Tax curvature*

Albert Jan Hummel†

June 4, 2023

Abstract

In a Mirrleesian environment, a monopsonist sets hourly wages and individuals choose how many hours to work. Labor market outcomes do not only depend on the level and slope of the income tax function, but also on its curvature. A more concave tax schedule raises the elasticity of labor supply, which boosts wages. Consequently, optimal marginal tax rates for low-skilled workers are declining in income. I derive an optimal tax formula in terms of sufficient statistics that accounts for the impact of tax curvature on labor market outcomes. Numerical simulations suggest that optimal marginal tax rates decline rapidly at low incomes.

JEL classification: H21, J38, J42
Keywords: monopsony, optimal taxation, tax curvature

*I would like to thank the editor Nathaniel Hendren, three anonymous referees, Thomas Gaube, Aart Gerritsen, Bas Jacobs, Florian Scheuer, Kevin Spiritus, Clemence Tricaud and seminar participants at University of Amsterdam, CESifo Area Conference on Public Economics 2022, and IIPF Annual Congress 2022 for helpful comments and suggestions.

†University of Amsterdam, Tinbergen Institute, CESifo and Instituut voor Publieke Economie. E-mail: a.j.hummel@uva.nl.

1 Introduction

Firms exercise significant market power when setting wages. How much market power they can exert depends critically on the elasticity of labor supply (Robinson, 1933, Manning, 2003). This statistic is not policy-invariant. I show that this insight has an implication for tax policy. In particular, if wages are set by a monopsonist and individuals choose how many hours to work, the government can use the *curvature* of the tax function to boost wages of low-skilled workers by raising their elasticity of labor supply. This is achieved by setting marginal tax rates that are declining in income. Consequently, the optimal tax schedule is concave at the bottom of the income distribution.

To reach this conclusion, I study a Mirrleesian environment where individuals differ in their ability, which is not observed by the government. Individuals take their hourly wage and the tax schedule as given and optimally choose how many hours to work. Wages are not determined competitively as in Mirrlees (1971), but instead set by a monopsonist that observes ability and maximizes profits taking into account how wages affect individual labor supply. As a result, wages are marked down relative to productivity, which generates downward distortions in labor supply. The government has a preference for redistribution and levies a nonlinear tax on labor earnings and an exogenous tax on profits.

A novel feature is that labor market outcomes do not only depend on the level and slope of the tax function, but also on its curvature, i.e., the the second derivative. The latter determines how responsive individuals are to wage changes. To illustrate, suppose the tax function is convex and marginal tax rates increase steeply with income. Individuals then have weak incentives to work longer hours following an increase in the hourly wage. A low elasticity of labor supply, in turn, implies that a profit-maximizing monopsonist sets low wages. Consequently, a local increase in the *second* derivative of the tax function reduces the hourly wage, hours worked and labor earnings.

A government that is interested in redistribution can exploit this feature to boost wages of low-skilled workers. I characterize the allocation that maximizes welfare subject to resource and incentive constraints, and show that the tax schedule that implements it is concave at the bottom of the income distribution. Declining marginal tax rates at low earnings raise the labor supply elasticity of low-skilled workers, which increases their wages. The government sets the curvature of the tax function in such a way that the monopsonist does not extract any rents from hiring the least-skilled workers: if the tax system is optimized, the monopsonist pays the least-skilled workers a wage equal to their productivity. At the bottom of the income distribution, the optimal tax system thus offsets the negative impact of monopsony power on wages.

The finding that optimal marginal tax rates for low-skilled workers are declining in income is a local result. I also derive an optimal tax formula that holds at each point in the income distribution. To that end, I study the welfare effects of increasing the *second* derivative of the tax function just below a particular income level, and decreasing it right above. This reform induces a local increase in the marginal tax rate, cf. Saez (2001) and Golosov et al. (2014). The optimal tax formula is obtained by setting the sum of the welfare-relevant effects from this reform equal to zero. Compared to existing results from the literature, the additional sufficient statistics that characterize optimal tax policy are the impact of tax curvature on labor earnings and hourly wages, and the impact of the level and slope of the tax function on hourly wages and profits.

To illustrate the implications of the monopsonistic forces for optimal income taxes, I calibrate a structural version of the model to the US economy, assuming the latter is subject to a very high degree of monopsony power. At the optimal tax system, the government exploits curvature effects to boost the wages of low-skilled workers. It does so by letting marginal tax rates decline rapidly at low income levels. In addition, to alleviate the downward distortions from monopsony power on labor supply, most workers face a negative marginal tax rate, which never occurs if labor markets are competitive (cf. Mirrlees, 1971 and Saez, 2001). It should be noted, however, that the negative marginal tax rates are a consequence of imposing a degree of monopsony power that is much larger than in the actual economy, as in the model individuals can only work at one firm. If that assumption is relaxed, optimal marginal tax rates are unlikely to be negative.

A number of recent papers study optimal redistributive taxation in an environment where firms have market power, either in the market for goods (Kaplow, 2019, Boar and Midrigan, 2020, Kushnir and Zubrickas, 2020, Eeckhout et al., 2021, Gürer, 2021) or, as in the current paper, the market for labor (Hariton and Piaser, 2007, Cahuc and Laroque, 2014, da Costa and Maestri, 2019, Hummel, 2021). Cahuc and Laroque (2014) study the desirability of minimum wages alongside an optimal income tax if firms observe ability and labor supply responds on the extensive (participation) margin. By contrast, Hariton and Piaser (2007) and da Costa and Maestri (2019) characterize optimal tax policy in an environment where firms do not observe ability and workers supply labor on the intensive (hours, effort) margin. Hummel (2021) analyzes the implications of monopsony power for optimal income taxation and welfare if firms observe ability and labor supply responds on both the intensive and extensive margin. As in Hariton and Piaser (2007) and da Costa and Maestri (2019), firms offer workers contracts that specify labor effort and earnings. The key difference with these studies is that in the current paper, the monopsonist only sets hourly wages as individuals choose how many hours to work. Consequently, unlike in the aforementioned studies, labor market outcomes depend on the curvature of the tax function, which is the main focus of this paper.[1]

It is well known that measures of labor supply or earnings responses to wage or tax changes, such as the elasticity of taxable income (ETI), depend on the curvature of the tax function. See, among others, Saez (2001) and Jacquet and Lehmann (2021) for a discussion of this issue. A key difference is that in my model, the curvature of the tax function affects labor market outcomes directly and not only the behavioral responses to wage or tax changes. Slemrod and Kopczuk (2002) derive an optimal elasticity of taxable income in an environment where the government can affect this elasticity using administrative instruments. By contrast, in my model the government uses the curvature of the tax function to affect the elasticity of labor supply.

The wage responses to tax changes that modify optimal tax formulas are reminiscent of the literature that studies optimal income taxation with general equilibrium effects on wages. See, for example, Stiglitz (1982), Rothschild and Scheuer (2013) and Sachs et al. (2020). In these studies, labor markets are perfectly competitive and wages respond to tax changes because labor types are imperfect substitutes in production. By contrast, in my model labor types are perfect substitutes in production and wages respond to tax changes because firms have labor market power.

The remainder of this paper is organized as follows. Section 2 presents the model and analyzes the impact of tax curvature on labor market outcomes. Section 3 studies the implications for optimal marginal tax rates at the bottom. Section 4 derives an optimal tax formula that holds at each point in

[1] The only study I am aware of that also derives an optimal tax formula in a model where outcomes are affected by the curvature of the tax (or subsidy) schedule is Moore (2022), who characterizes the optimal tax schedule on political donations in a framework with political externalities.

the income distribution. Section 5 calculates and compares the optimal tax system assuming wages are either set by a monopsonist or determined competitively. Finally, Section 6 concludes.

2 Model

There is a continuum of individuals who differ in their ability $n \in [n_0, n_1]$ with $n_1 > n_0 > 0$. Ability measures how much output an individual produces per hour worked. The cumulative distribution of ability is denoted by $F(n)$ with density $f(n)$. Individuals supply labor on the intensive margin to a single monopsonist. The monopsonist observes ability and sets the hourly wage at each ability level to maximize profits. The government does not observe individual ability but only their labor earnings. It has a preference for redistribution and levies a nonlinear tax on labor earnings. To focus on labor income taxation, I assume profits are taxed at an exogenous rate and after-tax profits flow back to firm-owners, who do not play an important role in the analysis. The timing is as follows.

1. The government chooses the tax schedule $T(\cdot)$ on labor income $z(n) = w(n)l(n)$.
2. The monopsonist sets the hourly wage $w(n)$ at each ability level $n \in [n_0, n_1]$.
3. Individuals choose how many hours $l(n)$ to work.

Before characterizing the equilibrium, a few remarks are in place. First, the assumption of a single monopsonist that perfectly observes ability is of course extreme. However, it captures that (i) as in Mirrlees (1971), firms have an informational advantage compared to the government and (ii) the labor supply curve a firm faces is less than perfectly elastic (Manning, 2003). Second, despite that the firm is a monopsonist, it can only offer (linear) compensation contracts that specify an hourly wage, as individuals choose their own working hours.[2] This leads to outcomes that can be Pareto improved upon. By offering a combination of earnings and working hours, the monopsonist could generate higher profits without making individuals worse off. See Hariton and Piaser (2007), da Costa and Maestri (2019) and Hummel (2021) for an analysis of this case. If instead of setting an hourly wage, the monopsonist offers a combination of earnings and working hours, the tax curvature effects I focus on in this paper do not emerge.[3] While it seems plausible that firms can also influence working hours through other means than the hourly wage, I conjecture that the main effects and results derived below also hold in richer, more realistic environments, provided individuals have some say over their working hours and provided firms with labor market power have an informational advantage compared to the government about their workers' abilities.

Working backward, an individual with ability n takes the hourly wage $w(n)$ and the tax schedule $T(\cdot)$ as given. Preferences over consumption c and labor effort l are described by a separable utility

[2]A potential micro-foundation for these contracts is that firms might be restricted to pay a worker the same wage for every hour worked. For example, in the Netherlands, individuals can formally request to change the working hours stipulated in their employment contract. Employers can only deny this request if they have good cause (e.g., financial reasons). Also, in some sectors it is common that (weekly) hours worked are not constant. In both cases, labor income typically scales with the number of hours worked, which implies a constant hourly wage. The (legal) requirement that firms must pay a worker the same wage for every hour worked may appear inconsistent with the traditional assumption from Mirrlees (1971) that the government cannot observe wages. However, as explained by Lee and Saez (2012), the informational inconsistency is overcome if one assumes that hourly wages are observable after a costly audit following a whistle blowing claim.

[3]However, if individuals can adjust their working hours based on the hourly wage $w = z/l$ that is *implied* by the employment contract, the equilibrium is as described in the current paper. See also the previous footnote.

function $U(c, l) = u(c) - \phi(l)$, with $u'(\cdot), \phi'(\cdot), \phi''(\cdot) > 0$ and $u''(\cdot) \leq 0$. The individual chooses how many hours to work to maximize utility:

$$v(n) = \max_{l} \left\{ u(w(n)l - T(w(n)l)) - \phi(l) \right\}. \tag{1}$$

The first-order necessary condition determines the optimal choice of labor effort $l(n)$:

$$u'(w(n)l(n) - T(w(n)l(n)))w(n)(1 - T'(w(n)l(n))) = \phi'(l(n)). \tag{2}$$

At the optimum, the marginal benefits of working an extra hour (on the left-hand side) are equal to the marginal costs (on the right-hand side). In what follows, I assume the tax function is sufficiently smooth and such that the second-order condition for utility maximization is satisfied. Equation (2) then pins down the number of hours worked as a function of the hourly wage.

The monopsonist sets the wage $w(n)$ at each ability level in order to maximize profits. It treats the tax schedule $T(\cdot)$ as given, but has to take into account that wages affect hours worked, cf. equation (2). The monopsonist thus maximizes aggregate profits $\Pi = \int_{n_0}^{n_1} \pi(n) f(n) dn$, where the profits from employing a worker with ability n are

$$\pi(n) = \max_{w,l} \left\{ (n - w)l \quad \text{s.t.} \quad u'(wl - T(wl))w(1 - T'(wl)) = \phi'(l) \right\}. \tag{3}$$

Again, I assume the tax schedule is such that the second-order conditions are satisfied. Combining the first-order necessary conditions with respect to the wage and hours worked gives, after rearranging,

$$\frac{w(n)}{n} = \frac{e_{lw}(n)}{1 + e_{lw}(n)}. \tag{4}$$

The monopsonist marks down wages relative to productivity, which, through the labor supply equation (2), generates downward distortions in labor supply. How much wages are marked down relative to productivity and hence, how large the downward distortions from monopsony power on labor supply are depends on the (firm-level) elasticity $e_{lw}(n)$ of hours worked with respect to the hourly wage, which varies across the skill distribution. Ignoring function arguments to save on notation, the elasticity of labor supply can be found by implicitly differentiating the first-order condition (2):

$$e_{lw} = \frac{dl}{dw}\frac{w}{l} = \frac{w}{l} \frac{u''wl(1 - T')^2 + u'(1 - T') - u'wlT''}{-u''w^2(1 - T')^2 + \phi'' + u'w^2T''}. \tag{5}$$

Because there is a single monopsonist, the *firm-level* elasticity of labor supply is equal to the aggregate, or *market-level* elasticity of labor supply: both are given by e_{lw}.[4] As stated before, the assumption of a single monopsonist is of course extreme, but captures that the labor supply curve a firm faces is less than perfectly elastic (Manning, 2003).

The elasticity of labor supply (5) depends on properties of the tax function, in particular the level T (which enters u' and u''), the slope T' and the *curvature* T'', which shows up both in the numerator and denominator. This leads to the following result.

[4]By contrast, if labor markets are perfectly competitive, the market-level elasticity of labor supply is given by equation (5), whereas the firm-level elasticity of labor supply is infinite.

