Monopsony power, income taxation and welfare*

Albert Jan Hummel†

April 14, 2020

Abstract

This paper studies the implications of monopsony power for optimal income taxation and welfare if monopsony power affects the distribution of income without generating efficiency losses. Firms observe workers' abilities while the government does not and monopsony power determines what share of the labor market surplus is translated into pure economic profits. Monopsony power makes labor income taxes less effective in redistributing labor income, but more effective in redistributing capital income as part of the incidence falls on firms. Monopsony power alleviates the equity-efficiency trade-off that occurs because the government does not observe ability, but at the expense of exacerbating inequality in capital income. I illustrate these findings by calibrating the model to the US economy.

JEL classification: H21, H22, J42, J48
Keywords: monopsony, optimal taxation, tax incidence

*I would like to thank Aart Gerritsen, Bas Jacobs, Marcelo Pedroni, Sander Renes, Dominik Sachs, Dirk Schindler, Kevin Spiritus and Christian Stoltenberg for many helpful comments and suggestions. This paper has benefited greatly from discussions at NED 2019 Amsterdam, the University of Amsterdam and Erasmus University Rotterdam.

†University of Amsterdam. E-mail: a.j.hummel@uva.nl.

1 Introduction

There is growing concern among economists and policymakers that firms exercise monopsony power (or buyer power) in labor markets. Recently, the Council of Economic Advisers published an issue brief on labor market monopsony (CEA (2016)) and the topic was extensively discussed during hearings held by the Federal Trade Commission (FTC (2018a,b)) and the House of Representatives.[1] The report and hearings cite a growing body of evidence documenting that (i) labor markets are highly concentrated and (ii) labor market concentration is associated with significantly lower wages (see, e.g., Azar et al. (2017, 2018, 2019), Benmelech et al. (2018), Lipsius (2018), Rinz (2018)). In addition to the potentially adverse effects on employment, output and economic efficiency, many people have voiced concerns about the *distributional* implications of monopsony power.[2]

This paper studies how monopsony power affects optimal income taxation and welfare if monopsony power changes the distribution of income without generating efficiency losses. To do so, I extend the non-linear tax framework from Mirrlees (1971) with monopsony power. Monopsony power determines what share of the labor market surplus is translated into pure economic profits. These profits are taxed at an exogenous rate and after-tax profits flow back as capital income to individuals according to their heterogeneous shareholdings. The model features inequality in labor income driven by differences in ability and inequality in capital income driven by differences in shareholdings. The government has a preference for redistribution and optimizes a non-linear tax on labor earnings.

The model generates two predictions that are of particular relevance to policymakers. First, monopsony power raises the incidence of labor income taxes that falls on firms and reduces the incidence that falls on workers. Intuitively, income taxes lower the labor market surplus and monopsony power determines what share of the surplus accrues to firms. As a result, income taxes reduce profits if firms have monopsony power. Second, monopsony power reduces inequality in labor income but increases inequality in capital income. This is because monopsony power raises aggregate profits and lowers the aggregate wage bill. As a result, any dispersion in labor (capital) income generated by differences in ability (shareholdings) is mitigated (exacerbated) if firms capture a larger share of the surplus.

Turning to the optimal tax problem, I derive an intuitive expression for the marginal tax rate on labor income at each point in the income distribution and how it is affected by

[1] The hearing on "Antitrust and Economic Opportunity: Competition in Labor Markets" was held on October 29, 2019. See `https://docs.house.gov/Committee/Calendar/ByEvent.aspx?EventId=110152`.

[2] For example, Alan Krueger noted in his address at the 2018 Fed conference in Jackson Hole:

> "... I would argue that the main effects of the increase in monopsony power and decline in worker bargaining power over the last few decades have been to shrink the slice of the pie going to workers and increase the slice going to employers, not to reduce the size of the pie overall." (Krueger (2018))

monopsony power. Income taxes are not only used to redistribute labor income, but also to redistribute capital income. The reason is that part of the tax burden is borne by firms if they have monopsony power. As a result, monopsony power makes labor income taxes less effective in redistributing labor income, but more effective in redistributing capital income. Whether monopsony power raises or lowers optimal tax rates is *a priori* ambiguous and depends on the covariance between welfare weights and shareholdings, which reflects the government's preference for redistributing capital income. Monopsony power raises optimal tax rates if the government cares strongly about redistributing capital income.

Monopsony power has an ambiguous effect on welfare. On the one hand, it increases inequality in capital income driven by differences in shareholdings. The associated impact on welfare is negative and proportional to the covariance between welfare weights and capital income. On the other hand, monopsony power decreases inequality in labor income driven by differences in ability. The associated impact on welfare is positive and proportional to the covariance between welfare weights and labor market payoffs (i.e., after-tax labor income minus the disutility of working). The reason why monopsony power might raise welfare is that firms observe ability, while the government does not. If firms have monopsony power, they reduce inequality in labor market payoffs generated by differences in ability. This reduction in inequality comes at zero efficiency costs, which can never be achieved with distortionary taxes on labor income. Monopsony power thus alleviates the equity-efficiency trade-off that occurs because the government does not observe ability, but at the expense of exacerbating inequality in capital income.

In the baseline version of the model, I assume all workers suffer to the same extent from monopsony power in the sense that with linear taxes on labor income, firms capture a constant (i.e., non ability-specific) share of the labor market surplus. I analyze an extension where this share is declining in ability, which may reflect that individuals with higher ability also have more bargaining power. Compared to the case where monopsony power does not vary with ability, optimal marginal tax rates are higher and the welfare effect of raising monopsony power is lower. Intuitively, inequality driven by differences in ability is exacerbated if individuals with higher ability suffer less from monopsony.

To illustrate the implications of monopsony power for optimal income taxation and welfare, I calibrate the baseline version of the model to the US economy. The degree of monopsony power is used to target an estimate of the pure profit share from Barkai and Benzell (2018). I find that monopsony power raises (lowers) optimal marginal tax rates at low (high) earnings levels. Moreover, taking monopsony power into account when designing tax policy leads to modest welfare gains that range between 0.07% and 1.04% of GDP in the calibrated economy depending on the covariance between welfare weights and shareholdings. By contrast, changing the degree of monopsony power to zero can have a large negative or positive impact on welfare (ranging between between –1.78% and +8.37% of GDP), again depending on the covariance between welfare weights and shareholdings. Finally, if the current tax system is optimal, an increase in monopsony power raises welfare only if the negative covariance between welfare weights and after-tax labor income is at least 2.85 times as large as the negative covariance between welfare weights and after-tax capital income.

Related literature. A few papers study optimal income taxation in an environment where firms have monopsony power. As I do, Hariton and Piaser (2007) and da Costa and Maestri (2019) analyze a model where labor supply responds on the intensive (hours, effort) margin, whereas Cahuc and Laroque (2014) focus on the extensive (participation) margin. These studies assume that firms – like the government – do *not* observe workers' abilities (Hariton and Piaser (2007) and da Costa and Maestri (2019)) or their reservation wages (Cahuc and Laroque (2014)). Monopsony power then leads to a downward distortion in employment, either in hours worked or the number of individuals employed. To partly off-set this distortion, the government finds it optimal to subsidize employment. This requires *negative* marginal (participation) tax rates if labor supply responds on the intensive (extensive) margin. By contrast, in my model firms observe ability and there is no distortion in employment. Optimal marginal tax rates only serve to redistribute income and are generally *positive*. Moreover, in my model monopsony power might raise welfare. This is not possible in Hariton and Piaser (2007), Cahuc and Laroque (2014) and da Costa and Maestri (2019), since firms do not have an informational advantage compared to the government.

This paper is also related to Kaplow (2019), who studies optimal income taxation in a model with multiple goods where firms sell their products at an exogenous, good-specific mark-up over labor costs. As in the classic model of monopoly, employment and output are inefficiently low. This calls for a downward adjustment in optimal tax rates on labor income. Without variation in mark-ups, such an adjustment would "undo the wrongs" of monopoly and market power has no impact on welfare.[3] The most important difference compared to Kaplow (2019) is that I assume firms offer employees a combination of earnings and labor effort instead of charging consumers a constant mark-up over labor costs. As a result, the outcome in the absence of taxation is efficient. Hence, tax policy is exclusively aimed at redistribution – not to restore efficiency. Moreover, tax policy cannot be used to off-set the impact of monopsony power. Therefore, monopsony power affects welfare even if there is only one good and hence, no variation in mark-ups.

The model of labor market monopsony I analyze features important similarities and differences with the classic monopsony model from Robinson (1933) and the new monopsony

[3] If mark-ups vary across goods, market power does affect welfare. Kaplow (2019) shows that optimal policy is aimed at reducing the *spread* in mark-ups.

models introduced in Manning (2003). The first similarity is that firms can exercise monopsony power because they face an upward-sloping labor supply curve. In Robinson (1933) and Manning (2003), this is because firms attract more workers if they pay higher wages. In my model, the number of workers available to each firm is fixed, but a firm can increase their labor effort by offering contracts that imply a higher wage per hour. Second, the mark-up of productivity over wages, the measure of "exploitation" due to Pigou (1920), is decreasing in the elasticity of labor supply. Third and in line with empirical evidence, the pass-through of productivity gains into wages is less than one-for-one.[4] The most important difference is that in Robinson (1933) and Manning (2003), monopsony power generates distortions. By contrast, in my model the equilibrium in the absence of taxation is efficient. The same is true in Sandmo (1994), who analyzes a setting where a monopsonist chooses a payment schedule that consists of a fixed income and a wage proportional to output. Sandmo (1994) discusses the distortionary effects and incidence of income taxes, but he does not analyze how monopsony power affects optimal tax policy or welfare, which is the main goal of this paper.

Outline. The remainder of this paper is organized as follows. Section 2 presents the model. Section 3 analyzes how monopsony power affects optimal income taxation and welfare. Section 4 explores quantitatively the policy and welfare implications of monopsony power by calibrating the model to the US economy. Section 5 concludes. An appendix contains all proofs and additional details of the analysis.

2 A Mirrleesian model with monopsony power

The basic structure of the model follows Mirrlees (1971). There is a continuum of individuals who differ in their ability. They supply labor on the intensive margin to identical firms, which produce output using a linear technology with labor as the only input. The government has a preference for redistribution but – unlike firms – does not observe individuals' abilities. Instead it can only observe and hence, tax labor earnings. The main departure from the standard model is that I allow for the possibility that firms have monopsony power. Whenever this is the case, firms earn pure economic profits. These profits are taxed at an exogenous rate and after-tax profits flow back to individuals according to their heterogeneous shareholdings. Consequently, the model features inequality in labor income generated by differences in ability and inequality in capital income generated by differences in shareholdings. Both types of inequality play an important role in the remainder of the analysis.

[4] See, e.g., Kline et al. (2019) for recent evidence on the pass-through from productivity gains into wages.

2.1 Individuals

There is a unit mass of individuals who differ in their ability $n \in [n_0, n_1]$ and shareholdings $\sigma \in [\sigma_0, \sigma_1]$ with $n_0 > 0$ and $\sigma_0 \geq 0$. Ability measures how much output an individual produces per unit of effort and shareholdings determine how aggregate profits are dissipated. Let $H(n, \sigma)$ denote the joint distribution over ability and shareholdings and $h(n, \sigma)$ the corresponding density. The latter is assumed to be positive on its entire support. Moreover, denote by $F(n)$ the marginal distribution of ability with density $f(n)$.

