2017): “The taxation of superstars,” *The Quarterly Journal of Economics*, 132, 211–270.
- SEIERSTAD, A. AND K. SYDSAETER (1986): *Optimal control theory with economic applications*, Elsevier North-Holland, Inc.
- SHOVEN, J. B. AND J. WHALLEY (1984): “Applied general-equilibrium models of taxation and international trade: an introduction and survey,” *Journal of Economic Literature*, 22, 1007–1051.
- STIGLITZ, J. E. (1982): “Self-selection and Pareto efficient taxation,” *Journal of Public Economics*, 17, 213–240.
- (2015): “In Praise of Frank Ramsey’s Contribution to the Theory of Taxation,” *The Economic Journal*, 125, 235–268.
- TEULINGS, C. N. (2005): “Comparative advantage, relative wages, and the accumulation of human capital,” *Journal of Political Economy*, 113, 425–461.
- TRICOMI, F. (1985): “Integral equations, Interscience, London, 1957, New edition,” .
- ZEMYAN, S. M. (2012): *The classical theory of integral equations: a concise treatment*, Springer Science & Business Media.