Proposition 1. *A local increase in the curvature of the tax function $T(\cdot)$ at earnings $z(n)$ leads to a lower equilibrium wage $w(n)$, hours worked $l(n)$ and labor earnings $z(n)$.*

Proof. See Appendix A. □

According to Proposition 1, a local increase in the *second* derivative of the tax function $T(\cdot)$ at earnings $z(n)$ reduces the hourly wage, hours worked and labor earnings of individuals with ability n. The reason is that an increase in the curvature of the tax function makes labor supply less responsive to a change in the hourly wage. To illustrate, suppose the tax schedule is convex and marginal tax rates are steeply increasing in income. In that case, individuals have weak incentives to work longer hours following an increase in the hourly wage. Put differently, the elasticity of labor supply e_{lw} is low. A low elasticity of labor supply makes it attractive for the monopsonist to pay low wages as well, cf. equation (4). Consequently, a local increase in the curvature of the tax function reduces the hourly wage, hours worked and labor earnings. The following example illustrates this.

Example 1. *Suppose the individual utility function is $U(c,l) = c - \frac{l^{1+1/\varepsilon}}{1+1/\varepsilon}$ and the tax schedule has a constant rate of progressivity $p \in (0,1)$: $T(z) = z - \frac{1-\tau}{1-p} z^{1-p}$. Equilibrium labor supply follows from equation (2): $l(n) = (1-\tau)^{\frac{\varepsilon}{1+p\varepsilon}} w(n)^{\frac{(1-p)\varepsilon}{1+p\varepsilon}}$. Hence, the elasticity of labor supply is $e_{lw} = \frac{(1-p)\varepsilon}{1+p\varepsilon}$ and the equilibrium wage is, cf. equation (4):*

$$w(n) = (1-p)\frac{\varepsilon}{1+\varepsilon} n. \tag{6}$$

A higher rate of tax progressivity $p \in (0,1)$ means that marginal tax rates are more quickly increasing in income. This makes it less attractive for individuals to work longer hours if their wage increases: the elasticity of labor supply e_{lw} is decreasing in p. An increase in the rate of tax progressivity thus amplifies the negative impact of monopsony power on wages and labor supply by making individuals less responsive to wage changes.

For many countries, the tax schedule is not smooth as in the previous example, but piecewise linear, which implies the tax curvature T'' is either zero or undefined.[5] As is well understood for the case where labor markets are competitive, a convex kink in the tax schedule (i.e., rising marginal tax rates) leads to bunching, whereas a concave kink (i.e., falling marginal tax rates) generates a hole in the income distribution (Kleven, 2016). In the former case, individuals whose wages fall in a certain range choose the same income. It is then optimal for the monopsonist to pay some individuals with different skills the lowest wage that induces them to bunch at the income level where the marginal tax rate rises. Hence, with a monopsonist, there is bunching not only in the income, but also in the *hourly wage* distribution. Conversely, if there is a concave kink in the tax schedule, for a certain range of wages, a small increase in the wage leads to a discrete jump in working hours. The monopsonist then finds it optimal to pay some individuals with different skills a high wage. This generates a bunch in the hourly wage and income distribution, right after the hole. Hence, while the result from Proposition 1 is derived assuming the tax function is smooth, a similar logic holds for piecewise linear taxes. Specifically, if the tax schedule has a convex (concave) kink, the equilibrium wage for some individuals is lower (higher) than would be the case in the absence of taxation. Appendix F illustrates this point formally by working out a case similar to Example 1 where the tax schedule is piecewise linear.

[5] A noteworthy exception is Germany, where *marginal* tax rates increase linearly with income from 14% to 24% and then from 24% to 42%. This implies that over certain ranges, the tax function is quadratic in income.

The key reason why tax curvature affects labor market outcomes is that individuals can adjust their working hours, and the monopsonist takes this into account when setting wages. If, by contrast, labor supply responds only on the extensive, or participation margin (so that working hours are fixed, but a monopsonist attracts more employees when it raises wages), then a change in tax curvature does not affect equilibrium wages. This is because participation responses are driven by average, and not marginal tax rates. While the extensive (participation) margin is the most relevant one especially for low-income workers, the literature consistently finds small but positive hours responses as well (see Evers et al. (2008), Chetty (2012) and Bargain and Peichl (2016) for recent surveys). In a hybrid model with labor supply responses on both margins, the result from Proposition 1 remains valid provided individuals can adjust their working hours. Changes in equilibrium wages, in turn, also generate participation responses. Through its impact on wages, a local increase in tax curvature could therefore also result in lower labor participation.[6]

The effect described in Proposition 1 differs from the wage-moderating effect of tax progressivity. The latter states that a local increase in the *marginal* tax rate lowers the equilibrium wage. This is a robust prediction in models where labor markets are imperfectly competitive: it holds in the context of union bargaining (Hersoug, 1984), search frictions (Pissarides, 1985) and efficiency wages (Pisauro, 1991). The main difference with Proposition 1 is that the latter concerns the impact of the second (as opposed to the first) derivative of the tax function. In the aforementioned studies, labor market outcomes only depend on the level and slope of the tax function. Consequently, a local change in the curvature of the tax function that leaves the tax burden and marginal tax rate unaffected has no impact. By contrast, a change in the curvature does affect labor market outcomes if, as in my model, a monopsonist sets hourly wages and individuals choose how many hours to work.

Before turning to the optimal tax problem, it is useful to highlight the following result.

Lemma 1. *At an interior solution characterized by the first-order conditions (2) and (4), hourly wages* $w(n)$*, hours worked* $l(n)$ *and labor earnings* $z(n) = w(n)l(n)$ *are increasing in ability* n*.*

Proof. See Appendix B. □

In words, for a given nonlinear tax schedule $T(\cdot)$, individuals with a higher ability n earn a higher hourly wage $w(n)$ and work longer hours $l(n)$. These observations imply that labor earnings $z(n)$ are increasing in ability n as well.

The government taxes profits at an exogenous rate $\tau \in [0, 1]$ and chooses the tax schedule $T(\cdot)$ on labor earnings to maximize social welfare

$$\mathcal{W} = \int_{n_0}^{n_1} \Psi(v(n), n) f(n) dn + \omega(1 - \tau) \int_{n_0}^{n_1} \pi(n) f(n) dn, \tag{7}$$

where the function $\Psi(v, n)$ is increasing and weakly concave in its first argument, and weakly declining in its second argument: $\Psi_v(v, n) > 0$, $\Psi_{vv}(v, n), \Psi_n(v, n) \leq 0$. Together with the concavity in the individual utility function $u(\cdot)$, this function determines the government's preferences for redistributing income from high-skilled to low-skilled workers. After-tax profits flow back to firm-owners

[6] An extension of the model where individuals supply labor on both the intensive and extensive margin, which also studies the implications for optimal income taxation, is available on request.

and receive a weight of $\omega \geq 0$ in the welfare function.[7] For simplicity, I assume the weight is such that transferring income from the government budget to firm-owners does not raise welfare. The government chooses the tax schedule $T(\cdot)$ to maximize social welfare (7), taking into account how a change in the tax function affects labor market outcomes and subject to the budget constraint

$$\int_{n_0}^{n_1} T(z(n))f(n)dn + \tau \int_{n_0}^{n_1} \pi(n)f(n)dn = G, \tag{8}$$

where G denotes some exogenous spending.

The literature offers two methods to solve an optimal tax problem of this kind: the mechanism design approach (see, e.g., Mirrlees, 1971) and the tax perturbation approach (see, e.g., Saez, 2001, Golosov et al., 2014 and Gerritsen, 2023). See Jacquet and Lehmann (2021) for a formal discussion and comparison of both methods. In what follows, I use the mechanism design approach to derive properties of marginal tax rates at the bottom of the income distribution (Section 3) and the tax perturbation approach to derive an optimal tax formula in terms of sufficient statistics that holds at each point in the income distribution (Section 4).

3 Declining marginal tax rates

Solving the optimal tax problem using the mechanism design approach requires finding the allocation $(v(n), \pi(n), l(n))$ for each $n \in [n_0, n_1]$ that maximizes welfare (7), subject to resource and incentive constraints.[8] The resource constraint is obtained by inverting the relationship $v(n) = u(z(n) - T(z(n))) - \phi(l(n))$ with respect to $T(z(n))$ and using the property $z(n) = nl(n) - \pi(n)$. Substituting these in the government budget constraint (8) gives

$$\int_{n_0}^{n_1} \Big[nl(n) - u^{-1}(v(n) + \phi(l(n))) - (1-\tau)\pi(n) \Big] f(n)dn = G. \tag{9}$$

The incentive constraints, in turn, describe the optimizing behavior of the monopsonist and individuals. To derive the first of these, differentiate the expression for profits (3) with respect to ability n and apply the envelope theorem:[9]

$$\pi'(n) = l(n). \tag{10}$$

[7] A potential micro-foundation is that there exists a group of identical firm-owners, who do not exert effort and whose utility is linear in their net income.

[8] This approach relies on the taxation principle, which states that an incentive compatible allocation can be implemented using a nonlinear tax on labor income. See, e.g., Hammond (1979). While formally establishing this result in the current environment with non-competitive labor markets is beyond the scope of this paper and an interesting topic for future research, I verify numerically in Section 5 that the necessary and sufficient conditions for implementation are satisfied.

[9] To see this, write the Lagrangian of the firm's problem as

$$\Lambda(n) = (n-w)l + \lambda\Big[u'(wl - T(wl))w(1 - T'(wl)) - \phi'(l)\Big],$$

where λ is the multiplier on the constraint (2). By the envelope theorem, $\pi'(n) = \Lambda'(n) = l(n)$.

To derive the second incentive constraint, differentiate the expression for individual utility (1) and again apply the envelope theorem:

$$\begin{aligned} v'(n) &= u'(w(n)l(n) - T(w(n)l(n)))l(n)(1 - T'(w(n)l(n)))w'(n) \\ &= \phi'(l(n))l(n)\frac{w'(n)}{w(n)} = \phi'(l(n))\left(\frac{\pi(n)}{nl(n) - \pi(n)}\right)b(n), \text{ where } b(n) = l'(n). \end{aligned} \tag{11}$$

The second step uses equation (2) and the final step uses the relationship $\pi(n) = (n - w(n))l(n)$.[10]

As will be made clear below, in order to find the allocation that maximizes social welfare subject to resource and incentive constraints, it is important to take the constraint $\pi(n_0) \geq 0$ explicitly into account. Equation (10) then implies the monopsonist makes non-negative profits from hiring each worker. Lastly, to make sure that the allocation can be implemented using a non-linear tax $T(z(n))$, labor earnings must satisfy the monotonicity constraint $z'(n) \geq 0$. Differentiating $z(n) = nl(n) - \pi(n)$ and imposing equation (10), the latter requires $b(n) = l'(n) \geq 0$. According to Lemma 1, this condition is satisfied at any interior solution described by the labor market equilibrium conditions (2) and (4).[11]

The problem of finding the optimal tax schedule $T(\cdot)$ can now be written as an optimal control problem, with state variables $(v(n), \pi(n), l(n))$ and control variable $b(n)$. Appendix C formulates the problem and derives the first-order necessary conditions. This leads to the following result.

Proposition 2. *If $l(n_0) > 0$, optimal marginal tax rates for the least-skilled workers are declining in income. Put differently, the optimal tax schedule is concave at the bottom of the income distribution: $T''(z(n_0)) < 0$.*

Proof. See Appendix C. □

The main insight from Proposition 2 is that the government can use the curvature of the tax function to boost wages of low-skilled workers. This is achieved by setting a tax schedule that is concave at the bottom of the income distribution. A more concave tax schedule, i.e., a decrease in the second derivative of the tax function, positively affects wages: see Proposition 1. Declining marginal tax rates make it attractive for individuals to work longer hours following an increase in the hourly wage. A high elasticity of labor supply, in turn, induces the monopsonist to pay high wages as well, cf. equation (4). A government that is interested in redistribution can exploit this feature to raise the wages of low-skilled workers, and finds it optimal to do so.

At the optimal tax system, the monopsonist does not extract any rents from hiring the least productive workers: $\pi(n_0) = 0$. Provided $l(n_0) > 0$, these workers get paid a wage equal to their productivity, despite that there is a single monopsonist.[12] The optimal tax system thus offsets any negative impact from monopsony power on the wages of the least-skilled workers. To achieve this, the government sets the curvature of the tax function at the bottom of the income distribution in such a way that labor supply of the least-skilled workers becomes infinitely elastic: $e_{lw}(n_0) \to \infty$. Doing so requires that marginal tax rates at the bottom are declining in income: $T''(z(n_0)) < 0$.

[10] Differentiate both sides with respect to ability to find $\pi'(n) = l(n) + (n - w(n))l'(n) - w'(n)l(n)$. Equation (10) then implies $w'(n)l(n) = (n - w(n))l'(n)$.

[11] These conditions are only necessary but not sufficient. Sufficiency can be checked *ex post* after having numerically solved for the optimal tax system.

[12] The property $l(n_0) > 0$ does not necessarily hold at the optimum. This happens if the non-negativity constraint $l(n) \geq 0$ is binding at low ability levels. For analytical convenience, I focus on interior solutions where $l(n_0) > 0$.

In Appendix C, I also show that the optimal marginal tax rate at the lowest income level is positive: $T'(z(n_0)) > 0$. Hence, despite that the government offsets the negative impact of monopsony power on the wages for the least-skilled workers, their labor supply is still downward distorted at the optimum due to a positive marginal tax rate. By discouraging labor supply, the marginal tax rate negatively affects profits. A positive marginal tax rate at the bottom therefore contributes to relaxing the constraint $\pi(n_0) \geq 0$, which is binding at the optimal tax system. The benefits of relaxing this constraint are then weighted against the costs of distorting labor supply for the least-skilled workers. Combined with the result from Proposition 2, this means that marginal tax rates at the lowest income levels are positive but declining. This contrasts the behavior of optimal marginal tax rates at the bottom if labor markets are competitive. In that case, a well-known result is that the optimal marginal tax rate is zero at the bottom provided $z(n_0) > 0$ and there is no mass point (Seade, 1977), and positive at interior earnings. From these observations it follows that the optimal tax schedule is convex at the bottom if labor markets are competitive, whereas it is concave if wages are set by a monopsonist.