Individuals derive utility from consumption c and disutility from providing labor effort l. Their preferences are described by a quasi-linear utility function $u(c, l) = c - \phi(l)$, where $\phi(\cdot)$ is strictly increasing, strictly convex and satisfies $\phi(0) = \phi'(0) = 0$. The assumption of quasi-linearity is made for analytical convenience and ensures that all variables except capital income vary only with ability (and not with shareholdings).[5] I denote by $l(n) \geq 0$ the labor effort exerted by an individual with ability n. In exchange for her services, she receives labor income $z(n) \geq 0$, which is subject to a labor income tax $T(\cdot)$. Individuals also generate income from holding shares in a diversified portfolio. Each individual's capital income is therefore proportional to the economy's aggregate profits. Denote by $\pi(n) = nl(n) - z(n) \geq 0$ the profits firms generate from hiring a worker with ability n. Aggregate profits are given by

$$\bar{\pi} = \int_{n_0}^{n_1} \pi(n) f(n) dn. \tag{1}$$

Profits are taxed linearly at an exogenous rate $\tau \in [0, 1]$ and after-tax profits flow back to individuals according to how many shares they own. Normalizing aggregate shareholdings to one, the utility of an individual with ability n and shareholdings σ is

$$\mathcal{U}(n, \sigma) = v(n) + \sigma(1 - \tau)\bar{\pi}. \tag{2}$$

Here, $\sigma(1 - \tau)\bar{\pi}$ is after-tax capital income and $v(n) = z(n) - T(z(n)) - \phi(l(n))$ is the payoff from working, or labor market payoff.

2.2 Firms

Firms produce output using an identical, linear technology with labor as the only input. Each firm is matched exogenously with a number of workers. I make the important assumption that each firm observes the ability of the workers with whom it is matched. To a (potential) employee, a firm offers a bundle (l, z) which consists of an effort (or hours) requirement $l \geq 0$ and labor earnings $z \geq 0$. Firms choose the bundle to maximize profits, subject to

[5] This would also be the case with Greenwood-Hercowitz-Huffman (GHH) preferences, so that the utility function is of the form $u(c, l) = V(c - \phi(l))$, where $V(\cdot)$ is increasing. I briefly comment on this alternative specification when describing the welfare function below.

the requirement that the employee's labor market payoff exceeds some threshold, or outside option $\underline{v}(n)$. The latter is takes as given by firms and allowed to vary with ability. As will be made clear below, in equilibrium the outside option is related to firms' monopsony power. If a firm is matched to a worker with ability n, it solves

$$\max_{l\geq 0, z\geq 0} \quad \pi(n) = nl - z, \tag{3}$$
$$\text{s.t.} \quad z - T(z) - \phi(l) \geq \underline{v}(n).$$

I assume the tax function $T(\cdot)$ is such that the first-order conditions are both necessary and sufficient and denote the solution to the maximization problem (3) by $l(n)$ and $z(n)$. At an interior solution, labor effort and earnings are related through

$$n = \frac{\phi'(l(n))}{1 - T'(z(n))}. \tag{4}$$

In the optimum, firms offer bundles which equate an individual's productivity (on the left-hand side) to her willingness to substitute between labor effort and earnings (on the right-hand side). Without taxes on labor income, there is no distortion in labor supply as the marginal rate of substitution between consumption and labor effort equals the marginal rate of transformation. The reason why the equilibrium without taxation is efficient is that firms take into account how labor earnings and effort affect the utility of its workers. As a result, there are no unexploited gains from trade and workers and firms divide the full labor market surplus. How this is done depends on the degree of monopsony power.

2.3 Monopsony power

Monopsony power determines what share of the labor market surplus is translated into pure economic profits. If labor markets are competitive as in Mirrlees (1971), the full labor market surplus accrues to workers as profits are driven to zero: $\pi(n) = 0$ and labor earnings satisfy $z(n) = nl(n)$.[6] Conversely, if firms have full monopsony power, the labor market surplus is translated into profits as workers are put on their participation constraint. The outside option then equals $\underline{v}(n) = -T(0)$ and the Lagrangian associated with the firm's maximization problem (3) is

$$\mathcal{L}(n) = nl - z + \kappa_1 \Big[z - T(z) - \phi(l) + T(0) \Big] + \kappa_2 l + \kappa_3 z, \tag{5}$$

[6] This equilibrium occurs if individuals can always find a job where they work their preferred number of hours at an hourly wage equal to their productivity. The outside option $\underline{v}(n)$ is then given by

$$\underline{v}(n) = \max_{l} \quad \{nl - T(nl) - \phi(l)\}.$$

where the κ's are Lagrange multipliers. I assume the non-employment benefit $-T(0)$ is such that firms do not make profits from hiring the least productive workers: $\pi(n_0) = 0$.[7] To derive an expression for the profits from hiring *any* worker, differentiate the objective (5) with respect to ability n and apply the Envelope theorem to find $\mathcal{L}'(n) = \pi'(n) = l(n)$, where $l(n)$ is the optimal choice of labor effort offered to an individual with ability n. Integrating this relationship and imposing the boundary condition $\pi(n_0) = 0$ gives an expression for profits if firms have full monopsony power:

$$\pi(n) = \int_{n_0}^{n} l(m)dm. \tag{6}$$

For any intermediate degree of monopsony power, firms capture part of the labor market surplus. To study the welfare effects of monopsony power and to keep the optimal tax problem tractable, I choose a specific way to operationalize monopsony power. It is formally defined as follows.

Definition 1. ***Monopsony power*** *$\mu(n) \in [0,1]$ and the profits $\pi(n) = nl(n) - z(n)$ firms generate from hiring a worker with ability n are related through*

$$\pi(n) = \mu(n) \int_{n_0}^{n} l(m)dm. \tag{7}$$

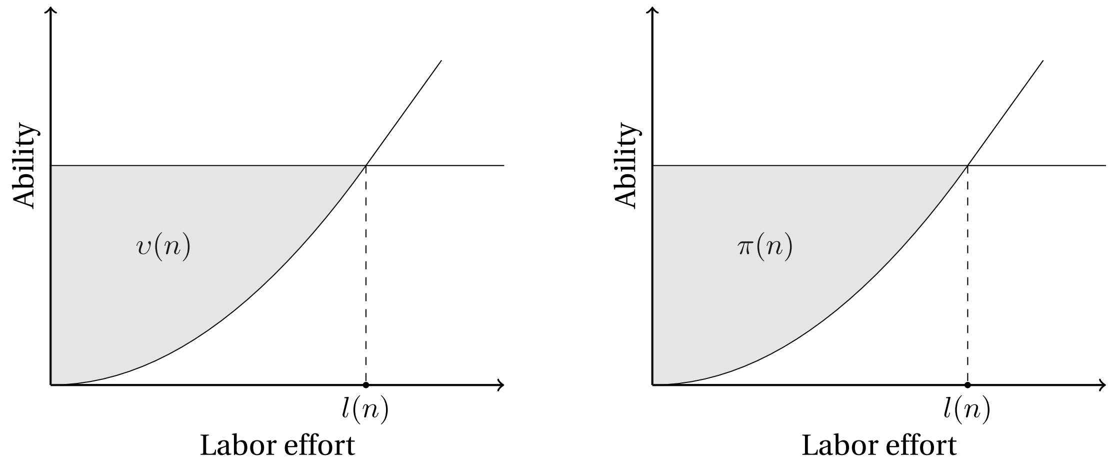
Clearly, profits are zero if labor markets are competitive (i.e., if $\mu(n) = 0$). Conversely, if firms have full monopsony power (i.e., if $\mu(n) = 1$), equations (6) and (7) coincide. In this case, the full labor market surplus is translated into profits as workers are put on their participation constraint. The degree of monopsony power $\mu(n)$ might vary with ability, which captures that individuals with different abilities might suffer more or less from monopsony.

If taxes on labor income are linear, monopsony power $\mu(n) \in [0,1]$ equals the share of the labor market surplus that is translated into pure economic profits if firms hire a worker with ability n. The payoffs for workers and firms then coincide with those obtained under the weighted Kalai-Smorodinsky bargaining solution introduced in Thomson (1994), where the payoff of each party is proportional to her ideal ('utopia') pay-off.[8] The weights $\mu(n)$ and $1 - \mu(n)$ can therefore be interpreted as the bargaining power of firms and workers, respectively. To make sure high-ability workers are not worse off, I assume individuals with higher ability do not have a lower bargaining power (i.e., do not suffer more from monopsony): $\mu'(n) \leq 0$.

[7] As is formally demonstrated in Appendix V, from an optimal tax perspective the assumption that firms do not earn profits from hiring the least productive workers is without loss of generality.

[8] Strictly speaking, the payoffs no longer necessarily coincide with those from the weighted Kalai-Smorodinsky solution if taxes on labor income are non-linear. The reason is that with non-linear taxes, labor effort generally depends on the degree of monopsony power as it is no longer pinned down only by the first-order condition (4) (as would be the case with linear income taxes). As stated above, the reason for choosing to operationalize monopsony power in this specific way is to guarantee the optimal tax problem remains tractable and to make it possible to study the welfare effects of monopsony power.

Figure 1 graphically illustrates how monopsony power affects the payoffs of workers and firms. Here, I assume income taxes are absent: $T(\cdot) = 0$. The horizontal line plots an individual's ability and corresponds to the labor demand schedule if labor markets are competitive. The upward-sloping line plots the relationship $\phi'(l) = n$, which – under perfect competition – corresponds to the labor supply schedule. The shaded area shows the labor market surplus. Monopsony power does not affect the size of the surplus (i.e., does not generate efficiency losses), but determines how it is split between workers and firms. If labor markets are competitive, firms earn zero profits and the full surplus accrues to workers. The shaded area then corresponds to the individual's labor market payoff $v(n)$: see Figure 1a. Conversely, if labor markets are fully monopsonistic, all surplus accrues to firms. The shaded area then corresponds to profits $\pi(n)$: see Figure 1b.[9]



(a) Perfect competition: $\mu(n) = 0$ (b) Full monopsony power: $\mu(n) = 1$

Figure 1: Labor market equilibrium

2.4 Government

The government's preferences are described by the following welfare function:

$$\mathcal{W} = \int_{\sigma_0}^{\sigma_1} \int_{n_0}^{n_1} \gamma(n,\sigma)\mathcal{U}(n,\sigma)h(n,\sigma)dnd\sigma. \tag{8}$$

Here, $\gamma(n,\sigma) \geq 0$ is the welfare weight (or Pareto weight) the government attaches to an individual with ability n and shareholdings σ. The average welfare weight is normalized to one. To make sure the government wishes to redistribute from individuals with high to individuals with low capital income, I assume the average welfare weight of individuals with the same shareholdings $\mathbb{E}[\gamma(n,\sigma)|\sigma]$ is weakly decreasing in σ. Similarly, to generate a motive to redistribute from individuals with high to individuals with low labor income, I assume the average

[9]The equilibrium with full monopsony power also occurs if firms engage in first-degree price discrimination. In that case, firms pay workers their reservation wage *for every hour worked* and demand labor effort up to the point where the worker's productivity is high enough to compensate for the marginal disutility of working.

welfare weight of individuals with the same ability $g(n) = \mathbb{E}[\gamma(n,\sigma)|n]$ is weakly decreasing in n.[10] Using the welfare weights $g(n)$, it is instructive to write the welfare function as follows.