The finding from Proposition 2 that optimal marginal tax rates are declining in income is a local result: it holds at the bottom of the income distribution. Unfortunately, using the mechanism design approach to derive optimal tax rules that that hold at each point in the income distribution and that can be meaningfully interpreted turns out to be particularly challenging. The next section derives such a result using an alternative approach to solving the optimal tax problem.

4 Optimal tax formula

The optimal tax problem can also be solved using the tax perturbation approach. See Saez (2001) and Golosov et al. (2014), among many others. The idea behind this approach is to study a perturbation, or reform of the nonlinear tax schedule $T(\cdot)$. Such a reform induces welfare-relevant effects. Optimal tax formulas can then be derived from the requirement that if the tax schedule is optimized, the welfare-relevant effects of the reform sum to zero.

An important advantage of the tax perturbation approach is that it allows for a derivation of optimal tax formulas in terms of sufficient statistics (Chetty, 2009). In the current setting, these are: i) the income distribution, ii) behavioral responses and iii) welfare weights. Starting with the first, let $H(z)$ denote the cumulative distribution of earnings, with corresponding density $h(z)$. The behavioral responses, in turn, capture how an increase in the level, slope or curvature of the tax function affect labor market outcomes. With a slight abuse of notation, I denote by y_x the impact of a local increase in $x \in \{T, T', T''\}$ on outcome $y \in \{z, l, w, \pi\}$. Lastly, the welfare weight of an individual with earnings z and firm-owners is

$$g(z) = \frac{1}{\eta}\Psi_v(v(\hat{n}(z)), \hat{n}(z))u'(z - T(z)), \quad g_f = \frac{\omega}{\eta}. \tag{12}$$

where η is the multiplier on the government budget constraint and $\hat{n}(z)$ denotes the ability level that corresponds to earnings z. In words, the welfare weights $g(z)$ and g_f measure by how much social welfare increases if an individual with earnings z and firm-owners receive an additional unit of after-tax income. These weights summarize in a reduced-form way the government's preferences for redistribution. Because $\Psi(\cdot)$ and $u(\cdot)$ are concave, the welfare weights $g(z)$ are declining in income z. Moreover, from the assumption that transferring income from the government budget to firm-owners does not raise welfare, it follows that $g_f \leq 1$.

The next Proposition states an optimal tax formula in terms of sufficient statistics.

Proposition 3. *At the optimal tax system, the following condition must hold at each point $z' \in [z_0, z_1]$ in the income distribution:*

$$0 = -\frac{\partial}{\partial z}\left[\Big(T'(z')z_{T''}(z') + g(z')(1-T'(z'))\hat{l}(z')w_{T''}(z')\Big)h(z')\right] \tag{13}$$
$$+ \Big(T'(z')z_{T'}(z') + (g_f(1-\tau)+\tau)\pi_{T'}(z') + g(z')(1-T'(z'))\hat{l}(z')w_{T'}(z')\Big)h(z')$$
$$+ \int_{z'}^{z_1} \Big(1 - g(z) + T'(z)z_T(z) + (g_f(1-\tau)+\tau)\pi_T(z) + g(z)(1-T'(z))\hat{l}(z)w_T(z)\Big)h(z)dz.$$

Here, it is explicitly taken into account which terms vary along the income distribution and $\hat{l}(z)$ denotes the labor effort associated with earnings z.

Proof. See Appendix D. □

To understand this result, Figure 1 graphically illustrates the reform that gives rise to optimal tax formula (13). The black, dotted line shows the original tax schedule. For simplicity, it is drawn as a straight line. The red, solid line shows the perturbed tax schedule after the reform is implemented. Below earnings z', the two tax functions are the same. In the small interval $[z', z'+\zeta]$, the government increases the curvature (i.e., the second derivative) of the tax function by dT''. This is shown by the convex part of the solid line. Following the increase in tax curvature, the marginal tax rates above earnings $z'+\zeta$ increase by an amount equal to $dT' = dT''\zeta$. As a result, the perturbed tax schedule is steeper than the original tax schedule: see Figure 1. In the interval $[z'+\delta, z'+\delta+\zeta]$ with $\delta \gg \zeta$, the government reverses the increase in the curvature by lowering the second derivative of the tax function by dT''. This is shown by the concave part of the solid line. Following this reversal, the marginal tax rates of the perturbed and original tax schedule are the same at earnings above $z'+\delta+\zeta$: the dotted and solid line are parallel. However, the reform does increase the tax burden for individuals with earnings above this level by an amount equal to $dT = dT'\delta = dT''\zeta\delta$.[13]

The reform graphically illustrated in Figure 1 generates three types of welfare-relevant effects. First, there are behavioral responses due to a change in the curvature of the tax function in the small intervals $[z', z'+\zeta]$ and $[z'+\delta, z'+\delta+\zeta]$. In Figure 1, these are labeled 'curvature effects'. A change in tax curvature affects earnings z and wages w for individuals in these intervals, as captured by $z_{T''}$ and $w_{T''}$.[14] Earnings responses have an impact on government finances proportional to the marginal tax rate T', whereas wage responses have an impact on individual utilities proportional to the number

[13]Let $\rho(z)$ denote the difference between the perturbed and the original tax schedule. A formal definition of the tax reform graphically illustrated in Figure 1 is

$$\rho(z) = \begin{cases} 0 & \text{if } z \leq z', \\ \frac{1}{2}dT''(z-z')^2 & \text{if } z \in (z', z'+\zeta], \\ -\frac{1}{2}dT''\zeta^2 + dT''\zeta(z-z') & \text{if } z \in (z'+\zeta, z'+\delta], \\ -\frac{1}{2}dT''\zeta^2 + dT''\zeta(z-z') - \frac{1}{2}dT''(z-(z'+\zeta))^2 & \text{if } z \in (z'+\delta, z'+\delta+\zeta], \\ dT''\zeta\delta & \text{if } z > z'+\delta+\zeta. \end{cases}$$

[14]A change in tax curvature also affects hours worked (cf. Proposition 1), but by the envelope theorem, these responses have no first-order impact on individual utilities and hence, welfare. Moreover, a change in tax curvature does not affect

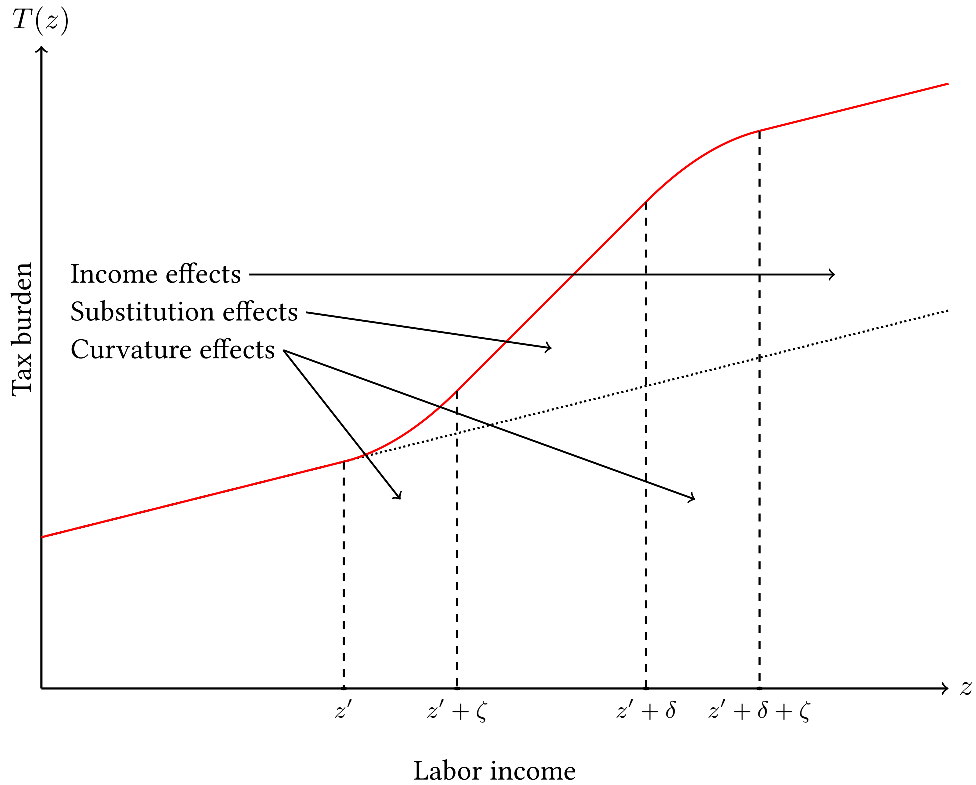


Figure 1: Tax reform and behavioral responses

of hours worked l and the net-of-tax rate $1 - T'$. To obtain the impact on social welfare, these wage responses are multiplied by the social welfare weights g. The first term on the right-hand side of equation (13) then captures the difference between the welfare effects of increasing the tax curvature and subsequently decreasing it, cf. Figure 1. Upon making the intervals arbitrarily small, this boils down to taking a derivative of the welfare impact with respect to income.

Second, the tax reform raises the marginal tax rate in the interval $[z' + \zeta, z' + \delta]$: see Figure 1 and recall that $\delta \gg \zeta$. A higher marginal tax rate not only affects earnings z and wages w, but also profits π. Figure 1 refers to these as 'substitution effects'. As before, changes in earnings affect government finances whereas changes in wages have a direct impact on individual utilities. Profit responses, in turn, affect both government finances and the utilities of firm-owners. To obtain the impact on welfare, the first of these effects is multiplied by τ and the second by $g_f(1 - \tau)$. The second line of equation (13) sums the welfare-relevant effects associated with a local increase in the marginal tax rate. These effects are proportional to the density of the income distribution at the point where the marginal tax rate is increased.

Third, the reform increases the tax burden for individuals with earnings above $z' + \delta + \zeta$. This generates two types of welfare-relevant effects. First, the reform mechanically transfers income from individuals with earnings above this level to the government. The net welfare impact is captured by $1 - g$. Second, a change in the tax burden generates behavioral responses on earnings z, wages w and

profits. To see this, note that profits can be written as

$$\pi = \max_{l,z} \left\{ nl - z \quad \text{s.t.} \quad u'(z - T(z))z(1 - T'(z)) = \phi'(l)l \right\},$$

which depends on the level and slope of the tax function, but not on its second derivative.

profits π. Figure 1 labels these 'income effects'. The third line of equation (13) integrates the welfare impact of a higher tax burden over all individuals who see their tax burden increase as a result of the reform.

If the tax system optimized, any reform should leave welfare unaffected. Equating to zero the sum of the welfare effects from the reform graphically illustrated in Figure 1 leads to equation (13). This optimal tax formula can be contrasted with the one that holds if labor markets are competitive. See, for instance, Saez (2001) and Golosov et al. (2014). Under perfect competition, labor market outcomes are not affected by a local increase in the curvature of the tax function: $z_{T''} = w_{T''} = 0$. Moreover, the wage of an individual with ability n is $w(n) = n$ and firms make zero profits: $\pi(n) = 0$. Consequently, wages and profits do not respond to a change in the level or slope of the tax function: $w_T = w_{T'} = \pi_T = \pi_{T'} = 0$. Equation (13) then becomes:

$$0 = T'(z')z_{T'}(z')h(z') + \int_{z'}^{z_1} \Big(1 - g(z) + T'(z)z_T(z)\Big)h(z)dz. \tag{14}$$

Apart from differences in presentation, this optimal tax formula coincides with the one derived in Saez (2001) and Golosov et al. (2014), among others.

Compared to the competitive benchmark, the additional sufficient statistics that show up in the optimal tax formula (13) are the effects of tax curvature on earnings and hourly wages, and the effects of the tax burden and the marginal tax rate on hourly wages and profits. Intuitively, the additional responses of earnings and profits to tax changes have budgetary effects, as the government taxes both labor income and profits. In equation (13), these effects are captured by $T'z_{T''}$, $\tau\pi_{T'}$ and $\tau\pi_T$. Profit responses also impact the net incomes of firm-owners, with a welfare impact equal to $g_f(1-\tau)\pi_{T'}$ and $g_f(1-\tau)\pi_T$. The wage responses $w_{T''}$, $w_{T'}$ and w_T, in turn, have an effect on workers' disposable incomes. This effect is proportional to l, the number of hours worked, and the after-tax rate $1 - T'$. Because changes in disposable income affect individual utilities, the welfare effects associated with wage changes in equation (13) are weighted by the individual welfare weights g.

As it turns out, implementing the optimal tax formula (13) using estimates of sufficient statistics to calculate optimal taxes is a very challenging task. To the best of my knowledge, there are no estimates available of the impact of tax curvature on labor market outcomes. Moreover, to implement equation (13), one also requires knowledge of how these statistics vary across the earnings distribution. An alternative approach to characterizing the optimal tax system is to calibrate a structural version of the model, which can subsequently be used to solve numerically the optimal tax problem that is described in Section 3 and formally defined in Appendix C. The resulting tax system can then be compared to the one that maximizes welfare under the assumption that labor markets are perfectly competitive, which gives rise to optimal tax formula (14). The next section takes this approach.

5 Numerical simulations

This section calculates the optimal tax system in a structural version of the model that is calibrated to the US economy, and compares it to the optimal tax system in the competitive benchmark. The goal is to study the shape of the optimal tax schedule, and to investigate to what extent it matters for optimal income taxation if wages are set by a monopsonist or determined competitively. It should be emphasized, however, that these simulations only serve an illustrative purpose. The reason is that, as mentioned before, the model is highly stylized in the sense that there is a single monopsonist. Consequently, the degree of monopsony power in the calibrated economy is much larger than in the actual economy. Keeping this caveat in mind, I start by briefly describing the data and calibration, and then turn to the optimal tax results.