Lemma 1. *The welfare function (8) can be written as*

$$\mathcal{W} = \int_{n_0}^{n_1} \Big[g(n)\upsilon(n) + (1-\Sigma)(1-\tau)\pi(n) \Big] f(n)dn, \tag{9}$$

where $\Sigma = -\text{Cov}[\boldsymbol{\sigma}, \boldsymbol{\gamma}] \in [0,1]$ *is the negative covariance between shareholdings and welfare weights, which is bounded between zero and one.*

Proof. See Appendix I. □

Individuals derive utility from earning labor income and capital income. Welfare is therefore increasing in the labor market payoff and after-tax profits. Importantly, the extent to which after-tax profits contribute to welfare depends on the covariance between shareholdings and welfare weights. This is because the government wishes to redistribute from individuals with high to individuals with low capital income. A higher concentration of firm-ownership (captured by a higher Σ) therefore *ceteris paribus* lowers welfare. It is worth pointing out that the covariance term Σ is exogenous and bounded between zero and one. It depends only on welfare weights and the distribution of ability and shareholdings. As such, it reflects properties of the joint distribution of capital and labor income and the government's desire to redistribute capital income. An increase in the government's desire to redistribute capital income raises Σ and thereby lowers the contribution of profits to welfare.

Turning to the instrument set, as in Mirrlees (1971) I assume the government does not observe individuals' abilities but only their labor earnings, which are subject to a non-linear tax $T(\cdot)$. In addition, the government observes aggregate profits, which are taxed linearly (either at the firm or the individual level) at an exogenous rate $\tau \in [0,1]$. The government's budget constraint reads

$$\int_{n_0}^{n_1} \Big[T(z(n)) + \tau\pi(n) \Big] f(n)dn = G, \tag{10}$$

where $G \geq 0$ denotes some exogenous government spending. Because the government wishes to redistribute from individuals with high to individuals with low shareholdings, levying a non-distortionary tax on pure economic profits is a very efficient way to redistribute capital income. One can therefore interpret the exogenous rate τ as the *maximum* share of

[10] An alternative way to generate a motive for redistribution (without the need to specify exogenous Pareto weights) is to assume the individual utility function is of the GHH-form $u(c,l) = V(c-\phi(l))$, where $V(\cdot)$ is strictly increasing and strictly concave: see footnote 5. Doing so is slightly more complicated and does not generate additional, substantive insights. Another advantage of using exogenous welfare weights is that in some cases it is possible to derive a closed-form solution for the optimal marginal tax rate, as will be made clear below.

pure economic profits that can be taxed. Without a restriction on profit taxation, $\tau = 1$. Conversely, if profit taxation is restricted, $\tau < 1$. Such a restriction may reflect the existence of tax havens and profit-shifting opportunities or the government's inability to distinguish between normal and above-normal returns.[11] The restriction could also reflect that levying a confiscatory tax on pure economic profits is not optimal, for example due to adverse effects on investment and firm entry.

2.5 Equilibrium

An equilibrium is formally defined as follows.

Definition 2. *An **equilibrium** consists of levels of labor effort $l(n) \geq 0$, earnings $z(n) \geq 0$ and profits $\pi(n) = nl(n) - z(n) \geq 0$ such that, for given monopsony power $\mu(n)$ and given labor income taxes $T(\cdot)$, profit taxes τ and government spending G,*

(i) for all n, labor effort and earnings are related through (4) or $l(n) = z(n) = 0$,

(ii) for all n, profits satisfy (7),

(iii) the government runs a balanced budget: (10).

Definition 2 describes the equilibrium for a given profile of monopsony power and a given set of tax instruments. Because of the specific way of modeling monopsony power, finding the equilibrium outcomes requires solving an integral equation if labor income taxes $T(\cdot)$ are non-linear.[12] As stated before, the main advantage of this modeling choice is that it keeps the optimal tax problem tractable and makes it possible to study the welfare effects of monopsony power. A disadvantage is that it is generally not possible to obtain sharp results when studying the impact of tax reforms or monopsony power on equilibrium outcomes. Keeping this caveat in mind, it is useful to highlight two implications of monopsony power. First, monopsony power increases the incidence of labor income taxes that falls on firms and decreases the incidence that falls on workers. To see this, compare the equilibria with $\mu(n) = 0$ (perfect competition) and $\mu(n) = 1$ (full monopsony power) for all n. If labor markets are competitive, firms earn zero profits – irrespective of the level of taxation. The full incidence of labor income taxes then falls on workers. Conversely, if firms have full monopsony power, all workers are put on their participation constraint. An increase in the tax burden must then be compensated one-for-one by higher labor earnings as otherwise workers prefer non-employment. In this case, the full incidence of labor income taxes falls on firms.

[11] In the model there is no productive capital. As a result, all income generated from firm-ownership is above-normal. In reality, distinguishing between normal and above-normal returns is very cumbersome.

[12] The integral equation is $\pi(n) = \mu(n) \int_{n_0}^{n} l(m)dm$, where $l(m)$ solves the first-order condition for profit maximization $m(1 - T'(ml(m) - \pi(m))) = \phi'(l(m))$ at an interior solution. See also footnote 8.

Second, monopsony power decreases inequality in labor income generated by differences in ability but increases inequality in capital income generated by differences in shareholdings. Intuitively, monopsony power determines what share of the labor market surplus is translated into labor income and what share is translated into capital income. An increase in monopsony power raises aggregate profits and lowers the aggregate wage bill. As a result, monopsony power mitigates inequality in labor income driven by differences in ability but exacerbates inequality in capital income driven by differences in shareholdings.

I am not aware of any direct evidence either in favor or against these hypotheses. A key challenge is that one needs variation in monopsony power, which should then be linked to measures of tax incidence and inequality. Webber (2015) and Rinz (2018) attempt to do the latter. They find that a lower elasticity of labor supply at the firm level and a higher labor market concentration (the two most commonly used measures of monopsony power: see Azar et al. (2019)) are associated with higher inequality in labor earnings. At first sight, these findings appear inconsistent with the hypothesis that monopsony power reduces inequality in labor income. However, Section 4 illustrates that the model presented here does not make clear-cut predictions on the impact of monopsony power on the measures of inequality used in these papers, i.e., the variance in log earnings and the P90/P10 earnings ratio. Moreover, the model can accommodate these findings if individuals with higher ability suffer less from monopsony (i.e., if $\mu'(n) < 0$). Regarding the impact of monopsony power on tax incidence, Saez et al. (2019) find that a payroll tax cut in Sweden raised profits without affecting net-of-tax wages. This result suggests firms have substantial monopsony power, but cannot be used to test if monopsony power increases the tax incidence borne by firms. By contrast, Benmelech et al. (2018) find support for the closely related hypothesis that the pass-through from productivity gains into wages is lower when labor markets are more concentrated.

3 Optimal tax policy and the welfare effects of monopsony power

This Section analyzes how monopsony power affects optimal income taxation and welfare. For analytical convenience, I start by considering the case where monopsony power does not vary with ability: $\mu'(n) = 0$. Section 3.1 derives results for optimal income taxation and Section 3.2 analyzes the welfare impact of increasing monopsony power. Section 3.3 generalizes the main findings to the case where monopsony power varies with ability.

3.1 Optimal income taxation

The government's problem consists of choosing the tax function $T(\cdot)$ that maximizes welfare. To solve this problem, I follow the approach pioneered by Mirrlees (1971) and characterize the allocation that maximizes welfare subject to resource and incentive constraints. The details can be found in Appendix II. Here, I directly state the first main result of this paper.

Proposition 1. *Consider the case where monopsony power does not vary with ability:* $\mu(n) = \mu$ *for all* n*. At the optimal allocation, the marginal tax rate at earnings level* $z(n)$ *satisfies*

$$T'(z(n)) = \left[\mu(1-\tau)\Sigma + (1-\mu)(1-T'(z(n)))\left(1+\frac{1}{\varepsilon(n)}\right)(1-\bar{g}(n))\right]\left(\frac{1-F(n)}{nf(n)}\right), \quad (11)$$

provided the local Pareto parameter of the ability distribution $a(n) = nf(n)/(1-F(n)) \geq \mu(1-\tau)\Sigma$*. Here,* $\bar{g}(n) \in [0,1]$ *is the average welfare weight for individuals with ability at least equal to* n *and* $\varepsilon(n) = \frac{\phi'(l(n))}{\phi''(l(n))l(n)} > 0$ *is the elasticity of labor supply. The marginal tax rate is generally positive and zero at the top:* $T'(z(n_1)) = 0$*. Individuals with ability levels where* $a(n) < \mu(1-\tau)\Sigma$ *do not work at the optimal allocation:* $l(n) = z(n) = 0$*.*

Proof. See Appendix V. □

Proposition 1 gives an expression for the optimal marginal tax rate at each point in the income distribution, which is generally positive and zero only at the top.[13] At the optimum, the marginal tax rate equals a weighted average between two components, where the weights depend on the degree of monopsony power. To understand this result, consider the case where firms have full monopsony power: $\mu = 1$. The optimal marginal tax rate is then given by

$$T'(z(n)) = \frac{(1-\tau)\Sigma(1-F(n))}{nf(n)}. \quad (12)$$

If labor markets are fully monopsonistic, taxes on labor earnings are used exclusively to redistribute capital income and not to redistribute labor income. This is because the full incidence of the tax burden falls on firms as all workers are put on their participation constraint. An increase in taxes on labor earnings must then be compensated one-for-one by higher earnings as otherwise workers prefer non-employment. The purpose of the *marginal* tax rate at earnings level $z(n)$ is to raise the *tax burden* of all individuals with earnings at least equal to $z(n)$.[14] The mass of individuals for whom this is the case equals $1-F(n)$, which shows up in the numerator of equation (12). Because labor earnings for these workers are increased one-for-one with an increase in the tax burden, the government indirectly taxes profits. This is valuable provided profit taxation is restricted and the negative covariance between welfare weights and shareholdings is positive: $\tau < 1$ and $\Sigma > 0$. The benefits of indirectly taxing

[13] Hence, the famous result from Seade (1977) that the optimal marginal tax rate equals zero at *both* end-points does not apply. As will be explained below, the reason is that the marginal tax rate at the bottom can be used to redistribute capital income by indirectly taxing profits.

[14] Note that individuals with different abilities do not earn the same labor income if firms have full monopsony power. This is because firms demand more labor effort from individuals with higher ability. To compensate them (i.e., to ensure the participation constraint holds), firms must pay higher labor earnings to these individuals.

profits by raising the marginal tax rate $T'(z(n))$ should be weighed against the distortions in labor effort: see equation (4). The efficiency costs are proportional to ability n and the density $f(n)$, which determines for how many individuals labor effort is distorted. Both terms show up in the denominator of equation (12).

The second component in the optimal tax formula (11) is as in the benchmark model without monopsony power. To see this, suppose labor markets are perfectly competitive: $\mu = 0$. The optimal marginal tax rate then satisfies

$$\frac{T'(z(n))}{1 - T'(z(n))} = \left(1 + \frac{1}{\varepsilon(n)}\right)(1 - \bar{g}(n))\left(\frac{1 - F(n)}{nf(n)}\right). \tag{13}$$

This is the well-known ABC-formula from Diamond (1998). Because profits are zero if labor markets are competitive, the sole purpose of income taxes is to redistribute labor income and not to redistribute capital income.