The model is calibrated on the basis of US data. To that end, I use the March release of the 2018 Current Population Survey (CPS). For a large number of individuals, the CPS provides detailed information on income and taxes. In the final sample, I include individuals between 25 and 65 years old, whose hourly wage exceeds half the federal minimum wage of $7.25 and whose annual income from wage and salary payments is at least $11,419.[15] I use the income from wage and salary payments as the empirical counterpart of z and multiply incomes that are top-coded by a factor of 2.67, based on an estimate of the Pareto parameter of 1.6 for top labor incomes (Saez and Stantcheva, 2018).[16]

To calibrate the model, I require a specification of preferences and the current tax schedule. Following Saez (2001), the latter is approximated by a linear function

$$T(z) = -g + tz, \tag{15}$$

where g is a lump-sum transfer and t is the constant marginal tax rate. Values for $g = \$4,818$ and $t = 0.33$ are obtained by regressing the sum of federal and state taxes on taxable income. Regarding preferences, I use the specification from Example 1:

$$U(c, l) = c - \frac{l^{1+1/\varepsilon}}{1 + 1/\varepsilon}. \tag{16}$$

From equation (5), it follows that with a linear tax function, ε equals the elasticity of labor supply e_{lw}. The latter determines the responsiveness of hours worked to changes in the net wage, both at the market-level and, because there is a single monopsonist, at the firm-level. When selecting a value of ε, a challenge is that the small values for the elasticity of labor supply typically found using micro data imply, through the lens of this model, very large markdowns and implausibly high profits relative to wages.[17] The reason is that individuals can only work at one firm, which implies a very substantial degree of monopsony power. Because in the model there is a direct link between the market-level elasticity of labor supply and the markdown of wages relative to productivity, I decide to split the difference and select a value of $\varepsilon = 1$.[18] This value is larger than what is commonly used in optimal tax models, but still implies a substantial markdown of $1 - \varepsilon/(1 + \varepsilon) = 0.5$.

With the above tax schedule and utility function, it is possible to calculate the ability n for each

[15] This is the income of an individual who earns the minimum wage and works full-time (on average 35 hours for 45 weeks). Cutting off a smaller part of the income distribution by choosing a lower threshold results in a smaller value of n_0. In that case, the non-negativity constraint $l(n) \geq 0$ in the optimal tax problem might become binding at low ability levels. However, the result from Proposition 2 requires $l(n_0) > 0$. Therefore, I prefer to work with a formulation of the optimal tax problem that generates a solution where labor supply at the bottom is positive.

[16] If incomes at the top are Pareto distributed, the expected value above a threshold z' is $E[z|z \geq z'] = \tilde{a}z'/(\tilde{a} - 1)$, where $\tilde{a} = 1.6$ is an estimate of the tail parameter.

[17] For example, a value of $\varepsilon = 0.33$, as recommended by Chetty (2012), would mean that wages are marked down $1 - \varepsilon/(1 + \varepsilon) = 75\%$ relative to productivity, which implies profits are three times as large as wages.

[18] This tight link can be relaxed, for instance, by introducing multiple firms that engage in Cournot competition in the labor market, as in Berger et al. (2022), or by introducing an extensive margin of labor supply. In both cases, the elasticity of labor supply that is relevant for the firm is larger than the elasticity of hours worked to a change in the wage.

individual from their income from wage and salary payments z, using the first-order conditions (2) and (4). In the spirit of Saez (2001), this gives an ability distribution that is consistent with the empirical income distribution. To smooth the distribution, I estimate a kernel density, and append a Pareto tail starting at an income level of \$350,000. The tail parameter of the ability distribution is chosen in such a way that the resulting income distribution has a tail parameter of 1.6 and the scale parameter is set to prevent a jump in the density at the point where the Pareto tail is pasted.[19]

Turning to the government, the social welfare function is

$$\Psi(v(n), n) = \tilde{g}(n)v(n), \qquad \tilde{g}(n) = \zeta n^{-\beta}, \tag{17}$$

where $\tilde{g}(n)$ denotes the welfare weight attached to individuals with ability n, $\beta \geq 0$ determines how quickly these weights are declining in ability, and $\zeta > 0$ is set in such a way that the average welfare weight equals one. In the benchmark, I select a value of $\beta = 0.5$. Combined with the specification $u(c) = c$, this value implies the government has rather modest preferences for redistribution. The reason for making this choice is that with stronger inequality aversion, the constraint $l(n) \geq 0$ in the optimal tax problem binds at low ability levels, while the monotonicity constraint $z'(n) \geq 0$ binds for a non-negligible share of high-ability individuals.[20] Furthermore, I set the welfare weight of firm-owners to $g_f = 1$ to make sure the government does not wish to redistribute on average from firm-owners to workers or *vice versa*. Lastly, I assume profits are taxed at an exogenous rate of $\tau = 0.36$, based on Trabandt and Uhlig (2011). The government budget constraint (8) then implies a value of $G = \$42,228$.[21]

With the calibrated set of primitives, I numerically solve the optimal tax problem formally defined in Appendix C, using the GPOPS-II software developed by Patterson and Rao (2014). The red, solid line in Figure 2 plots the marginal tax rates against current income levels below \$200,000 if the degree of inequality aversion is $\beta = 0.5$. Optimal marginal tax rates start out very high at 94%, but then decrease rapidly. This confirms the result from Proposition 2 that the optimal tax schedule is concave at the bottom of the income distribution, and that individuals with the lowest income face a positive marginal tax rate. Optimal marginal tax rates continue to decline, and are *negative* for individuals whose current labor income exceeds \$18,800. This property can be formally established at the top (see equation (54) in Appendix D), but it turns out that the vast majority of individuals receive a subsidy on the last dollar earned if the tax schedule is optimized. For individuals whose current labor earnings exceed \$86,000, optimal marginal tax rates are roughly constant at -70%. These negative marginal tax rates imply substantial *upward* redistribution toward high-income individuals.

To understand the shape of the optimal tax schedule, suppose the government only cares about efficiency: $\beta = 0$. It then wishes to implement an efficient allocation where, for each individual,

[19] With the current tax function and preference specification, the tail parameter of the ability distribution, a, is linked to the tail parameter of the earnings distribution, $\tilde{a}$, through $a = \tilde{a}(1 + \varepsilon)$.

[20] With $\beta = 0.5$, the monotonicity constraint still binds for some ranges that correspond to current incomes above \$419,000, which is the case for less than 0.5% of individuals in the final sample. This fraction increases if inequality aversion β increases, but also if the labor supply elasticity ε is reduced.

[21] This large value of spending means the government needs to generate substantial revenues. Consequently, at the optimal tax system, the consumption of individuals whose current labor earnings are below approximately \$13,300 is actually *negative*. While this does not violate any of the optimality conditions from Appendix C, this is of course not feasible in practice. However, this issue can easily be resolved by choosing a lower value of G or by letting after-tax profits flow back uniformly to workers instead of an exogenous group of firm-owners. For this reason, I suffice with this note, cautioning against interpreting the results for consumption implied by the optimal tax system in the calibrated economy.

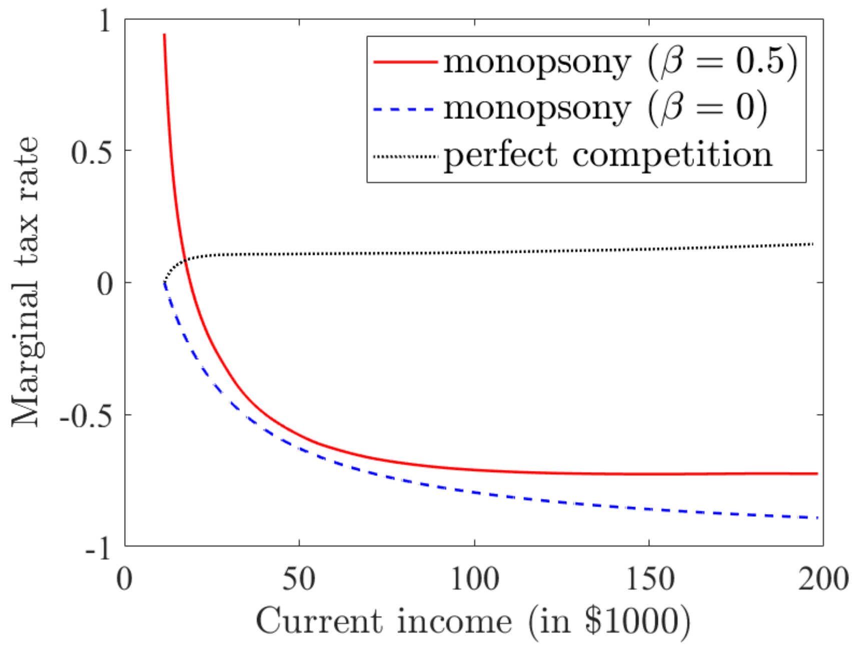


Figure 2: Optimal marginal tax rates

the marginal productivity of labor is equated to their marginal disutility of working an extra hour: $\phi'(l(n)) = l(n)^{1/\varepsilon} = n$, or $l(n) = n^{\varepsilon}$. Given the specification of preferences (16), the government can achieve this by setting a tax schedule

$$T(z) = A + B\ln(z) - \frac{1}{\varepsilon}z, \tag{18}$$

where A and B are constants that should jointly ensure the government budget constraint (8) is satisfied.[22] This requirement does not uniquely pin down these constants. For example, the government can implement an efficient allocation with a *linear* tax schedule by setting $B = 0$. A constant subsidy rate of $-T'(z) = 1/\varepsilon$ on labor income offsets the negative impact of monopsony power on labor supply. This is, however, not the only tax schedule that implements an efficient allocation. While labor effort $l(n) = n^{\varepsilon}$ is the same in all cases, the choice of B affects wages $w(n)$ and profits $\pi(n)$, which determine how quickly payoffs $v(n)$ are increasing in ability: see equation (11). Selecting the tax function that generates the lowest inequality in payoffs $v(n)$ requires setting B such that the monopsonist does not make profits from hiring the least productive workers: $\pi(n_0) = 0$. These workers then receive a wage equal to their productivity: $w(n_0) = n_0$. The blue, dashed line in Figure 2 shows the corresponding tax schedule. Among the tax schedules that implement an efficient allocation, this is the one that maximizes the payoff $v(n_0)$ for the least-skilled workers. Marginal tax rates start at zero, then decline according to equation (18), and converge to a level of $-1/\varepsilon = -1$, or -100%.

The optimal tax schedule with inequality aversion $\beta = 0.5$ inherits some of these properties. In particular, most individuals face a negative marginal tax rate. This largely reflects efficiency considerations. Subsidies on labor income alleviate the downward distortions in labor supply that occur because

[22] To verify this tax function implements an efficient allocation, substitute equation (18) in equation (32) from Appendix D, and use the property that utility is linear in consumption: $u(c) = c$. It follows directly that labor effort is pinned down at its efficient level: $l(n) = n^{\varepsilon}$.

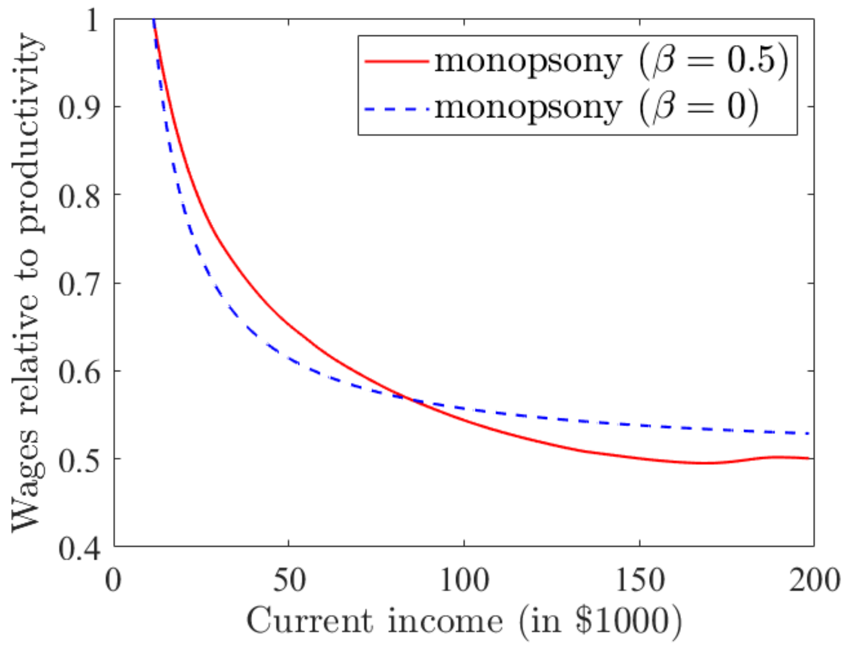


Figure 3: Optimal markdowns

the monopsonist marks down wages relative to productivity. The desire to boost wages of low-skilled workers for equity considerations, in turn, explains why marginal tax rates are rapidly declining at the bottom, i.e., why the tax schedule is concave at low income levels.

The most important difference between the red, solid and blue, dashed line from Figure 2 is that optimal marginal tax rates are higher if the government has a preference for redistribution ($\beta = 0.5$). In addition, marginal tax rates are positive at low income levels, which does not occur if the government is only concerned with efficiency ($\beta = 0$). This reduces the labor income subsidies paid to high-ability individuals and lowers the tax burden for individuals with lower incomes, which is desirable for redistributive reasons. Furthermore, marginal tax rates decline more (less) rapidly at low (high) incomes if the degree of inequality aversion is $\beta = 0.5$. By exploiting curvature effects, the government reduces wage inequality, thereby effectively engaging in 'predistribution' (i.e., equalizing the distribution of hourly wages). Figure 3 illustrates this by plotting hourly wages relative to productivity, $w(n)/n$. Under both tax schedules, the least-skilled workers receive a wage equal to their productivity: $w(n_0) = n_0$. However, wages are higher at low incomes if the degree of inequality aversion is $\beta = 0.5$, and *vice versa* at higher incomes.