According to equation (11), the higher the degree of monopsony power, the more taxes on labor earnings are geared toward redistributing capital income and the less they are geared toward redistributing labor income. Intuitively, monopsony power increases the incidence of income taxes that falls on firms and decreases the incidence that falls on workers. Monopsony power therefore makes labor income taxes less (more) effective in redistributing labor (capital) income. Whether monopsony power raises or lowers optimal marginal tax rates is *a priori* ambiguous and depends crucially on the government's preferences for redistribution. This insight is formalized in the next Corollary.

Corollary 1. *Suppose the utility function is iso-elastic: $\phi(l) = l^{1+1/\varepsilon}/(1+1/\varepsilon)$, so that $\varepsilon(n) = \varepsilon$ for all n. At ability levels where the local Pareto parameter $a(n) \geq \mu(1-\tau)\Sigma$, the closed-form solution for the optimal marginal tax rate is*

$$T'(z(n)) = \frac{\mu(1-\tau)\Sigma + (1-\mu)(1+1/\varepsilon)(1-\bar{g}(n))}{a(n) + (1-\mu)(1+1/\varepsilon)(1-\bar{g}(n))}. \tag{14}$$

If $(1-\tau)\Sigma > 0$, an increase in monopsony power unambiguously raises the marginal tax rate $T'(z(n_0))$ at the bottom of the income distribution if $a(n_0) \geq \mu(1-\tau)\Sigma$. Moreover, at ability levels where $\bar{g}(n) < 1$ an increase in monopsony power raises $T'(z(n))$ if and only if

$$((1-\tau)\Sigma))^{-1} < ((1+1/\varepsilon)(1-\bar{g}(n)))^{-1} + a(n)^{-1}. \tag{15}$$

Proof. See Appendix VI. □

Equation (14) gives a closed-form solution for the optimal marginal tax rate. It follows directly from rearranging equation (11) and plays an important role when exploring the quantitative implications of monopsony power for tax policy in Section 4. Equation (15), in turn, gives a precise condition which can be used to determine if an increase in monopsony power raises the optimal marginal tax rate at each point in the income distribution. Because monopsony power makes income taxes more (less) effective in redistributing capital (labor) income, the impact of monopsony power on optimal tax rates is generally ambiguous. According to equation (15), the first (positive) effect dominates if profit taxation is severely restricted (i.e., if τ is low) and if the government has a strong preference for redistributing capital income (i.e., if Σ is high). Conversely, the second (negative) effect dominates if redistributing from individuals with high to individuals with low ability is very important (i.e., if $\bar{g}(n)$ is low).[15]

The impact of monopsony power on optimal tax rates varies along the income distribution depending on the behavior of $\bar{g}(n)$ and the local Pareto parameter $a(n)$. Because the average welfare weight of all individuals equals one (i.e., $\bar{g}(n_0) = 1$), condition (15) is always satisfied at the bottom of the income distribution. Hence, monopsony power unambiguously raises $T'(z(n_0))$ if $a(n_0) \geq \mu(1-\tau)\Sigma$. Intuitively, the marginal tax rate at the bottom only serves to indirectly tax profits as it does not help to redistribute labor income from individuals with high to individuals with low ability. This becomes more important if monopsony power increases. At higher levels of income, redistributing labor income from individuals above to individuals below that level becomes on average more valuable: $\bar{g}(n)$ is decreasing. Monopsony power makes income taxes less effective in redistributing labor income as part of the tax incidence falls on firms. *Ceteris paribus*, monopsony power therefore has a smaller positive or a larger negative impact on optimal tax rates at higher income levels.

Finally, it is worth pointing out that the marginal tax rate according to equation (14) exceeds 100% if the local Pareto parameter $a(n) < \mu(1-\tau)\Sigma$. Clearly, this violates the first-order condition for profit maximization (4). In this case, the non-negativity constraint on labor effort $l(n) \geq 0$ in the government's optimization problem is binding: see Appendix V for details. Hence, at the optimal tax system some individuals may not work if firms have monopsony power. Empirically, this is only a relevant issue at the bottom of the ability distribution, where the local Pareto parameter $a(n)$ is low. The reason why the government may find it optimal to have some individuals not work is that stimulating participation by lowering the tax liability raises aggregate profits if $\mu > 0$, which has a negative impact on welfare if $(1-\tau)\Sigma > 0$.

3.2 Welfare impact of raising monopsony power

I now turn to analyze how an increase in monopsony power affects welfare. The following Proposition states the second main result of this paper.

[15]Monopsony power also lowers the optimal marginal tax rate if the local Pareto parameter $a(n)$ is high. The reason is quite mechanical. In the second component of equation (11), monopsony power affects optimal marginal tax rates through the term $T'(z(n))/(1 - T'(z(n)))$. The latter changes faster (and hence, implies a smaller change in the marginal tax rate), the higher is $T'(z(n))$. This is the case if the local Pareto parameter is low. Therefore, a lower Pareto parameter makes it easier for condition (15) to be satisfied.

Proposition 2. *Suppose monopsony power does not vary with ability and the tax function $T(\cdot)$ is optimized. An increase in monopsony power raises welfare if and only if*

$$\mu\Sigma^{\upsilon} > (1-\mu)\Sigma^{k}, \tag{16}$$

where $\Sigma^{\upsilon} = -\text{Cov}[\boldsymbol{\upsilon}, \boldsymbol{\gamma}] \geq 0$ is the negative covariance between labor market payoffs and welfare weights and $\Sigma^{k} = -\text{Cov}[\boldsymbol{\sigma}(1-\tau)\bar{\pi}, \boldsymbol{\gamma}] = \Sigma(1-\tau)\bar{\pi} \geq 0$ is the negative covariance between capital income and welfare weights.

Proof. See Appendix VII. □

Monopsony power raises aggregate profits and lowers the aggregate wage bill. The associated impact on welfare is ambiguous. On the one hand, monopsony power reduces inequality in labor income generated by differences in ability. The positive welfare effect is captured by the left-hand side of equation (16). On the other hand, it increases inequality in capital income generated by differences in shareholdings. The negative welfare effect is captured by the right-hand side of equation (16).

To gain further intuition why monopsony power might raise welfare, recall that firms observe ability while the government does not. If labor markets are competitive, firms do not benefit from this information as profits are driven to zero. By contrast, profits are positive if firms have monopsony power. Moreover, the profits firms generate from hiring a worker are increasing in ability. An increase in monopsony power thus reduces inequality in labor market payoffs generated by differences in ability. Importantly, unlike with distortionary taxes on labor income, the reduction in inequality comes at zero efficiency costs. An increase in monopsony power thus alleviates the equity-efficiency trade-off that occurs because the government does not observe ability, cf. Mirrlees (1971).

The negative welfare effect of monopsony power that occurs because it exacerbates inequality in capital income depends critically on the extent to which pure economic profits are taxed. Without a restriction on profit taxation (i.e., if $\tau = 1$), an increase in monopsony power unambiguously raises welfare as there is no inequality in capital income that is exacerbated by monopsony power. Welfare is therefore highest if firms have full monopsony power (i.e., if $\mu = 1$ as well). In this case, there is no inequality in labor market payoffs either, as all individuals are put on their identical participation constraint.[16] The government optimally uses the proceeds from the confiscatory tax on profits to finance a universal basic income $-T(0)$ that should not be taxed away if individuals earn labor income.[17] If profit taxation is unrestricted, monopsony power reduces inequality generated by differences in ability

[16]Despite that all individuals are put on their identical participation constraint, there is still inequality in labor income. This is because firms demand more labor effort from individuals with higher ability: see footnote 14.

[17]Put differently, optimal marginal tax rates are zero. To see this, substitute $\tau = \mu = 1$ in equation (11).

without exacerbating inequality generated by differences in shareholdings, which is welfare-enhancing. However, if profits cannot be taxed at a confiscatory rate, the welfare effect of monopsony power is generally ambiguous.

A few remarks are in order. First, equation (16) depends on capital income and labor market payoffs, which are both endogenous. I show in Appendix VII that the welfare effect of raising monopsony power can be written as a function of exogenous variables if the labor supply elasticity is constant. Second, the result from Proposition 2 is derived assuming income taxes are optimized. Hence, the result can only be used to assess the welfare effect of raising monopsony power at the *current* tax system under the additional assumption that the latter reflects the government's preferences for redistribution.[18] Third, labor market payoffs depend on the disutility of working, which is difficult to measure. It is also possible to derive a necessary condition that depends on the covariance between welfare weights and after-tax labor income.

Corollary 2. *Suppose monopsony power does not vary with ability and the tax function* $T(\cdot)$ *is optimized. If labor effort is weakly increasing in ability at the optimal allocation, i.e.,* $l'(n) \geq 0$*, an increase in monopsony power raises welfare only if*

$$\mu\Sigma^{\ell} > (1-\mu)\Sigma^{k}, \tag{17}$$

where $\Sigma^{\ell} = -\text{Cov}[\boldsymbol{z} - T(\boldsymbol{z}), \boldsymbol{\gamma}] > \Sigma^{v} \geq 0$ *is the negative covariance between welfare weights and after-tax labor income.*

Proof. See Appendix VII. □

If individuals with higher ability exert more effort, the negative covariance between welfare weights and after-tax labor income exceeds the negative covariance between welfare weights and labor market payoffs. Therefore, equation (17) gives a necessary condition which can be used to determine if an increase in monopsony power could raise welfare. The advantage compared to the necessary and sufficient condition from Proposition (2) is that condition (17) is arguably easier to assess for policymakers, as it depends on after-tax labor income and not on the disutility of working.