The optimal tax schedule if wages are set by a monopsonist looks very different than if wages are determined competitively, as in Mirrlees (1971) and Saez (2001). To illustrate this, I recalibrate the ability distribution and government spending G assuming workers receive a wage equal to their productivity, and profits are zero: $w(n) = n$ and $\pi(n) = 0$. The value for the lump-sum transfer g and tax rate t at the current tax system, and the labor supply elasticity ε are the same as before. I subsequently solve the optimal tax problem assuming labor markets are competitive, which is formally defined in Appendix E. The black, dotted line in Figure 2 shows the results. Optimal marginal tax rates start at zero cf. Seade (1977), quickly increase to 11%, and then increase slowly to around 25% in the Pareto tail (not plotted).[23] The reason why optimal marginal tax rates are quite low is that the government does not have a strong preference for redistribution ($u(c) = c$ and $\beta = 0.5$) and labor supply is responsive to changes in the after-tax wage ($\varepsilon = 1$).

The finding that if wages are set by a monopsonist, optimal marginal tax rates are negative for most workers and can be as low as -70%, which never occurs if labor markets are competitive, largely reflects that the model from Section 2 is highly stylized. Because individuals can only work at one firm, the degree of monopsony power in the calibrated economy is very high, and much larger than in the actual economy. As a result, wages are marked down substantially relative to productivity. This generates large downward distortions in labor supply, which the optimal tax system partly offsets. If the assumption of a single monopsonist is relaxed (e.g., by introducing multiple firms), it seems plausible that the monopsonistic forces reduce optimal marginal tax rates compared to the competitive benchmark, but unlikely that they become negative for most workers, as this would imply substantial upward redistribution toward high-income individuals. As a final remark, the finding that the optimal tax schedule looks very different compared to the case where labor markets are competitive should *not* be taken as evidence that curvature effects are important empirically. Rather, the main lesson from these simulations is that curvature effects could have important implications for optimal income taxation if firms have significant wage-setting power. As such, the results from this section should be seen as a motivation for future empirical work on the relevance of curvature effects.

6 Conclusion

If a monopsonist sets hourly wages and individuals choose how many hours to work, labor market outcomes do not only depend on the level and the slope of the tax function, but also on its curvature. I use that insight to obtain the following three results. First, a local increase in the curvature (i.e., the second derivative) of the tax function reduces the hourly wage, hours worked and labor earnings. Intuitively, a more convex or less concave tax schedule lowers the elasticity of labor supply, which induces the monopsonist to pay lower wages. Second, the optimal tax schedule is concave at the bottom of the income distribution. Declining marginal tax rates at the bottom make low-skilled workers more responsive to wage changes, which leads the monopsonist to pay higher wages. Third, I derive an optimal tax formula that accounts for the impact of tax curvature on labor market outcomes. Compared to existing results in the literature, the additional sufficient statistics that characterize the optimal tax system are the impact of tax curvature on earnings and hourly wages, and the impact of the tax burden and the marginal tax rate on hourly wages and profits. In a structural version of the model that is calibrated to the US economy, the government uses curvature effects to boost wages of low-skilled workers by letting marginal tax rates decline rapidly at low incomes.

How important the effects of tax curvature are on labor market outcomes is an open question, but one that can be investigated empirically. One approach would be to construct a measure of tax curvature $T''(z)$ throughout the income distribution, which can then be used as an explanatory variable in a regression framework with wages, hours worked or labor earnings as the dependent variable. Alternatively, one can study the effect of tax kinks on wages (see Appendix F). Yet another approach would

[23] Hence, optimal marginal tax rates do not follow the U-shape pattern from Diamond (1998) and Saez (2001). The reason is that $z(n_0) > 0$ and there is no mass point at the bottom. Consequently, the optimal marginal tax rate starts at zero cf. Seade (1977), and not at a high level as in Saez (2001).

be to regress wages, hours worked or labor earnings on (instrumented) measures of an individual's own tax burden, marginal tax rate *and* the marginal tax rate an individual would face if the individual's earnings increase or decrease. The model from this paper predicts that a higher (lower) marginal tax rate of one's 'neighbor' in the income distribution with slightly higher (lower) earnings negatively affects wages, hours worked and labor earnings. I leave an empirical investigation of this issue as a topic for future research.

References

Bargain, O. and A. Peichl (2016). Own-wage labor supply elasticities: variation across time and estimation methods. *IZA Journal of Labor Economics 5*, 1–31.

Berger, D., K. Herkenhoff, and S. Mongey (2022). Labor market power. *American Economic Review 112*(4), 1147–93.

Boar, C. and V. Midrigan (2020). Markups and inequality. NBER Working Paper No. 25952, Cambridge-MA: NBER.

Cahuc, P. and G. Laroque (2014). Optimal taxation and monopsonistic labor market: does monopsony justify the minimum wage? *Journal of Public Economic Theory 16*(2), 259–273.

Chetty, R. (2009). Sufficient statistics for welfare analysis: a bridge between structural and reduced-form methods. *Annual Review of Economics 1*, 451–488.

Chetty, R. (2012). Bounds on elasticities with optimization frictions: a synthesis of micro and macro evidence on labor supply. *Econometrica 80*(3), 969–1018.

da Costa, C. and L. Maestri (2019). Optimal Mirrleesian taxation in non-competitive labor markets. *Economic Theory 68*, 845–886.

Diamond, P. (1998). Optimal income taxation: an example with a U-shaped pattern of optimal marginal tax rates. *American Economic Review 88*(1), 83–95.

Eeckhout, J., C. Fu, W. Li, and X. Weng (2021). Optimal taxation and market power. mimeo, UPF Barcelona.

Evers, M., R. De Mooij, and D. Van Vuuren (2008). The wage elasticity of labour supply: a synthesis of empirical estimates. *De Economist 156*, 25–43.

Gerritsen, A. (2023). Optimal nonlinear taxation: a simpler approach. *International Tax and Public Finance*. Forthcoming.

Golosov, M., A. Tsyvisnki, and N. Werquin (2014). A variational approach to the analysis of tax systems. NBER Working Paper No. 20780, Cambridge-MA: NBER.

Gürer, E. (2021). Rising markups and optimal redistributive taxation. mimeo, Goethe University Frankfurt.

Hammond, P. (1979). Straightforward individual incentive compatibility in large economies. *Review of Economic Studies 46*(2), 263–282.

Hariton, C. and G. Piaser (2007). When redistribution leads to regressive taxation. *Journal of Public Economic Theory 9*(4), 589–606.

Hersoug, T. (1984). Union wage responses to tax changes. *Oxford Economic Papers 36*(1), 37–51.

Hummel, A. J. (2021). Monopsony power, income taxation and welfare. CESifo Working Paper No. 9128, Munich: CESifo.

Jacquet, L. and E. Lehmann (2021). Optimal income taxation with composition effects. *Journal of the European Economic Association 19*(2), 1299–1341.

Kaplow, L. (2019). Market power and income taxation. NBER Working Paper No. 25578, Cambridge-MA: NBER.

Kleven, H. (2016). Bunching. *Annual Review of Economics 8*, 435–464.

Kushnir, A. and R. Zubrickas (2020). Optimal income taxation with endogenous prices. mimeo, Carnegie Mellon University.

Lee, D. and E. Saez (2012). Optimal minimum wage policy in competitive labor markets. *Journal of Public Economics 96*(9-10), 739–749.

Manning, A. (2003). *Monopsony in motion: imperfect competition in labor markets.* Princeton: Princeton University Press.

Mirrlees, J. (1971). An exploration in the theory of optimum income taxation. *Review of Economic Studies 38*(2), 175–208.

Moore, D. (2022). Optimal taxation with political externalities. mimeo, University of Hawaii.

Patterson, M. and A. Rao (2014). GPOPS-II: a matlab software for solving multiple-phase optimal control problems using hp-adaptive Gaussian quadrature collocation methods and sparse nonlinear programming. *ACM Transactions on Mathematical Software 41*(1), 1–37.

Pisauro, G. (1991). The effect of taxes on labour in efficiency wage models. *Journal of Public Economics 46*(3), 329–345.

Pissarides, C. (1985). Taxes, subsidies and equilibrium unemployment. *Review of Economic Studies 52*(1), 121–133.

Robinson, J. (1933). *The economics of imperfect competition.* London: Palgrave Macmillan UK.

Rothschild, C. and F. Scheuer (2013). Redistributive taxation in the Roy model. *Quarterly Journal of Economics 128*(2), 623–668.

Sachs, D., A. Tsyvinski, and N. Werquin (2020). Nonlinear tax incidence and optimal taxation in general equilibrium. *Econometrica 88*(2), 469–493.

Sadka, E. (1976). On income distribution, incentive effects and optimal income taxation. *Review of Economic Studies 43*(2), 261–267.

Saez, E. (2001). Using elasticities to derive optimal income tax rates. *Review of Economic Studies 68*(1), 205–229.

Saez, E. and S. Stantcheva (2018). A simpler theory of optimal capital taxation. *Journal of Public Economics 162*, 120–142.

Seade, J. (1977). On the shape of optimal tax schedules. *Journal of Public Economics 7*(2), 203–235.

Slemrod, J. and W. Kopczuk (2002). The optimal elasticity of taxable income. *Journal of Public Economics 84*(1), 91–112.

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

Trabandt, M. and H. Uhlig (2011). The Laffer curve revisited. *Journal of Monetary Economics 58*(4), 305–327.

A Proof Proposition 1

Use the first-order condition (2) to write $l(n) = l(w(n))$, or simply $l = l(w)$. Substituting this in the firm's objective $\pi(n) = (n - w)l(w)$ gives an unconstrained optimization problem in the wage w. The first-order condition is

$$0 = -l(w) + (n-w)l'(w) = -l(w) + (n-w)\times \quad (19)$$

$$\frac{u''(R(wl(w)))wl(w)R'(wl(w))^2 + u'(R(wl(w)))(R'(wl(w)) + wl(w)R''(wl(w)))}{-u''(R(wl(w)))w^2R'(wl(w))^2 + \phi''(l(w)) - u'(R(wl(w)))w^2R''(wl(w))},$$

where the term $l'(w)$ is obtained from implicitly differentiating the first-order condition (2) and $R(x) = x - T(x)$ denotes the retention function, with $R'(x) = 1 - T'(x)$ and $R''(x) = -T''(x)$. Equation (19) implicitly pins down the equilibrium wage $w(n)$ for an individual with ability n and can be used to study how the latter is affected by a local increase in the curvature of the tax function. To that end, introduce a reform parameter γ and define the function

$$\Delta(w,\gamma) = -l(w) + (n-w)\times \quad (20)$$

$$\frac{u''(R(wl(w)))wl(w)R'(wl(w))^2 + u'(R(wl(w)))(R'(wl(w)) + wl(w)(R''(wl(w)) - \gamma))}{-u''(R(wl(w)))w^2R'(wl(w))^2 + \phi''(l(w)) - u'(R(wl(w)))w^2(R''(wl(w)) - \gamma)}.$$

Equation (19) implies $\Delta(w, 0) = 0$. Because $R''(x) = -T''(x)$, an increase in γ can be interpreted as a local increase in the curvature of the tax function that leaves the tax burden $T(x)$ and marginal tax rate $T'(x)$ unaffected. The impact on the equilibrium wage follows from the implicit function theorem:

$$\left.\frac{dw}{d\gamma}\right|_{\gamma=0} = -\frac{\Delta_\gamma(w,0)}{\Delta_w(w,0)}. \quad (21)$$

The second-order condition of the profit maximization problem implies $\Delta_w(w,0) < 0$. Consequently, the sign of $dw/d\gamma$ is the same as the sign of $\Delta_\gamma(w,0)$. Ignoring function arguments to save on notation, the latter is

$$\Delta_\gamma(w,0) = \frac{(w-n)u'wl}{-u''w^2R'^2 + \phi'' - u'w^2R''} + \frac{(w-n)(u''wlR'^2 + u'R' + u'wlR'')u'w^2}{(-u''w^2R'^2 + \phi'' - u'w^2R'')^2}. \tag{22}$$

The second-order condition of the utility maximization problem (1) implies that the denominator of the first term is positive. Moreover, at an interior solution with positive profits, the labor supply curve is upward sloping and the wage is below productivity: see equation (4). As a result, the first term of equation (22) is negative. The second term is negative as well. To see why, note that equation (19) implies $u''wlR'^2 + u'(R' + wlR'') > 0$. It follows that $\Delta_\gamma(w,0) < 0$ and hence, $dw/d\gamma < 0$. Because the labor supply curve is upward sloping at an interior solution, a local increase in the curvature of the tax function also leads to a reduction in hours worked l and labor earnings $z = w \times l$.

B Proof Lemma 1

As explained in the proof of Proposition 1, the equilibrium wage $w(n)$ is pinned down by equation (19). Define the right-hand side of this equation as $\Omega(w,n)$. By the implicit function theorem,

$$w'(n) = -\frac{\Omega_n(w,n)}{\Omega_w(w,n)} = -\frac{1}{\Omega_w(w,n)}\frac{l(w)}{n-w}. \tag{23}$$

The second-order condition implies $\Omega_w(w,n) < 0$. Consequently, at an interior solution with positive profits, $w'(n) > 0$. Because the labor supply curve $l(w(n))$ is upward sloping at an interior solution (i.e., $l'(w) > 0$), it follows immediately that labor effort $l(n)$ and labor earnings $z(n) = w(n) \times l(n)$ are increasing in ability as well: $l'(n), z'(n) > 0$.

C Proof Proposition 2

The key behind the result from Proposition 2 lies in demonstrating that the constraint $\pi(n_0) \geq 0$ is binding. To do so, write the optimal tax problem in terms of allocation variables:

$$\max_{[v(n),\pi(n),l(n),b(n)]_{n_0}^{n_1}} \mathcal{W} = \int_{n_0}^{n_1} \Psi(v(n),n)f(n)dn + \omega(1-\tau)\int_{n_0}^{n_1} \pi(n)f(n)dn, \tag{24}$$

$$\begin{aligned}
\text{s.t.} \quad & \int_{n_0}^{n_1} \Big[nl(n) - u^{-1}(v(n) + \phi(l(n))) - (1-\tau)\pi(n)\Big] f(n)dn = G, \\
\forall n: \quad & v'(n) = \phi'(l(n))\left(\frac{\pi(n)}{nl(n) - \pi(n)}\right) b(n), \\
\forall n: \quad & \pi'(n) = l(n), \\
\forall n: \quad & l'(n) = b(n), \\
\forall n: \quad & b(n) \geq 0, \\
& \pi(n_0) \geq 0.
\end{aligned}$$

Using integration by parts on the incentive constraints, the Lagrangian is given by

$$\mathcal{L} = \int_{n_0}^{n_1} \Bigg[\bigg(\Psi(v(n), n) + \omega(1-\tau)\pi(n) + \eta\Big(nl(n) - u^{-1}(v(n) + \phi(l(n))) - (1-\tau)\pi(n) - G \Big) \bigg)$$
$$\times f(n) + \lambda(n)\phi'(l(n)) \left(\frac{\pi(n)}{nl(n) - \pi(n)} \right) b(n) + \lambda'(n)v(n) + \mu(n)l(n) + \mu'(n)\pi(n)$$
$$+ \chi(n)b(n) + \chi'(n)l(n) + \psi(n)b(n) \Bigg] dn + \lambda(n_0)v(n_0) - \lambda(n_1)v(n_1) + \mu(n_0)\pi(n_0)$$
$$- \mu(n_1)\pi(n_1) + \chi(n_0)l(n_0) - \chi(n_1)l(n_1) + \xi\pi(n_0). \tag{25}$$

The first-order necessary conditions with respect to the states $v(n)$, $\pi(n)$, $l(n)$ and the control $b(n)$ are:

$$v(n): \quad \left(\Psi_v(v(n), n) - \frac{\eta}{u'(c(n))} \right) f(n) + \lambda'(n) = 0, \tag{26}$$

$$\pi(n): \quad (\omega - \eta)(1-\tau)f(n) + \lambda(n)\phi'(l(n)) \frac{nl(n)}{(nl(n) - \pi(n))^2} b(n) + \mu'(n) = 0, \tag{27}$$

$$l(n): \quad \eta \left(n - \frac{\phi'(l(n))}{u'(c(n))} \right) f(n) + \mu(n) + \chi'(n)$$
$$+ \lambda(n) \left[\phi''(l(n)) \frac{\pi(n)}{nl(n) - \pi(n)} - \phi'(l(n)) n \frac{\pi(n)}{(nl(n) - \pi(n))^2} \right] b(n) = 0, \tag{28}$$

$$b(n): \quad \lambda(n)\phi'(l(n)) \frac{\pi(n)}{nl(n) - \pi(n)} + \chi(n) + \psi(n) = 0. \tag{29}$$

Here, $c(n) = u^{-1}(v(n) + \phi(l(n)))$ denotes the consumption of an individual with ability n. The latter is increasing in ability, because monotonicity of labor earnings, i.e., $b(n) \geq 0$, implies labor effort $l(n)$ and utility $v(n)$ are increasing in ability as well. As with the utility and profit maximization problems (1) and (3), I assume the second-order conditions are satisfied.[24]

The transversality conditions imply $\lambda(n_0) = \lambda(n_1) = \mu(n_1) = 0$ and $\mu(n_0) + \xi = 0$. Because $\Psi(\cdot)$ and $u(\cdot)$ are concave in $v(n)$ and $c(n)$, and because $v(n)$ and $c(n)$ are increasing in ability, equation (26) implies that $\lambda(n) \leq 0$ for all n with a strict equality only at the end-points. Moreover, the assumption that transferring income from the government budget to firm-owners does not raise welfare means $\omega \leq \eta$. Equation (27) then implies $\mu(n)$ is increasing in ability. From $\mu(n_1) = 0$, it follows that $\mu(n_0) < 0$ and hence, $\xi > 0$. The constraint $\pi(n_0) \geq 0$ thus holds with equality.

If $\pi(n_0) = 0$ and $l(n_0) > 0$, it follows that $w(n_0) = n_0$ at the optimal tax system. Equation (4) then implies $e_{lw}(n_0) \to \infty$. This requires that the denominator in equation (5) approaches zero. Therefore, at the lowest skill level $-u''w^2(1-T')^2 + \phi'' + u'w^2T'' \to 0$. Because $u'' \leq 0$ and $\phi'' > 0$, it follows that the tax function is concave at the bottom of the income distribution: $T''(z(n_0)) < 0$.

Another result that is worth highlighting is that, at an interior solution where the monotonicity constraint $b(n) \geq 0$ is not binding, the optimal marginal tax rate is positive at the bottom of the income distribution. To see this, evaluate equation (28) at n_0 and substitute the transversality condition $\lambda(n_0) = 0$. To derive an expression for $\chi'(n_0)$, differentiate equation (29) with respect to ability. Because $\lambda(n_0) = \pi(n_0) = 0$ and $\psi(n) = 0$ for all n if the monotonicity constraint $b(n) \geq 0$ is not binding, it follows that $\chi'(n_0) = 0$ as well. Using equation (2) and the property that $w(n_0) = n_0$,

[24] As explained before, these conditions can be checked *ex post* after having solved numerically the optimal tax problem.

equation (28) can be written as

$$w(n_0)T'(z(n_0))f(n_0) = -\frac{\mu(n_0)}{\eta}. \tag{30}$$

Because $\mu(n_0) < 0$ and $\eta > 0$, the optimal marginal tax rate is positive at the bottom of the income distribution: $T'(z(n_0)) > 0$.

D Proof Proposition 3

To derive the result from Proposition 3, perturb the tax function $T(z)$ in the direction $\rho(z)$ by a magnitude m, where $\rho(z)$ is a twice continuously differentiable reform function. After the reform, the tax function is $T^*(z, m) = T(z) + m\rho(z)$. With this tax function, equilibrium earnings and labor effort are pinned down by

$$u'(z - T(z) - m\rho(z))z(1 - T'(z) - m\rho'(z)) = \phi'(l)l, \tag{31}$$

$$nu''(z - T(z) - m\rho(z))z(1 - T'(z) - m\rho'(z))^2 + nu'(z - T(z) - m\rho(z))(1 - T'(z) - m\rho'(z))$$
$$= nu'(z - T(z) - m\rho(z))z(T''(z) + m\rho''(z)) + \phi'(l) + \phi''(l)l. \tag{32}$$

The first of these is obtained from multiplying equation (2) by l and the second from combining equations (4)–(5), using the property $z = w \times l$. Denote the solution to these equations by $z^R(n, m)$ and $l^R(n, m)$, respectively. Equilibrium wages are then given by $w^R(n, m) = z^R(n, m)/l^R(n, m)$ and equilibrium profits are $\pi^R(n, m) = nl^R(n, m) - z^R(n, m)$. By construction, for each outcome $y \in \{z, l, w, \pi\}$ it follows that $y^R(n, 0) = y(n)$.

The impact of the reform parameter m on $y(n)$ can be obtained by applying the implicit function theorem to (31)–(32) and using the relationship for wages $w^R(n, m) = z^R(n, m)/l^R(n, m)$ and profits $\pi^R(n, m) = nl^R(n, m) - z^R(n, m)$. Evaluated at $m = 0$,

$$\frac{\partial y^R(n, 0)}{\partial m} = \tilde{y}_T(n)\rho(z(n)) + \tilde{y}_{T'}(n)\rho'(z(n)) + \tilde{y}_{T''}(n)\rho''(z(n)). \tag{33}$$

Here, $\tilde{y}_T(n)$, $\tilde{y}_{T'}(n)$ and $\tilde{y}_{T''}(n)$ denote the response of outcome $y \in \{z, l, w, \pi\}$ to a local increase in the level, slope and curvature of the tax function, at ability level n.[25]

After the reform is implemented, social welfare is given by

$$\mathcal{W} = \int_{n_0}^{n_1} \Psi(u(z^R(n, m) - T(z^R(n, m)) - m\rho(z^R(n, m))) - \phi(l^R(n, m)), n)f(n)dn$$
$$+ \omega(1 - \tau)\int_{n_0}^{n_1} \pi^R(n, m)f(n)dn. \tag{34}$$

The government budget constraint, in turn, is

$$\int_{n_0}^{n_1} \left[T(z^R(n, m)) + m\rho(z^R(n, m)) + \tau\pi^R(n, m)\right]f(n)dn = G. \tag{35}$$

[25] To illustrate, from equation (21), $\tilde{w}_{T''}(n) = -\Delta_\gamma(w(n), 0)/\Delta_w(w(n), 0)$.

The Lagrangian associated with the optimal tax problem is then

$$\mathscr{L}(m) = \int_{n_0}^{n_1} \Big[\Psi(u(z^R(n,m) - T(z^R(n,m)) - m\rho(z^R(n,m))) - \phi(l^R(n,m)), n) + \omega(1-\tau) \\ \times \pi^R(n,m) + \eta\Big(T(z^R(n,m)) + m\rho(z^R(n,m)) + \tau\pi^R(n,m) - G \Big)\Big] f(n)dn + \xi\pi^R(n_0, m). \quad (36)$$

Here, η is the multiplier on the government budget constraint and ξ the multiplier on the constraint that the profits from hiring the least productive workers are non-negative.

To obtain the welfare impact of perturbing the tax function $T(z)$ in the direction $\rho(z)$, differentiate the Lagrangian (36) with respect to m and evaluate the result at $m = 0$. Denoting by $v(n) = u(z(n) - T(z(n))) - \phi(l(n))$ the utility for an individual with ability n, this gives

$$\frac{\partial \mathscr{L}(0)}{\partial m} = \int_{n_0}^{n_1} \Big[-\Psi_v(v(n),n)u'(z(n) - T(z(n))) + \eta \Big] \rho(z(n)) f(n)dn \\ + \int_{n_0}^{n_1} \Psi_v(v(n),n) \times \Big[u'(z(n) - T(z(n)))(1 - T'(z(n))) \frac{\partial z^R(n,0)}{\partial m} - \phi'(l(n)) \frac{\partial l^R(n,0)}{\partial m} \Big] f(n)dn \\ + \eta \int_{n_0}^{n_1} \Big[T'(z(n)) \frac{\partial z^R(n,0)}{\partial m} + \Big(\frac{\omega}{\eta}(1-\tau) + \tau \Big) \frac{\partial \pi^R(n,0)}{\partial m} \Big] f(n)dn + \xi \frac{\partial \pi^R(n_0,0)}{\partial m}. \quad (37)$$

The first line captures the mechanical welfare effect associated with transferring income from individuals to the government budget. In addition to the mechanical effect, the reform also affects equilibrium earnings, labor effort and profits. The second line captures the welfare effect due to changes in individual utilities, whereas the first term on the third line captures the welfare effect due to the impact on the government budget and the utilities of firm-owners. Lastly, the tax reform may also make the non-negativity constraint $\pi(n_0) \geq 0$ more or less binding. The final term on the third line captures the associated impact on welfare.

To proceed, divide equation (37) by η and define the welfare weight of an individual with ability n and the welfare weight of firm-owners by:

$$\tilde{g}(n) = \frac{1}{\eta}\Psi_v(v(n),n)u'(z(n) - T(z(n))), \qquad g_f = \frac{\omega}{\eta} \leq 1. \quad (38)$$

In words, $\tilde{g}(n)$ and g_f measure the monetized increase in social welfare if an individual with ability n or firm-owners receive an additional unit of after-tax income. From the assumption that transferring income from the government budget to firm-owners does not raise welfare, it follows that $g_f \leq 1$. Next, use the first-order condition (2) to substitute out for $\phi'(l(n)) = u'(z(n) - T(z(n)))w(n)(1 - T'(z(n)))$. Equation (37) then reads:

$$\frac{\partial \mathscr{L}(0)}{\partial m}\frac{1}{\eta} = \int_{n_0}^{n_1} \Big[(1 - \tilde{g}(n))\rho(z(n)) + \Big(\tilde{g}(n)(1 - T'(z(n))) + T'(z(n)) \Big) \frac{\partial z^R(n,0)}{\partial m} \\ - \tilde{g}(n)w(n)(1 - T'(z(n))) \frac{\partial l^R(n,0)}{\partial m} + (g_f(1-\tau) + \tau) \frac{\partial \pi^R(n,0)}{\partial m} \Big] f(n)dn + \frac{\xi}{\eta} \frac{\partial \pi^R(n_0,0)}{\partial m}. \quad (39)$$

From $z^R(n,m) = w^R(n,m)l^R(n,m)$, it follows that

$$\frac{\partial z^R(n,m)}{\partial m} - w^R(n,m)\frac{\partial l^R(n,m)}{\partial m} = l^R(n,m)\frac{\partial w^R(n,m)}{\partial m}. \tag{40}$$

Evaluate at $m = 0$ and substitute into equation (39) to find

$$\frac{\partial \mathscr{L}(0)}{\partial m}\frac{1}{\eta} = \int_{n_0}^{n_1} \bigg[(1-\tilde{g}(n))\rho(z(n)) + \tilde{g}(n)(1-T'(z(n)))l(n)\frac{\partial w^R(n,0)}{\partial m} + T'(z(n))\frac{\partial z^R(n,0)}{\partial m} + (g_f(1-\tau)+\tau)\frac{\partial \pi^R(n,0)}{\partial m}\bigg] f(n)dn + \frac{\xi}{\eta}\frac{\partial \pi^R(n_0,0)}{\partial m}. \tag{41}$$