3.3 Ability-specific monopsony power

The results from Propositions 1 and 2 are derived assuming all individuals suffer to the same extent from monopsony power. Hence, if labor income taxes are linear, firms capture a share of the labor market surplus that does not vary with ability: $\mu(n) = \mu$ for all n. I now generalize

[18] The welfare weights that make the current tax system optimal can be calculated using the inverse optimal tax method: see Bourguignon and Spadaro (2012).

these results by allowing for the possibility that individuals with higher ability also have more bargaining power (i.e., suffer less from monopsony): $\mu'(n) \leq 0$.[19]

Proposition 3. *Suppose monosony power is weakly decreasing in ability: $\mu'(n) \leq 0$. The optimal marginal tax rate satisfies*

$$T'(z(n)) = \frac{1-F(n)}{nf(n)}\Bigg[\bar{\mu}(n)(1-\tau)\Sigma + (1-\mu(n))(1-T'(z(n)))\left(1+\frac{1}{\varepsilon(n)}\right)(1-\bar{g}(n)) \tag{18}$$

$$-\frac{\mu'(n)\pi(n)(1-T'(z(n)))(1-\bar{g}(n))}{\mu(n)\varepsilon(n)l(n)} - \frac{\int_n^{n_1}\mu'(m)(1-T'(z(m)))\left(\int_m^{n_1}(1-g(s))f(s)ds\right)dm}{1-F(n)}\Bigg],$$

provided $a(n) \geq \bar{\mu}(n)(1-\tau)\Sigma - \int_n^{n_1}\mu'(m)\frac{\phi'(l(m))}{m}\int_m^{n_1}(1-g(s))f(s)dsdm/(1-F(n))$*, where* $\bar{\mu}(n)$ *denotes the average monopsony power for individuals with ability at least equal to* n*. Individuals with ability levels where this condition is not satisfied do not work:* $l(n) = z(n) = 0$*. The optimal marginal tax rate is generally positive and zero at the top:* $T'(z(n_1)) = 0$*. To assess the welfare effect of monopsony power, consider a proportional increase in monopsony power from* $\mu(n)$ *to* $\mu(n)(1+\alpha)$*. The associated impact on welfare is determined by*

$$\frac{\partial\mathcal{W}(\alpha)}{\partial\alpha} = \int_{n_0}^{n_1}\Bigg[\frac{\phi'(l(n))}{n}\int_n^{n_1}(1-g(m))f(m)dm - (1-\tau)\Sigma\int_n^{n_1}\frac{\mu(m)}{\mu(n)}f(m)dm$$

$$+\int_n^{n_1}\frac{\mu'(m)}{\mu(n)}\frac{\phi'(l(m))}{m}\left(\int_m^{n_1}(1-g(s))f(s)ds\right)dm\Bigg]\mu(n)l(n)dn. \tag{19}$$

Proof. See Appendix V and VII. □

Compared to the case where monopsony power is constant, inequality generated by differences in ability is higher if individuals with higher ability suffer less from monopsony. This explains why *ceteris paribus* optimal marginal tax rates are higher. Compared to the result from Proposition 1, two additional effects show up in equation (18). First, a reduction in monopsony power *at a particular ability level* implies the labor market payoff increases more quickly in ability. Second, a reduction in monopsony power *at higher ability levels* lowers the profits firms generate from hiring more productive workers. Hence, individuals with higher ability manage to capture a larger share of the labor market surplus. Both effects raise the distributional benefits of income taxes and hence, raise the optimal marginal tax rate.

Equation (19) gives an expression for the welfare effect of raising monopsony power. If monopsony power does not vary with ability, the first (positive) term is proportional to Σ^v and the second (negative) term is proportional to Σ^k. Hence, one additional effect shows up in equation (19) compared to the result from Proposition 2. As stated before, individuals

[19]In line with this assumption, the findings from Webber (2015) and Rinz (2018) suggest that individuals at lower parts of the earnings distribution suffer more from firms' ability to exercise monopsony power.

with higher ability capture a larger share of the labor market surplus if they suffer less from monopsony. This lowers the positive welfare effect of raising monopsony power that occurs because monopsony power mitigates inequality driven by differences in ability. Hence, an increase in monopsony power has a smaller positive or a larger negative impact on welfare compared to the case where monopsony power does not vary with ability.

4 Numerical illustration

This Section quantitatively explores the implications of monopsony power in the baseline version of the model where monopsony power does not vary with ability. After presenting the calibration (Section 4.1) and the welfare function (Section 4.2), I analyze how monopsony power affects optimal income taxation (Section 4.3) and welfare (Section 4.4).

4.1 Calibration

4.1.1 Data

I calibrate the model on the basis of US data. The primary data source is the March release of the 2018 Current Population Survey (CPS), which provides detailed information on income and taxes for a large sample of individuals. For each individual I observe taxable income, the tax liability (computed as the sum of federal and state taxes) and income from wage and salary payments. In the remainder the latter is referred to as labor income, or labor earnings. In the analysis I include individuals between 25 and 65 years who derive strictly positive labor income and whose hourly wage is at least half the federal minimum wage of $7.25. For individuals whose labor income is top-coded I multiply the reported income with a factor 2.67, consistent with an estimate of the Pareto parameter of 1.6 for the distribution of labor income at the top obtained by Saez and Stantcheva (2018).[20]

4.1.2 Functional forms

To calibrate the model I require a specification of the utility function and the current tax schedule. The utility function is assumed to be of the iso-elastic form

$$u(c,l) = c - \frac{l^{1+1/\varepsilon}}{1+1/\varepsilon}, \tag{20}$$

[20]If labor income at the top follows a Pareto distribution with tail parameter $\tilde{a}$, the expected value of income above a certain amount z' equals $\mathbb{E}[z|z \geq z'] = \left(\frac{\tilde{a}}{\tilde{a}-1}\right) z'$.

where ε is the constant elasticity of labor supply. The latter is set at a value $\varepsilon = 0.33$, as suggested by Chetty (2012). I approximate the current tax schedule using a linear specification

$$T(z(n)) = -g + tz(n). \tag{21}$$

Values for the lump-sum transfer g and the constant marginal tax rate t are obtained by regressing the tax liability on taxable income, see, e.g., Saez (2001). This gives $g = \$4,590$ and $t = 33.1\%$ with an R^2 of approximately 0.94. Figure 4 in Appendix VIII plots the actual and fitted values for incomes up to $500,000.

4.1.3 Equilibrium

If the utility function is iso-elastic and the tax function is linear, it is straightforward to derive the equilibrium (cf. Definition 2). Labor effort follows from equation (4):

$$l(n) = (1-t)^{\varepsilon} n^{\varepsilon}. \tag{22}$$

Labor earnings, in turn, are obtained by substituting labor effort in equation (7) and using the definition $\pi(n) = nl(n) - z(n)$. This gives

$$z(n) = \left(1 - \frac{\mu}{1+\varepsilon}\right)(1-t)^{\varepsilon} n^{1+\varepsilon} + \left(\frac{\mu}{1+\varepsilon}\right) z(n_0). \tag{23}$$

An individual's labor income equals a weighted average of the output she produces (first term) and the labor income of the individuals with the lowest ability (second term).[21] The profits $\pi(n) = nl(n) - z(n)$ firms generate from hiring a worker with ability n are given by

$$\pi(n) = \left(\frac{\mu}{1+\varepsilon-\mu}\right)(z(n) - z(n_0)). \tag{24}$$

Equations (23) and (24) give a mapping from (observable) labor income to (unobservable) ability and pure economic profits, respectively.

With this closed-form characterization of the equilibrium, a few remarks are in place. First, as in the classic and new monopsony models introduced in Robinson (1933) and Manning (2003), the mark-up of productivity over wages (or output over earnings) is decreasing in the elasticity of labor supply. To see this, denote by $w(n) = z(n)/l(n)$ the hourly wage of an individual with ability n and assume $z(n_0)$ is very small, as in the data. Using equation (23),

[21]The reason why the lowest income shows up in equation (23) is that, by assumption, firms make no profits from hiring individuals with the lowest ability: $\pi(n_0) = 0$. This can only be the case for *any* degree of monposony power if individuals with ability n_0 are indifferent between working and not working. Therefore, the lowest income level is informative about the outside option of non-employment. Note that the value of non-employment generally differs from the lump-sum transfer g, for example because non-employed individuals are entitled to an additional benefit or because of (non-modeled) utility costs or benefits of having a job.

the mark-up, i.e., the measure of "exploitation" introduced by Pigou (1920), is

$$\frac{n - w(n)}{w(n)} = \frac{\mu}{1 + \varepsilon - \mu}. \tag{25}$$

Clearly, the latter is increasing in monopsony power μ and decreasing in the elasticity of labor supply ε. Second, equation (23) implies that if firms have monposony power, productivity gains (captured by an increase in ability n) are not translated one-for-one into higher wages. This is a standard prediction from models where firms have monopsony power that is supported by empirical evidence (see Kline et al. (2019) for a recent example). Third, from equation (23) it is clear that monopsony power mitigates inequality in labor earnings driven by differences in ability. Despite this, monposony power has no impact on typical measures of inequality in labor earnings, such as the Gini coefficient, the variance in log earnings or the P90/P10 earnings ratio. The reason is that monopsony power simply scales down labor earnings for this particular choice of the utility and tax function. In the more general case where monopsony power, the marginal tax rate or the elasticity of labor supply vary with ability, the model does not make a clear-cut prediction on the impact of monopsony power on these measures of inequality.[22]

4.1.4 Monopsony power

Monopsony power μ determines how much pure economic, or above-normal profits firms make. In recent work, Barkai and Benzell (2018) and Barkai (2019) decompose US output into a labor share, a capital share and a profit share. The labor share is calculated as total compensation to employees as a fraction of gross value added. The capital share, in turn, is calculated as the product of the capital stock and the required (or normal) rate of return, again as a fraction of gross value added. The remainder, i.e., the profit share, is a measure of pure economic profits. Because my model abstracts from productive capital, I calibrate monopsony power μ to target the ratio of aggregate profits to aggregate labor income, or the ratio of the *profit share* to the *labor share*. For the most recent year 2015, Barkai and Benzell (2018) calculate that the ratio of aggregate profits to aggregate wages is approximately 24.2%. Using their estimate, the value for monopsony power μ can be calculated by integrating equation (24) over the ability distribution and dividing by aggregate labor income $\bar{z} = \int_{n_0}^{n_1} z(n) f(n) dn$. This gives

$$\left(\frac{\bar{\pi}}{\bar{z}}\right) = \left(\frac{\mu}{1 + \varepsilon - \mu}\right)\left(1 - \left(\frac{z(n_0)}{\bar{z}}\right)\right) \quad \Leftrightarrow \quad \mu = (1 + \varepsilon)\left[\frac{(\bar{\pi}/\bar{z})}{1 + (\bar{\pi}/\bar{z}) - (z(n_0)/\bar{z})}\right]. \tag{26}$$

[22]This could also explain why Webber (2015) and Rinz (2018) find a positive association between measures of monopsony power and the variance in log earnings or the P90/P10 earnings ratio, respectively.

Substituting out for the elasticity of labor supply and the ratio of profits to wages gives a value for monopsony power of approximately $\mu = 0.26$.[23]

4.1.5 Ability distribution

As in Saez (2001), I calibrate the ability distribution to match the empirical income distribution. To do so, I use equation (23) and calculate the ability n for each individual with positive labor earnings. This gives an empirical counterpart of the ability distribution $F(n)$. I subsequently smooth this distribution by estimating a kernel density. The empirical distribution and the kernel density are plotted in the top panel of Figure 5 in Appendix VIII. The bottom panel plots the distribution of labor earnings and the implied kernel density.

I make one adjustment to the density as plotted in the top panel of Figure 5. In particular, I append a right Pareto tail starting at an ability level associated with $350,000 in annual earnings. The reason for doing so is that individuals with very high labor earnings are significantly under-represented in the CPS data. I choose the tail parameter of the *ability* distribution to be consistent with a tail parameter of 1.6 of the *labor income* distribution at the top.[24] This is the estimate obtained by Saez and Stantcheva (2018) using tax returns data. The scale parameter of the Pareto distribution is set to ensure there is no jump in the density at the point where the Pareto tail is pasted.

4.1.6 Profit taxation and revenue requirement

In the model, there is no productive capital and τ is the rate at which pure economic, or above-normal profits are taxed. The current tax system does not distinguish between normal and above-normal returns. I therefore assume all capital income is taxed at a rate $\tau = 36\%$, taken from Trabandt and Uhlig (2011). This figure is very similar to the one that is obtained if the government levies a corporate tax rate of 21% at the firm level and a capital gains tax rate of 20% at the individual level. For a given value of τ, the government's budget constraint (10) can be used to calculate the revenue requirement. This gives $G = \$22,049$, which in the calibrated economy corresponds to approximately 28.6% of aggregate output. Table 1 summarizes the calibration strategy.

[23] In the CPS data, the lowest earnings level is very small compared to average earnings. Hence, the choice of $z(n_0)/\bar{z}$ only has a small effect on the calibrated value of μ.

[24] Let $\tilde{F}(z(n))$ denote the labor income distribution with density $\tilde{f}(z(n))$. Monotonicity of labor earnings implies $F(n) = \tilde{F}(z(n))$ for all n where $z(n) > 0$ and hence, $f(n) = \tilde{f}(z(n))z'(n)$. The local Pareto parameter of the ability distribution $a(n) = nf(n)/(1 - F(n))$ and income distribution $\tilde{a}(z(n)) = z(n)\tilde{f}(z(n))/(1 - \tilde{F}(z(n)))$ are related through $a(n) = \tilde{a}(z(n))e_{zn}$, where $e_{zn} = z'(n)n/z(n)$ is the elasticity of labor earnings with respect to ability. The latter equals approximately $1 + \varepsilon$ at high levels of labor earnings: see equation (23).

Variable	Target	Source	Value
μ	Aggregate profits over wages	Barkai and Benzell (2018)	0.26
ε	Elasticity of labor supply	Chetty (2012)	0.33
τ	Tax rate on capital income	Trabandt and Uhlig (2011)	0.36
G	Government budget constraint	Equilibrium condition	$22,049
$T(z)$	Tax liability	CPS 2018	Figure 4
$F(n)$	Income distribution	CPS 2018	Figure 5

Table 1: Calibration

4.2 Welfare function

The welfare function (9) depends on the average welfare weights $g(n)$ of individuals with the same ability and the negative covariance $\Sigma \in [0, 1]$ between welfare weights and shareholdings. The first (second) determines how much the government values reducing inequality generated by differences in ability (shareholdings). In the remainder, I let Σ vary between zero and one. If $\Sigma = 0$, the government does not value redistributing capital income. Conversely, if $\Sigma = 1$, the government cares a lot about redistributing capital income as all shares are held by individuals with a welfare weight of zero. Regarding the average welfare weights of individuals with the same ability, I use the following specification:

$$g(n) = \rho n^{-\beta}. \tag{27}$$

Here, $\rho > 0$ is a scaling parameter and $\beta \geq 0$ governs how much the government wishes to redistribute from individuals with high to individuals with low ability. If $\beta = 0$, the government attaches the same average weight to individuals of all ability levels. Conversely, if $\beta \to \infty$, the government only cares about individuals with the lowest ability.

Before selecting a value for ρ and β, I make one adjustment to the welfare function (9). In particular, I assume there is a mass of $\nu = 0.05$ non-participants, who earn zero labor and capital income and whose welfare weight equals twice the average welfare weight of all other individuals. The government optimizes a benefit for the non-participants, subject to the requirement that their utility does not exceed the labor market payoff of individuals with the lowest ability. Under these assumptions, the optimal marginal tax rate at the bottom of the income distribution is positive even if labor markets are competitive or there is no (desire to reduce) capital income inequality: $(1 - \tau)\Sigma = 0$. This avoids technical difficulties associated with steeply increasing marginal tax rates at very low earnings.[25] The parameter ρ is set

[25] These difficulties arise because a low value of the local Pareto parameter $a(n)$ at the bottom implies the optimal marginal tax rate jumps from $T'(z(n_0)) = 0$ immediately to a high value. Such a jump often leads to a violation of the monotonicity condition: see Appendix III.

to make sure the average welfare weight of all individuals (including the non-participants) equals one. Moreover, I choose the value of β such that the average marginal tax rate at the optimal tax system with competitive labor markets equals the current rate $t = 33.1\%$.

4.3 Optimal marginal tax rates

Figure 2 plots optimal marginal tax rates for different assumptions on the degree of monopsony power μ and the negative covariance between welfare weights and shareholdings Σ. To facilitate the comparison, the horizontal axis shows *current* labor earnings. The red, solid line plots the marginal tax rates a "naive" government would set that acts *as if* labor markets are competitive. The tax rates are calculated by substituting $\mu = 0$ in equation (14). Consistent with the calibrated value of β, the average marginal tax rate equals 33.1%. The conventional U-shape pattern (see, e.g., Diamond (1998) and Saez (2001)) follows from the behavior of the local Pareto parameter $a(n)$: see Figure 6 in Appendix VIII.

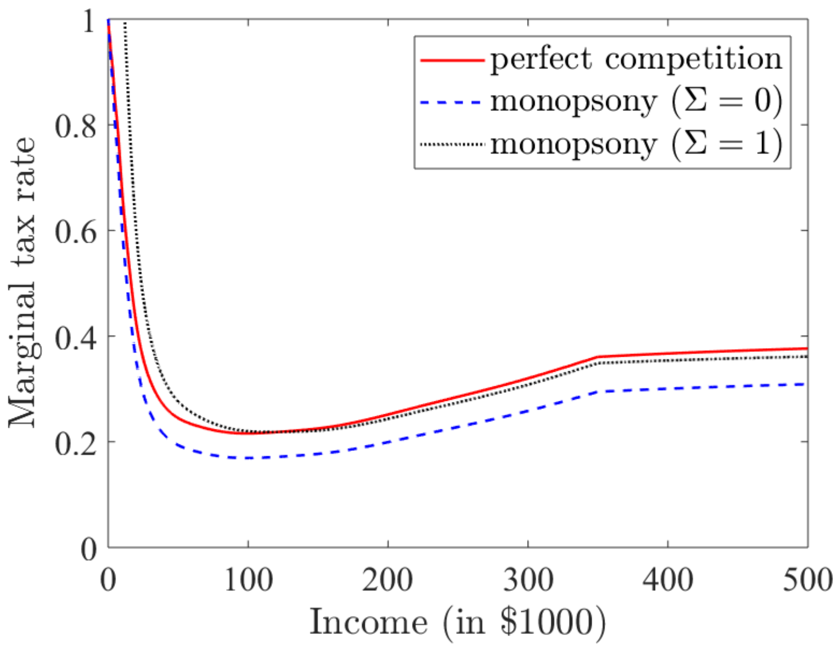


Figure 2: Optimal marginal tax rates

The blue, dashed line in Figure 2 plots the optimal marginal tax rates if the degree of monopsony power is as in the calibrated economy and the government does not value redistributing capital income: $\mu = 0.26$ and $\Sigma = 0$. Compared to the case with competitive labor markets, optimal marginal tax rates are lower, cf. Corollary 1. This is because monopsony power makes labor income taxes less effective in redistributing labor income as part of the incidence falls on firms. The average reduction in optimal marginal tax rates brought about by monopsony power is 5.6 percentage points.

The black, dotted line plots the optimal marginal tax rates if the government has a very strong preference for redistributing capital income: $\Sigma = 1$. Naturally, tax rates are higher compared to the case with $\Sigma = 0$. The average increase brought about by a change in the covariance between welfare weights and shareholdings is 13.2 percentage points. Compared to the case with competitive labor markets, optimal marginal tax rates are higher (lower) for individuals whose current labor earnings are below (above) approximately $122,000. On average, the optimal marginal tax rate with monopsony power is 7.7 percentage points higher. The increase is driven mostly by substantially higher marginal tax rates at low earnings levels, where the local Pareto parameter $a(n)$ is low: see Corollary 1 and Figure 6. The low Pareto parameter at the bottom also implies that some individuals do not work at the optimal allocation, as the constraint $l(n) \geq 0$ is binding. This is the case for individuals whose current labor earnings are below approximately $12,000.

According to Corollary 1, the impact of monopsony power on optimal tax rates is generally ambiguous. The analysis here suggests that if the government wishes to reduce inequality generated by differences in *both* ability and shareholdings, monopsony power tends to increase optimal marginal tax rates at lower earnings levels and to decrease optimal marginal tax rates at higher earnings levels. At what earnings level the impact changes from positive to negative depends critically on the covariance between welfare weights and shareholdings.

4.4 Implications for welfare

To assess the quantitative implications of monopsony power for welfare in the calibrated economy, I conduct two exercises. First, I calculate the welfare costs of ignoring monopsony power when designing tax policy. To do so, I compare the allocation that is obtained if the government sets income taxes optimally (cf. the dashed and dotted lines in Figure 2) with the one that is obtained if a "naive" government wrongfully sets tax policy *as if* labor markets are competitive (cf. the solid line in Figure 2). Second, I calculate how much the government is willing to pay for changing the degree of monopsony power to zero. The first exercise gives an indication of the welfare benefits of taking a *given* degree of monopsony power into account when designing tax policy, whereas the second exercise is informative about the costs or benefits of *changing* the degree of monopsony power.

Figure 3 shows the results of both exercises for different values of the covariance between welfare weights and shareholdings. The left axis plots the welfare costs of ignoring monopsony power when designing tax policy (i.e., the costs of "misoptimization"). The right axis plots the welfare effect of changing the degree of monopsony power from its value in the calibrated economy to zero. In both cases, the welfare impact is expressed in consumption equivalents as a percentage of current GDP in the calibrated economy. Regarding the first exercise, the welfare costs of ignoring monopsony power when designing tax policy range between \$57 and \$802 in consumption equivalents, or between 0.07% and 1.04% of GDP. These costs are small for low values of the negative covariance between welfare weights and shareholdings and largest if the government has a strong preference for redistributing capital income. To illustrate, moving from the solid to the dashed tax code plotted in Figure 2 generates a welfare gain equivalent to increasing all individuals' net income by \$70, or 0.09% of GDP. By contrast, moving from the solid to the dotted tax code plotted in Figure 2 generates a welfare gain equivalent to increasing all individuals' net income by \$802, or 1.04% of GDP.

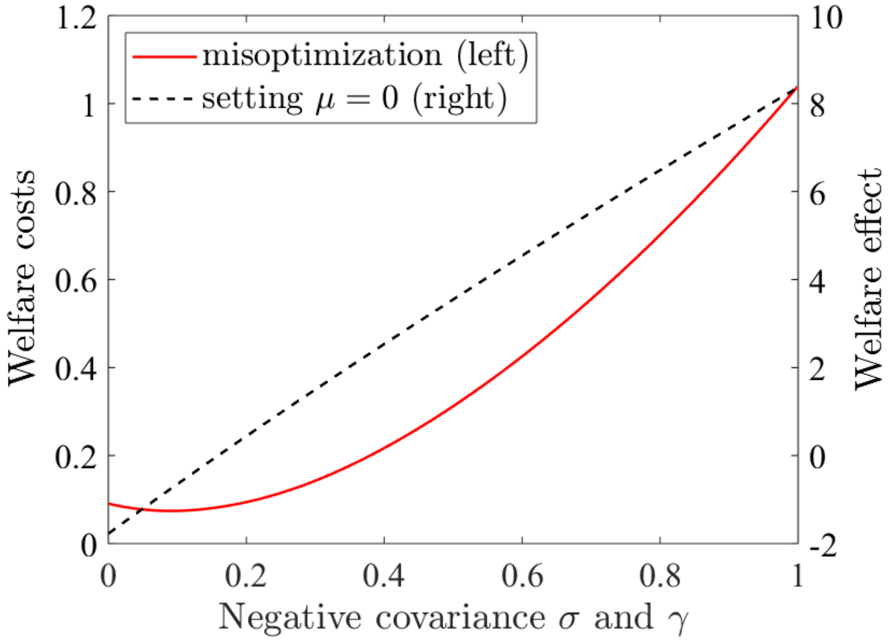


Figure 3: Welfare impact in consumption equivalents (% of GDP)

Regarding the second exercise, changing the degree of monopsony power from its value in the calibrated economy to zero can have a negative or positive impact on welfare depending on the covariance between shareholdings and welfare weights. If $\Sigma = 0$, getting rid of monopsony power leads to a welfare loss of \$1,370 in consumption equivalents, or 1.78% of GDP. This loss occurs because a reduction in monopsony power exacerbates labor income inequality and the government does not value the associated reduction in capital income inequality. By contrast, the welfare impact is positive if the government cares about redistributing capital income. In the calibrated economy, this happens whenever $\Sigma \geq 0.17$. If $\Sigma = 1$, the welfare gain of firms losing monopsony power is large and equals \$6,453 in consumption equivalents, or 8.37% of GDP.