Next, use equation (33) to substitute out for the responses of earnings, wages and profits to the tax reform:

$$\begin{aligned}\frac{\partial \mathscr{L}(0)}{\partial m}\frac{1}{\eta} = \int_{n_0}^{n_1} \bigg[&(1-\tilde{g}(n))\rho(z(n)) + \tilde{g}(n)(1-T'(z(n)))l(n)\Big(\tilde{w}_T(n)\rho(z(n)) + \tilde{w}_{T'}(n)\rho'(z(n)) \\ &+ \tilde{w}_{T''}(n)\rho''(z(n))\Big) + T'(z(n))\Big(\tilde{z}_T(n)\rho(z(n)) + \tilde{z}_{T'}(n)\rho'(z(n)) + \tilde{z}_{T''}(n)\rho''(z(n))\Big) \\ &+ (g_f(1-\tau)+\tau)(\tilde{\pi}_T(n)\rho(z(n)) + \tilde{\pi}_{T'}(n)\rho'(z(n)) + \tilde{\pi}_{T''}(n)\rho''(z(n)))\bigg] f(n)dn \\ &+ \frac{\xi}{\eta}\Big(\tilde{\pi}_T(n_0)\rho(z(n_0)) + \tilde{\pi}_{T'}(n_0)\rho'(z(n_0)) + \tilde{\pi}_{T''}(n_0)\rho''(z(n_0))\Big).\end{aligned} \tag{42}$$

Because profits $\pi^R(n,m)$ satisfy

$$\pi^R(n,m) = \max_{l,z}\left\{nl - z \text{ s.t. } u'(z - T(z) - m\rho(z))z(1-T'(z) - m\rho'(z)) = \phi'(l)l\right\}, \tag{43}$$

where the tax burden and marginal tax rate show up but tax curvature does not, it follows that $\tilde{\pi}_{T''}(n) = 0$. Substitute this in equation (42) and collect the terms with $\rho(z(n))$, $\rho'(z(n))$ and $\rho''(z(n))$:

$$\begin{aligned}\frac{\partial \mathscr{L}(0)}{\partial m}\frac{1}{\eta} = \int_{n_0}^{n_1} \bigg[&\Big(1-\tilde{g}(n) + T'(z(n))\tilde{z}_T(n) + (g_f(1-\tau)+\tau)\tilde{\pi}_T(n) + \tilde{g}(n)(1-T'(z(n)))l(n) \\ &\times \tilde{w}_T(n)\Big)\rho(z(n)) + \Big(T'(z(n))\tilde{z}_{T'}(n) + (g_f(1-\tau)+\tau)\tilde{\pi}_{T'}(n) + \tilde{g}(n)(1-T'(z(n)))l(n)\tilde{w}_{T'}(n)\Big) \\ &\times \rho'(z(n)) + \Big(T'(z(n))\tilde{z}_{T''}(n) + \tilde{g}(n)(1-T'(z(n)))l(n)\tilde{w}_{T''}(n)\Big)\rho''(z(n))\bigg] f(n)dn \\ &+ \frac{\xi}{\eta}\Big(\tilde{\pi}_T(n_0)\rho(z(n_0)) + \tilde{\pi}_{T'}(n_0)\rho'(z(n_0))\Big).\end{aligned} \tag{44}$$

To proceed, let $H(z)$ denote the earnings distribution with associated density $h(z)$. Both are defined on the support $[z_0, z_1]$, where $z_0 = z(n_0)$ and $z_1 = z(n_1)$. Monotonicity of labor earnings $z'(n) \geq 0$ implies the earnings and ability distribution are related through $H(z(n)) = F(n)$ for all n and hence, $h(z(n))z'(n) = f(n)$. Making a change of variables from n to z, equation (44) can be written as

$$
\begin{aligned}
&\frac{\partial \mathscr{L}(0)}{\partial m}\frac{1}{\eta} = \frac{\xi}{\eta}\Big(\pi_T(z_0)\rho(z_0) + \pi_{T'}(z_0)\rho'(z_0)\Big) \\
&+ \int_{z_0}^{z_1}\bigg[\Big(1 - g(z) + T'(z)z_T(z) + (g_f(1-\tau)+\tau)\pi_T(z) + g(z)(1-T'(z))\hat{l}(z)w_T(z)\Big)\rho(z) \\
&+ \Big(T'(z)z_{T'}(z) + (g_f(1-\tau)+\tau)\pi_{T'}(z) + g(z)(1-T'(z))\hat{l}(z)w_{T'}(z)\Big)\rho'(z) \\
&+ \Big(T'(z)z_{T''}(z) + g(z)(1-T'(z))\hat{l}(z)w_{T''}(z)\Big)\rho''(z)\bigg]h(z)dz.
\end{aligned} \tag{45}
$$

Here, $g(z)$ denotes the welfare weight at earnings z, so that $g(z(n)) = \tilde{g}(n)$, and similarly for the responses of earnings, wages and profits to changes in the tax burden, marginal tax rate and tax curvature, respectively. Furthermore, $\hat{l}(z)$ denotes labor effort at earnings z, so that $\hat{l}(z(n)) = l(n)$.

Integrate by parts the second line of equation (45):

$$
\begin{aligned}
&\int_{z_0}^{z_1}\Big(1 - g(z) + T'(z)z_T(z) + (g_f(1-\tau)+\tau)\pi_T(z) + g(z)(1-T'(z))\hat{l}(z)w_T(z)\Big)h(z)\rho(z)dz = \\
&-\rho(z_1)\int_{z_1}^{z_1}\Big(1 - g(z) + T'(z)z_T(z) + (g_f(1-\tau)+\tau)\pi_T(z) + g(z)(1-T'(z))\hat{l}(z)w_T(z)\Big)h(z)dz \\
&+\rho(z_0)\int_{z_0}^{z_1}\Big(1 - g(z) + T'(z)z_T(z) + (g_f(1-\tau)+\tau)\pi_T(z) + g(z)(1-T'(z))\hat{l}(z)w_T(z)\Big)h(z)dz \\
&+\int_{z_0}^{z_1}\bigg[\int_{z}^{z_1}\Big(1 - g(y) + T'(y)z_T(y) + (g_f(1-\tau)+\tau)\pi_T(y) + g(y)(1-T'(y))\hat{l}(y)w_T(y)\Big) \\
&\times h(y)dy\bigg]\rho'(z)dz.
\end{aligned} \tag{46}
$$

Next, integrate by parts the fourth line of equation (45):

$$
\begin{aligned}
&\int_{z_0}^{z_1}\Big(T'(z)z_{T''}(z) + g(z)(1-T'(z))\hat{l}(z)w_{T''}(z)\Big)h(z)\rho''(z)dz = \\
&\rho'(z_1)\Big(T'(z_1)z_{T''}(z_1) + g(z_1)(1-T'(z_1))\hat{l}(z_1)w_{T''}(z_1)\Big)h(z_1) \\
&-\rho'(z_0)\Big(T'(z_0)z_{T''}(z_0) + g(z_0)(1-T'(z_0))\hat{l}(z_0)w_{T''}(z_0)\Big)h(z_0) \\
&-\int_{z_0}^{z_1}\frac{\partial}{\partial z}\Big[\big(T'(z)z_{T''}(z) + g(z)(1-T'(z))\hat{l}(z)w_{T''}(z)\big)h(z)\Big]\rho'(z)dz.
\end{aligned} \tag{47}
$$

Substitute these in equation (45), and collect the terms with $\rho'(z)$:

$$\frac{\partial \mathscr{L}(0)}{\partial m}\frac{1}{\eta} = \int_{z_0}^{z_1}\Bigg[\int_{z}^{z_1}\Big(1 - g(y) + T'(y)z_T(y) + (g_f(1-\tau)+\tau)\pi_T(y) + g(y)(1-T'(y))\hat{l}(y)$$
$$\times w_T(y)\Big)h(y)dy + \Big(T'(z)z_{T'}(z) + (g_f(1-\tau)+\tau)\pi_{T'}(z) + g(z)(1-T'(z))\hat{l}(z)w_{T'}(z)\Big)h(z)$$
$$-\frac{\partial}{\partial z}\Big(\big(T'(z)z_{T''}(z) + g(z)(1-T'(z))\hat{l}(z)w_{T''}(z)\big)h(z)\Big)\Bigg]\rho'(z)dz$$
$$-\rho(z_1)\int_{z_1}^{z_1}\Big(1 - g(z) + T'(z)z_T(z) + (g_f(1-\tau)+\tau)\pi_T(z) + g(z)(1-T'(z))\hat{l}(z)w_T(z)\Big)h(z)dz$$
$$+\rho(z_0)\int_{z_0}^{z_1}\Big(1 - g(z) + T'(z)z_T(z) + (g_f(1-\tau)+\tau)\pi_T(z) + g(z)(1-T'(z))\hat{l}(z)w_T(z)\Big)h(z)dz$$
$$+\rho'(z_1)\Big(T'(z_1)z_{T''}(z_1) + g(z_1)(1-T'(z_1))\hat{l}(z_1)w_{T''}(z_1)\Big)h(z_1)$$
$$-\rho'(z_0)\Big(T'(z_0)z_{T''}(z_0) + g(z_0)(1-T'(z_0))\hat{l}(z_0)w_{T''}(z_0)\Big)h(z_0)$$
$$+\frac{\xi}{\eta}\Big(\pi_T(z_0)\rho(z_0) + \pi_{T'}(z_0)\rho'(z_0)\Big). \tag{48}$$

If the tax function $T(\cdot)$ is optimized, perturbing it must leave welfare unaffected, irrespective of the direction $\rho(\cdot)$. Consequently, at the optimum, the term in square brackets below the first integral sign that is multiplied by $\rho'(z)$ is equal to zero:

$$0 = -\frac{\partial}{\partial z}\Bigg[\Big(T'(z)z_{T''}(z) + g(z)(1-T'(z))\hat{l}(z)w_{T''}(z)\Big)h(z)\Bigg] \tag{49}$$
$$+\Big(T'(z)z_{T'}(z) + (g_f(1-\tau)+\tau)\pi_{T'}(z) + g(z)(1-T'(z))\hat{l}(z)w_{T'}(z)\Big)h(z)$$
$$+\int_{z}^{z_1}\Big(1 - g(y) + T'(y)z_T(y) + (g_f(1-\tau)+\tau)\pi_T(y) + g(y)(1-T'(y))\hat{l}(y)w_T(y)\Big)h(y)dy.$$

Evaluating the result at z' and changing the index of integration from y to z, this result coincides with equation (13) from Proposition 3.

The result from equation (48) can also be used to derive a number of boundary conditions. Equating to zero the terms proportional to $\rho(z_0)$, $\rho'(z_0)$ and $\rho'(z_1)$ gives (the term proportional to $\rho(z_1)$ is equal to zero because the boundaries of integration coincide):

$$\int_{z_0}^{z_1}\Big(1 - g(z) + T'(z)z_T(z) + \pi_T(z) + g(z)(1-T'(z))\hat{l}(z)w_T(z)\Big)h(z)dz + \frac{\xi}{\eta}\pi_T(z_0) = 0, \tag{50}$$
$$-\Big(T'(z_0)z_{T''}(z_0) + g(z_0)(1-T'(z_0))\hat{l}(z_0)w_{T''}(z_0)\Big)h(z_0) + \frac{\xi}{\eta}\pi_{T'}(z_0) = 0, \tag{51}$$
$$\Big(T'(z_1)z_{T''}(z_1) + g(z_1)(1-T'(z_1))\hat{l}(z_1)w_{T''}(z_1)\Big)h(z_1) = 0. \tag{52}$$

Combining equations (50)–(51) leads to the following optimality condition

$$\int_{z_0}^{z_1}\Big(1 - g(z) + T'(z)z_T(z) + \pi_T(z) + g(z)(1-T'(z))\hat{l}(z)w_T(z)\Big)h(z)dz$$
$$+\frac{\pi_T(z_0)}{\pi_{T'}(z_0)}\Big(T'(z_0)z_{T''}(z_0) + g(z_0)(1-T'(z_0))\hat{l}(z_0)w_{T''}(z_0)\Big)h(z_0) = 0, \tag{53}$$

which simplifies considerably if if utility is linear in consumption, i.e., if $u(c) = c$.[26] Moreover, from equation (52), the marginal tax rate at the top satisfies

$$\frac{T'(z_1)}{1-T'(z_1)} = -g(z_1)\frac{\hat{l}(z_1)w_{T''}(z_1)}{z_{T''}(z_1)}. \tag{54}$$

Given that $w_{T''}, z_{T''} < 0$ (see Proposition 1), the optimal top rate is non-positive, $T'(z_1) \leq 0$, and zero only if the welfare weight at the top is $g(z_1) = 0$.[27] To understand this result, consider a local increase in the curvature at earnings z_1 that leaves the marginal tax rate $T'(z_1)$ and the tax burden $T(z_1)$ unaffected. This reform lowers earnings and wages for individuals with ability n_1. The earnings response has a budgetary impact proportional to $T'(z_1)$ and the wage response has a welfare impact proportional to $g(z_1)$. If $T'(z_1) > 0$, this reform unambiguously reduces welfare. However, if the tax schedule is optimized, the welfare impact of the reform must be equal to zero. It follows that the optimal marginal tax rate at the top of the income distribution is non-positive, $T'(z(n_1)) \leq 0$, and zero only if the welfare weight $g(z_1)$ of these individuals is zero.

E Optimal tax problem under perfect competition

If labor markets are perfectly competitive, equilibrium wages are $w(n) = n$ and firms make zero profits: $\pi(n) = 0$. The government's objective (i.e., the counterpart of equation (7)) is then given by

$$\mathcal{W} = \int_{n_0}^{n_1} \Psi(v(n), n) f(n) dn. \tag{55}$$

In addition, the government's budget constraint, written in terms of the allocation variables (i.e., the counterpart of equation (9)), is

$$\int_{n_0}^{n_1} \Big[nl(n) - u^{-1}(v(n) + \phi(l(n))) \Big] f(n) dn = G. \tag{56}$$

The incentive constraint (i.e., the counterpart of equation (11)) can again be found by differentiating equation (1) with respect to ability. By the envelope theorem,

$$\begin{aligned} v'(n) &= u'(w(n)l(n) - T(w(n)l(n)))l(n)(1 - T'(w(n)l(n)))w'(n) \\ &= \phi'(l(n))l(n)\frac{w'(n)}{w(n)} = \frac{\phi'(l(n))l(n)}{n}, \end{aligned} \tag{57}$$

where the second step uses equation (2), which also holds if labor markets are competitive, and the last step uses the property $w(n) = n$. Combining the above, the optimal tax problem if labor market are

[26] If $u(c) = c$, from equations (31)–(32), it follows that $z_T = l_T = 0$ and hence, $w_T = \pi_T = 0$. This is because the terms that feature $\rho(z)$ no longer show up (unlike the terms with $\rho'(z)$ and $\rho''(z)$). Equation (53) then simplifies to

$$1 = \int_{z_0}^{z_1} g(z)h(z)dz.$$

[27] This contrasts the famous result that the optimal marginal tax rate is zero at the top of the income distribution if labor markets are competitive (irrespective of the value of $g(z_1)$), provided the income distribution is bounded. See Sadka (1976).

competitive can be written as follows:

$$\max_{[v(n),l(n)]_{n_0}^{n_1}} \quad \mathcal{W} = \int_{n_0}^{n_1} \Psi(v(n), n) f(n) dn, \tag{58}$$
$$\text{s.t.} \quad \int_{n_0}^{n_1} \Big[nl(n) - u^{-1}(v(n) + \phi(l(n)))\Big] f(n) dn = G,$$
$$\forall n: \quad v'(n) = \frac{\phi'(l(n))l(n)}{n}.$$

After recalibrating the ability distribution and the value of G assuming labor markets are competitive, I numerically solve the problem (58) using the GPOPS-II software developed by Patterson and Rao (2014), as with the optimal tax problem (24).