The previous exercise illustrates that changing the degree of monopsony power from its value in the calibrated economy to zero and simultaneously re-optimizing the tax code can have a large negative or positive impact on welfare. It is also possible to analyze the welfare effect of a marginal increase in monopsony power at the current tax system provided the latter reflects the government's preferences for redistribution. Because labor effort is increasing in ability (see equation (22)), the result from Corollary 2 applies. Hence, an increase in monopsony power raises welfare only if

$$\left(\frac{\Sigma^{\ell}}{\Sigma^{k}}\right) > \left(\frac{1-\mu}{\mu}\right). \tag{28}$$

In the calibrated economy, the right-hand side equals approximately 2.85. Hence, if the current tax system is optimal, an increase in monopsony power raises welfare only if the negative covariance between welfare weights and after-tax labor income exceeds the negative covariance between welfare weights and after-tax capital income by a factor of at least this amount. If the preferences for redistribution are such that this condition is not satisfied at the current tax system, an increase in monopsony power lowers welfare.

To summarize, correcting the sub-optimal tax code by taking monopsony power into account leads to welfare gains that vary between 0.07% and 1.04% of current GDP in the calibrated economy. Moreover, changing the degree of monopsony power to zero has a welfare impact that ranges between –1.78% to +8.37% of GDP depending on the covariance between welfare weights and shareholdings. Finally, if the current tax system is optimal, an increase in monopsony power raises welfare only if the negative covariance between welfare weights and labor income exceeds the negative covariance between welfare weights and capital income by a factor of at least 2.85.

5 Conclusion

This paper extends the non-linear tax framework of Mirrlees (1971) with monopsony power and studies the implications for optimal income taxation and welfare. In my model, monopsony power does not reduce the size of the labor market surplus but determines what share is translated into pure economic profits. These profits flow back as capital income to individuals who differ in their ability and shareholdings.

Monopsony power makes labor income taxes less effective in redistributing labor income, but more effective in redistributing capital income. This is because monopsony power raises the tax incidence that falls on firms and lowers the tax incidence that falls on workers. The impact of monopsony power on optimal marginal tax rates is ambiguous and depends on the covariance between welfare weights and shareholdings, which captures the government's preference for redistributing capital income. Monopsony power raises optimal tax rates if the government cares strongly about redistributing capital income. A calibration of the model to the US economy suggests that monopsony power raises (lowers) optimal marginal tax rates at low (high) earnings levels if the government wishes to reduce inequality generated by differences in both ability and shareholdings. The welfare costs of ignoring monopsony power when designing tax policy range between 0.07% and 1.04% of GDP in the calibrated economy depending on the covariance between welfare weights and shareholdings.

An increase in monopsony power might increase or decrease welfare, as it mitigates (exacerbates) inequality in labor (capital) income. The reason why monopsony power might raise welfare is that firms observe ability, while the government does not. Monopsony power reduces inequality generated by differences in ability. This alleviates the trade-off between equity and efficiency that occurs if the government does not observe ability, but at the expense of increasing capital income inequality. In the calibrated economy, eliminating monopsony power has a welfare effect that ranges between –1.78% and +8.37% of GDP depending on the covariance between welfare weights and shareholdings. Moreover, if the current tax system is optimal, an increase in monopsony power raises welfare only if the negative covariance between welfare weights and after-tax labor income is at least 2.85 times as high as the negative covariance between welfare weights and after-tax capital income.

The analysis from this paper can be extended in at least two directions. First, in order to focus sharply on distributional issues I have abstracted from efficiency costs associated with monopsony power. Recent evidence by Berger et al. (2019) suggests these costs are significant. A natural way to introduce distortions from monopsony power in my model is to assume firms do not perfectly observe ability (as in Hariton and Piaser (2007) and da Costa and Maestri (2019)) or cannot offer contracts that specify both earnings and labor effort (as in Robinson (1933)). Second, I have treated monopsony power as exogenously determined. In reality monopsony power is unlikely to be policy-invariant. Extending the analysis with distortions from monopsony power and a potential role for the government to affect monopsony power (e.g., through competition policy) seems highly policy-relevant.

References

Azar, J., I. Marinescu, and M. Steinbaum (2017). Labor market concentration. NBER Working Paper No. 24147, Cambridge-MA: NBER.

Azar, J., I. Marinescu, and M. Steinbaum (2019). Measuring labor market power two ways. *AEA Papers and Proceedings 109*, 317–321.

Azar, J., I. Marinescu, M. Steinbaum, and B. Taska (2018). Concentration in US labor markets: evidence from online vacancy data. NBER Working Paper No. 24395, Cambridge-MA: NBER.

Barkai, S. (2019). Declining labor and capital shares. *Journal of Finance*. Forthcoming.

Barkai, S. and S. Benzell (2018). 70 years of US corporate profits. Stigler Center Working Paper No. 22, Chicago: Stigler Center.

Benmelech, E., N. Bergman, and H. Kim (2018). Strong employers and weak employees: how does employer concentration affect wages? NBER Working Paper No. 24307, Cambridge-MA: NBER.

Berger, D., K. Herkenhoff, and S. Mongey (2019). Labor market power. NBER Working Paper No. 25719, Cambridge-MA: NBER.

Bourguignon, F. and A. Spadaro (2012). Tax–benefit revealed social preferences. *Journal of Economic Inequality 10*(1), 75–108.

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.

CEA (2016). Labor market monopsony: trends, consequences, and policy responses. Issue Brief, White House Council of Economic Advisers, Washington, DC.

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.

FTC (2018a). Hearings on competition and consumer protection in the 21st century, no. 2: monopsony and the state of US antitrust law. Federal Trade Commission, Washington, DC.

FTC (2018b). Hearings on competition and consumer protection in the 21st century, no. 3: multi-sided platforms, labor markets and potential competition. Federal Trade Commission, Washington, DC.

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

Jacobs, B. (2018). The marginal cost of public funds is one at the optimal tax system. *International Tax and Public Finance 25*(4), 883–912.

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

Kline, P., N. Petkova, H. Williams, and O. Zidar (2019). Who profits from patents? Rent-sharing at innovative firms. *Quarterly Journal of Economics 134*(3), 1343–1404.

Krueger, A. (2018). Reflections on dwindling worker bargaining power and monetary policy. mimeo, Luncheon Address at the Jackson Hole Economic Symposium.

Lipsius, B. (2018). Labor market concentration does not explain the falling labor share. mimeo, University of Michigan.

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*(114), 175–208.

Pigou, A. (1920). *The economics of welfare.* London: Palgrave Macmillan UK.

Rinz, K. (2018). Labor market concentration, earnings inequality and earnings mobility. CARRA Working Paper No. 2018-10, Washington, DC: US Census Bureau.

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

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

Saez, E., B. Schoefer, and D. Seim (2019). Payroll taxes, firm behavior, and rent sharing: evidence from a young workers' tax cut in Sweden. *American Economic Review 109*(5), 1717–63.

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

Sandmo, A. (1994). Monopsonistic wage discrimination, incentives and efficiency. *Labour Economics 1*(2), 151–170.

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

Thomson, W. (1994). Cooperative models of bargaining. In R. Aumann and S. Hart (Eds.), *Handbook of Game Theory with Economic Applications*, Volume 2, Chapter 35, pp. 1237–1284. Amsterdam: Elsevier.

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

Webber, D. (2015). Firm market power and the earnings distribution. *Labour Economics 35*, 123–134.

Appendix

I Rewriting the welfare function

The result from Lemma 1 can be obtained as follows. Substitute the utility function (2) in the welfare function (8) and rewrite the resulting expression in a number of steps:

$$\begin{aligned}
\mathcal{W} &= \int_{\sigma_0}^{\sigma_1} \int_{n_0}^{n_1} \gamma(n,\sigma)\mathcal{U}(n,\sigma)h(n,\sigma)dnd\sigma \\
&= \int_{\sigma_0}^{\sigma_1} \int_{n_0}^{n_1} \gamma(n,\sigma)\Big[v(n) + \sigma(1-\tau)\bar{\pi}\Big]h(n,\sigma)dnd\sigma \\
&= \int_{n_0}^{n_1} v(n) \underbrace{\left(\int_{\sigma_0}^{\sigma_1} \gamma(n,\sigma)h(n,\sigma)d\sigma\right)}_{=g(n)f(n)} dn + (1-\tau)\bar{\pi} \int_{\sigma_0}^{\sigma_1} \int_{n_0}^{n_1} \sigma\gamma(n,\sigma)h(n,\sigma)dnd\sigma \\
&= \int_{n_0}^{n_1} g(n)v(n)f(n)dn + (1-\tau)\bar{\pi}\Bigg(1 + \underbrace{\int_{\sigma_0}^{\sigma_1} \int_{n_0}^{n_1} (\sigma-1)(\gamma(n,\sigma)-1)h(n,\sigma)dnd\sigma}_{=\mathrm{Cov}[\boldsymbol{\sigma},\boldsymbol{\gamma}]=-\Sigma}\Bigg) \\
&= \int_{n_0}^{n_1} g(n)v(n)f(n)dn + (1-\tau)(1-\Sigma)\int_{n_0}^{n_1} \pi(n)f(n)dn,
\end{aligned} \tag{29}$$

which corresponds to equation (9). To show that $\Sigma \in [0,1]$, write

$$\Sigma = -\int_{\sigma_0}^{\sigma_1} \int_{n_0}^{n_1} (\sigma-1)(\gamma(n,\sigma)-1)h(n,\sigma)dnd\sigma = 1 - \int_{\sigma_0}^{\sigma_1} \int_{n_0}^{n_1} \sigma\gamma(n,\sigma)h(n,\sigma)dnd\sigma. \tag{30}$$

Given that $\sigma \geq 0$ and $\gamma(n,\sigma) \geq 0$, it follows that $\Sigma \leq 1$. Next, write the covariance as

$$\begin{aligned}
\Sigma &= \int_{\sigma_0}^{\sigma_1} \int_{n_0}^{n_1} (1-\sigma)\gamma(n,\sigma)h(n,\sigma)dnd\sigma = \int_{\sigma_0}^{\sigma_1} (1-\sigma) \int_{n_0}^{n_1} \gamma(n,\sigma)h(n,\sigma)dnd\sigma \\
&= \int_{\sigma_0}^{\sigma_1} (1-\sigma) \underbrace{\left(\frac{\int_{n_0}^{n_1} \gamma(n,\sigma)h(n,\sigma)dn}{\int_{n_0}^{n_1} h(n,\sigma)dn}\right)}_{=\mathbb{E}[\gamma(n,\sigma)|\sigma]} \underbrace{\left(\int_{n_0}^{n_1} h(n,\sigma)dn\right)}_{=k(\sigma)} d\sigma.
\end{aligned} \tag{31}$$