F Piecewise linear taxation

This Appendix works out the equilibrium in an example where the tax schedule is piecewise linear, as is the case in many countries. The utility function is the same as in Example 1 and Section 5:

$$U(c, l) = c - \frac{l^{1+1/\varepsilon}}{1 + 1/\varepsilon}. \tag{59}$$

The tax function, in turn, has two brackets:

$$T(z) = \begin{cases} -g + t_1 z & z \leq z_1, \\ -g + t_1 z_1 + t_2(z - z_1) & z > z_1. \end{cases} \tag{60}$$

The marginal tax rate is t_1 up to income z_1 and t_2 for incomes above this level.

Consider first the case where the tax function has a convex kink: $t_2 > t_1$. This the typical case in many countries. If the tax kink is convex, individuals with wages in a certain range bunch at income z_1 (see, e.g., Kleven, 2016). This means that, over this range, a higher wage leads to *lower* labor effort: $l = z_1/w$. From the perspective of a monopsonist, it is then optimal to pay workers with skills in a certain range of the ability distribution (to be determined) the lowest wage that induces them to earn income z_1. That wage is

$$w' = \left[\frac{z_1}{(1 - t_1)^{\varepsilon}}\right]^{\frac{1}{1+\varepsilon}}, \tag{61}$$

which follows from combining the relationship $z = w \times l = z_1$ and the first-order condition for labor supply $l = (w(1 - t_1))^{\varepsilon}$. For individuals with a much higher or lower ability, the monopsonist optimally selects the interior solution where the wage is a constant markdown relative to productivity. Since at an interior solution, $e_{lw} = \varepsilon$, the wage schedule is

$$w(n) = \begin{cases} \frac{\varepsilon}{1+\varepsilon} n & n \in [n_0, n'] \cup [n'', n_1], \\ w' & n \in (n', n''), \end{cases} \tag{62}$$

for certain thresholds n' and n''. The lower bound n' is such that an individual whose wage is $w(n') =$

$\varepsilon n'/(1+\varepsilon) = w'$ optimally chooses to earn income z_1. This gives:

$$n' = \frac{1+\varepsilon}{\varepsilon} w' = \frac{1+\varepsilon}{\varepsilon} \left[\frac{z_1}{(1-t_1)^{\varepsilon}} \right]^{\frac{1}{1+\varepsilon}} \tag{63}$$

The upper bound n'', in turn, follows from an indifference relationship for the monopsonist. For an individual with ability n'', the profits from paying the 'low' wage w' and inducing labor effort $l = z_1/w' = (w'(1-t_1))^{\varepsilon}$ are the same as the profits from selecting the interior equilibrium, where the 'high' wage is $w(n'') = \varepsilon n''/(1+\varepsilon)$ and labor supply is $l = (w(n'')(1-t_2))^{\varepsilon}$. The upper bound n'' is the largest root of the following equation:

$$(n'' - w')(w'(1-t_1))^{\varepsilon} = \frac{1}{1+\varepsilon} \left(\frac{\varepsilon}{1+\varepsilon} \right)^{\varepsilon} (1-t_2)^{\varepsilon} (n'')^{1+\varepsilon}. \tag{64}$$

It can readily be verified that the upper bound $n'' > n'$, provided $t_2 > t_1$.[28] The mass of individuals who bunch not only in the income, but also in the hourly wage distribution is then given by $\int_{n'}^{n''} f(n)dn$.

Given the values for the thresholds n' and n'', wages follow from equation (62). Figure 4a plots the wage schedule $w(n)$. It is increasing according to $w(n) = \varepsilon n/(1+\varepsilon)$ up to the threshold n', flat over the range (n', n''), and increasing again according to $w(n) = \varepsilon n/(1+\varepsilon)$ above the threshold n''. Because of the convex tax kink, individuals with ability $n \in (n', n'')$ receive a wage that is below the level they would receive without the kink. This leads to the first prediction.

Prediction 1. *Around a convex tax kink, some individuals receive a wage below the level they would receive without the kink.*

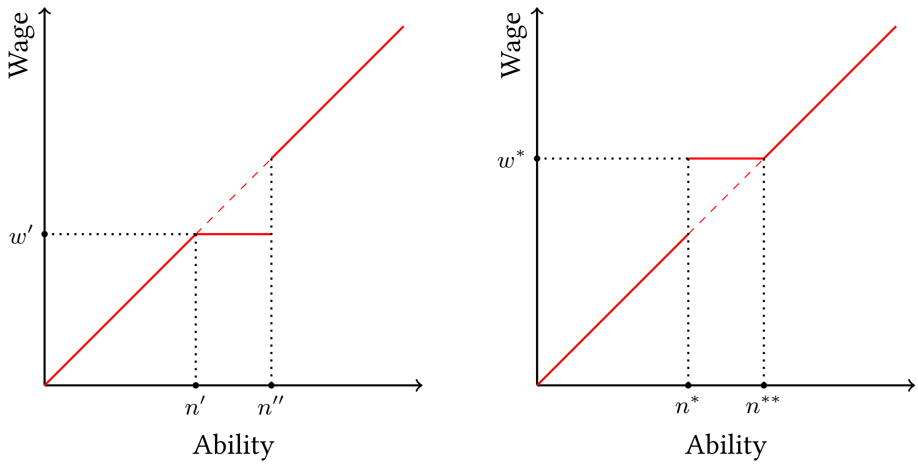
Given the wage schedule, labor supply is given by $l(n) = (w(n)(1-t_1))^{\varepsilon}$ for individuals with ability below n'', and $l(n) = (w(n)(1-t_2))^{\varepsilon}$ for individuals with ability above n''. Profits are given by $\pi(n) = (n - w(n))l(n)$, and labor income follows from $z(n) = w(n)l(n)$. The latter is increasing in ability up to n', flat over the range (n', n'') at the level z_1, then jumps to a level $z(n'') > z_1$, after which it increases again in ability. Hence, the example predicts that, around a convex tax kink, there is a mass point in the income distribution, followed by a hole. The same is true for equilibrium wages, see Figure 4a, which leads to the following.

Prediction 2. *Around a convex tax kink, there is a mass point, followed by a hole, in the distribution of both incomes and hourly wages.*

It is worth contrasting the above with the predictions from the competitive benchmark. If labor markets are competitive, individuals get paid according to their productivity. Hence, a convex tax kink does not generate bunching or a hole in the hourly *wage* distribution. A convex tax kink does, under perfect competition, lead to bunching at the *income* level where the tax rate rises. However, unlike what happens in the monopsony case, this bunch is not followed by a hole in the income distribution if labor markets are competitive (see, e.g., Kleven, 2016).

Next, consider the case where the tax schedule has a concave kink: $t_2 < t_1$. With a concave kink, there is a hole in the income distribution: some incomes are never selected (Kleven, 2016). This is

[28] Evaluated at n', the profits from paying a low wage (on the left-hand side) exceed the profits from paying a high wage (on the right-hand side), provided $t_2 > t_1$. Because the left-hand side increases linearly in n'', while the right-hand side is increasing and convex in n'', it follows that the largest root $n'' > n'$.



(a) Convex tax kink: $t_2 > t_1$

(b) Concave tax kink: $t_2 < t_1$

Figure 4: Wage schedule

because there is a wage w^* at which individuals are indifferent between working 'low' hours, $l = (w^*(1-t_1))^\varepsilon$, or 'high' hours, $l = (w^*(1-t_2))^\varepsilon$. That wage follows from equating the utility from working high and low hours:

$$w^* = \left[\frac{(1+\varepsilon)(t_1-t_2)z_1}{(1-t_2)^{1+\varepsilon}-(1-t_1)^{1+\varepsilon}}\right]^{\frac{1}{1+\varepsilon}}. \tag{65}$$

The monopsonist, however, is *not* indifferent as to whether individuals work high or low hours. Provided $n > w^*$, it prefers individuals to work long hours. The monopsonist can induce a high labor effort by paying a wage $w^* + dw$, with $dw > 0$ arbitrarily small.[29] As a result, for a certain range of the ability distribution (to be determined), it is optimal for the monopsonist to pay workers the lowest wage w^* that induces them to select the high working hours $l = (w^*(1-t_2))^\varepsilon$. For individuals with a much higher or lower ability, the monopsonist optimally selects the interior solution where the wage is a constant markdown relative to productivity. The wage schedule is therefore

$$w(n) = \begin{cases} \frac{\varepsilon}{1+\varepsilon}n & n \in [n_0, n^*] \cup [n^{**}, n_1], \\ w^* & n \in (n^*, n^{**}), \end{cases} \tag{66}$$

for certain thresholds n^* and n^{**}. The value of n^{**} is such that the wage at the interior solution is w^*. Using equation (65), this gives

$$n^{**} = \frac{1+\varepsilon}{\varepsilon}w^* = \frac{1+\varepsilon}{\varepsilon}\left[\frac{(1+\varepsilon)(t_1-t_2)z_1}{(1-t_2)^{1+\varepsilon}-(1-t_1)^{1+\varepsilon}}\right]^{\frac{1}{1+\varepsilon}}. \tag{67}$$

The value of n^* is again determined by an indifference relationship for the monopsonist. For an individual with ability n^*, the profits from paying the lowest wage w^* that induces labor effort $l = (w^*(1-t_2))^\varepsilon$ are the same as the profits from selecting the interior solution with $w(n^*) = \varepsilon n^*/(1+\varepsilon)$

[29]For simplicity, in the remainder I set $dw = 0$ and assume that if individuals are indifferent between working low hours, $l = (w^*(1-t_1))^\varepsilon$, or high hours, $l = (w^*(1-t_2))^\varepsilon$, they choose the latter.

and $l = (w(n^*)(1-t_1))^\varepsilon$. The lower bound n^* is the smallest root of

$$(n^* - w^*)(w^*(1-t_2))^\varepsilon = \frac{1}{1+\varepsilon}\left(\frac{\varepsilon}{1+\varepsilon}\right)^\varepsilon (1-t_1)^\varepsilon (n^*)^{1+\varepsilon}. \tag{68}$$

Provided $t_2 < t_1$, it can be verified that the lower bound $n^* < n^{**}$.[30] The mass of individuals who, as a result of the concave tax kink, bunch both in the income and the hourly wage distribution is therefore given by $\int_{n^*}^{n^{**}} f(n)dn$.

For given values of the thresholds n^* and n^{**}, wages follow from equation (66). The resulting wage schedule is shown in Figure 4b. For high and low ability levels, wages increase in ability according to $w(n) = \varepsilon n/(1+\varepsilon)$. Within the interval (n^*, n^{**}), individuals are paid the wage w^* that induces them to select, from the two effort levels that yield the same utility, the high one (which yields larger profits for the monopsonist). Because of the concave tax kink, these individuals receive a wage above the level they would receive in the absence of taxation. This leads to the counterpart of Prediction 1.

Prediction 3. *Around a concave tax kink, some individuals receive a wage above the level they would receive without the kink.*

Given the equilibrium wage schedule, labor effort for individuals with ability below n^* is given by $l(n) = (w(n)(1-t_1))^\varepsilon$ and for individuals with ability above n^*, labor effort is given by $l(n) = (w(n)(1-t_2))^\varepsilon$. As before, profits follow from $\pi(n) = (n - w(n))l(n)$ and labor income is given by $z(n) = w(n)l(n)$. The latter is increasing in ability up to n^*, then jumps to $z(n^{**})$ where it remains flat, and starts increasing again at ability levels above n^{**}. The pattern of labor income thus looks very similar to that of wages, see Figure 4b, which leads to the following.

Prediction 4. *Around a concave tax kink, there is a hole, followed by a mass point, in the distribution of both incomes and hourly wages.*

These predictions can again be contrasted to what would happen if labor markets are competitive. In that case, a concave tax kink does not generate a bunch or hole in the wage distribution. There is a hole in the *income* distribution around a concave tax kink, but, unlike the monopsony case, it is not followed by a mass point (see, e.g., Kleven, 2016).

To summarize, the main takeaways from this example are that, first, the logic from Proposition 1, which was derived for a smooth tax schedule, carries over to piecewise linear taxes. In particular, a convex (concave) tax kink leads to a lower (higher) wage for some individuals. Second, if the tax schedule is piecewise linear, individuals bunch not only in the income but also in the hourly wage distribution—irrespective of whether the tax kink is convex or concave. Some of these predictions, which can be tested empirically, contrast with what would happen if labor markets are competitive.

[30] Evaluated at n^{**}, the profits at the interior solution with the high tax rate (on the right-hand side) are below the profits from paying the wage w^* that induces labor effort $l = (w^*(1-t_2))^\varepsilon$ (on the left-hand side). Because the left-hand side increases linearly in n^*, while the right-hand side is increasing and convex in n^*, it follows that the smallest root of equation (68) satisfies $n^* < n^{**}$.