By assumption, $\mathbb{E}[\gamma(n,\sigma)|\sigma]$ is non-increasing and σ averages to one. Therefore,

$$\Sigma = \int_{\sigma_0}^{\sigma_1} (1-\sigma)\mathbb{E}[\gamma(n,\sigma)|\sigma]k(\sigma)d\sigma \geq \int_{\sigma_0}^{\sigma_1} (1-\sigma)k(\sigma)d\sigma = 0. \tag{32}$$

II Optimal tax problem

To solve the optimal tax problem, I follow the approach from Mirrlees (1971) and let the government choose the allocation variables to maximize welfare (9) subject to resource and incentive constraints. The allocation variables are labor effort $l(n)$, the labor market payoff $\upsilon(n)$ and the profits $\pi(n)$ firms make from hiring a worker with ability n. To derive the resource constraint in terms of the allocation variables, substitute $T(z(n)) = z(n) - \upsilon(n) - \phi(l(n)) = nl(n) - \pi(n) - \upsilon(n) - \phi(l(n))$ in the government's budget constraint (10) and rearrange to find

$$\int_{n_0}^{n_1} nl(n)f(n)dn = \int_{n_0}^{n_1} \Big[\upsilon(n) + \phi(l(n)) + (1-\tau)\pi(n)\Big] f(n)dn + G. \tag{33}$$

In words, aggregate output equals the sum of private consumption (first term) and public consumption (second term).

In addition to the resource constraint, the allocation must also satisfy incentive constraints. To derive the first of these, differentiate the labor market payoff $\upsilon(n) = z(n) - T(z(n)) - \phi(l(n))$ with respect to ability to find

$$\upsilon'(n) = (1 - T'(z(n)))z'(n) - \phi'(l(n))l'(n). \tag{34}$$

Next, use the first-order condition from the profit maximization problem (4) and the relationship $\pi(n) = nl(n) - z(n)$. Condition (34) can then be written as

$$\upsilon'(n) = \frac{\phi'(l(n))}{n}\Big[l(n) - \pi'(n)\Big]. \tag{35}$$

This condition differs from the incentive constraint in the Mirrlees (1971) problem through the occurrence of the term $\pi'(n)$, which is zero if labor markets are competitive. The labor market payoff increases less quickly in ability if firms generate more profits from hiring individuals with higher ability.

To derive the second incentive constraint, differentiate the condition for profits (7) with respect to ability to find

$$\pi'(n) = \mu(n)l(n) + \frac{\mu'(n)}{\mu(n)}\pi(n). \tag{36}$$

Intuitively, profits increase more rapidly in ability the higher is monopsony power and labor effort. Profits increase less quickly in ability if individuals with higher ability suffer less from monopsony (i.e., if $\mu'(n) < 0$). Combining equations (35) and (36) gives

$$\upsilon'(n) = \frac{\phi'(l(n))}{n}\Big[(1-\mu(n))l(n) - \frac{\mu'(n)}{\mu(n)}\pi(n)\Big]. \tag{37}$$

By assumption, more productive workers do not suffer more from monopsony: $\mu'(n) \leq 0$. Equation (37) then implies that the labor market payoff is weakly increasing in ability: $\upsilon'(n) \geq 0$. The labor market payoff does not vary with ability if firms have full monopsony power (i.e., if $\mu(n) = 1$ for all n). In that case, all individuals are put on their identical participation constraint and hence, $\upsilon'(n) = 0$.

The government's problem consists of choosing the allocation variables $\upsilon(n)$, $\pi(n)$ and $l(n)$ at each ability level n to maximize welfare (9), subject to the resource constraint (33) and the incentive constraints (36) – (37). As it turns out, it is important to take the non-negativity constraint $l(n) \geq 0$ explicitly into account.[26] The final restriction we need to impose is that the profits from hiring the least productive workers are non-negative: $\pi(n_0) \geq 0$. This condition guarantees that firms are willing to hire individuals of all ability levels.[27] It is shown in Appendix V that this constraint is always binding, which *ex post* validates the assumption that $\pi(n_0) = 0$ in the description of the equilibrium: see Definition 2 and equation (7). The optimal tax problem can now be formulated as a standard optimal control problem where $\upsilon(n)$ and $\pi(n)$ are the state variables and $l(n)$ is the control variable. The corresponding Lagrangian and first-order conditions can be found in Appendix IV.

To make sure that the optimal allocation (as implicitly characterized in Appendix IV) can be decentralized using a tax on profits τ and a non-linear tax on labor income $T(z(n))$, I assume that earnings $z(n) = nl(n) - \pi(n)$ are increasing in ability whenever the non-negativity constraint on labor effort is not binding: $z'(n) > 0$ if $l(n) > 0$. This condition serves two purposes. First, it guarantees that individuals with different abilities do not earn the same income and hence, are not required to face the same marginal tax rate. Second, the monotonicity condition also ensures that the second-order condition for profit maximization is satisfied – see Appendix III for details.

III Monotonicity condition

This Appendix demonstrates the equivalence between the monotonicity condition $z'(n) > 0$ and the requirement that the second order-condition for the profit maximization problem (3) is satisfied. To do so, note that the constraint in the firm's maximization problem (3) is always binding. If not, firms can raise profits by increasing labor effort. Invert the constraint with respect to labor effort to write $l = \hat{l}(z, \upsilon(n))$, where $\upsilon(n) = \underline{\upsilon}(n)$ for all n. The profit

[26] To ensure consumption is non-negative, one could also include the constraint $\upsilon(n) + \phi(l(n)) \geq 0$ for all n. I assume the revenue requirement G and preferences for redistribution are such that this constraint never binds.

[27] To see why, note that the general solution to the differential equation (36) is

$$\pi(n) = \mu(n) \left[\frac{\pi(n_0)}{\mu(n_0)} + \int_{n_0}^{n} l(m)dm \right], \tag{38}$$

which simplifies to equation (7) if $\pi(n_0) = 0$. Because labor effort is non-negative, it follows that $\pi(n_0) \geq 0$ implies $\pi(n) \geq 0$ for all n.

maximization problem is

$$\max_{z \geq 0} \quad n\hat{l}(z, \upsilon(n)) - z. \tag{39}$$

By the implicit function theorem, $\hat{l}_z = (1 - T')/\phi'$, where I ignore function arguments to save on notation. At an interior solution, the first-order condition is given by

$$\frac{n(1 - T'(z))}{\phi'(\hat{l}(z, \upsilon(n)))} - 1 = 0. \tag{40}$$

The second-order condition is strictly satisfied if the left-hand side of equation (40) is strictly decreasing in earnings z. The latter is true if and only if

$$-\phi''(l) - n^2 T''(z) < 0, \tag{41}$$

where I used the first-order condition (40) and substituted out for $\hat{l}(z, \upsilon(n)) = l$. Because $\phi(\cdot)$ is strictly convex, condition (41) is satisfied as long as the tax function is not too concave.

To determine how earnings z vary with ability, rewrite equation (40) and define

$$L(z, n) \equiv n(1 - T'(z)) - \phi'(\hat{l}(z, \upsilon(n))) = 0. \tag{42}$$

Next, apply the implicit function theorem and use the first-order condition (40) and the property $\hat{l}_\upsilon = -1/\phi'$ to find

$$z'(n) = -\frac{L_n(z, n)}{L_z(z, n)} = \frac{\phi'(l) + \frac{\phi''(l)}{\phi'(l)} n\upsilon'(n)}{\phi''(l) + n^2 T''(z)}. \tag{43}$$

From the incentive constraint (37), $\upsilon'(n) \geq 0$ as long as monopsony power is non-increasing in ability. The numerator in (43) is therefore unambiguously positive. Hence, $z'(n) > 0$ if and only if the denominator is positive as well. This is the case if and only if the second-order condition (41) is satisfied. Therefore, if the allocation satisfies the monotonicity condition $z'(n) > 0$, it follows that the first-order condition for profit maximization (40) is both necessary and sufficient.

IV Lagrangian and first-order conditions

Written in terms of the allocation variables, the optimal tax problem is

$$\max_{[\upsilon(n), \pi(n), l(n)]_{n_0}^{n_1}} \mathcal{W} = \int_{n_0}^{n_1} \Big[g(n)\upsilon(n) + (1 - \Sigma)(1 - \tau)\pi(n) \Big] f(n) dn, \tag{44}$$

$$\text{s.t.} \quad \int_{n_0}^{n_1} \Big[nl(n) - \upsilon(n) - \phi(l(n)) - (1 - \tau)\pi(n) \Big] f(n) dn = G,$$

$$
\begin{aligned}
\forall n: \quad & \upsilon'(n) = \frac{\phi'(l(n))}{n}\left[(1-\mu(n))l(n) - \frac{\mu'(n)}{\mu(n)}\pi(n)\right], \\
\forall n: \quad & \pi'(n) = \mu(n)l(n) + \frac{\mu'(n)}{\mu(n)}\pi(n), \\
\forall n: \quad & l(n) \geq 0, \\
& \pi(n_0) \geq 0.
\end{aligned}
$$

The corresponding Lagrangian is given by

$$
\begin{aligned}
&\mathscr{L} = \qquad (45)\\
&\int_{n_0}^{n_1} \Bigg[\bigg(g(n)\upsilon(n) + (1-\Sigma)(1-\tau)\pi(n) + \eta\Big(nl(n) - \upsilon(n) - \phi(l(n)) - (1-\tau)\pi(n) - G\Big)\bigg)f(n) \\
&+ \chi(n)\frac{\phi'(l(n))}{n}\left((1-\mu(n))l(n) - \frac{\mu'(n)}{\mu(n)}\pi(n)\right) + \chi'(n)\upsilon(n) + \lambda(n)\left(\mu(n)l(n) + \frac{\mu'(n)}{\mu(n)}\pi(n)\right) \\
&+ \lambda'(n)\pi(n) + \psi(n)l(n)\Bigg]dn + \chi(n_0)\upsilon(n_0) - \chi(n_1)\upsilon(n_1) + \lambda(n_0)\pi(n_0) - \lambda(n_1)\pi(n_1) + \xi\pi(n_0).
\end{aligned}
$$

Suppressing the function argument of $\phi'(\cdot)$ and $\phi''(\cdot)$ to save on notation, the first-order conditions are given by

$$
\begin{aligned}
\upsilon(n): \quad & (g(n) - \eta)\, f(n) + \chi'(n) = 0, && (46)\\
\pi(n): \quad & (1-\tau)(1-\Sigma-\eta)f(n) - \frac{\mu'(n)}{\mu(n)}\left(\chi(n)\frac{\phi'}{n} - \lambda(n)\right) + \lambda'(n) = 0, && (47)\\
l(n): \quad & \eta\left(n - \phi'\right) f(n) + \frac{\chi(n)}{n}\left((1-\mu(n))(\phi' + \phi'' l(n)) - \phi''\frac{\mu'(n)}{\mu(n)}\pi(n)\right) && (48)\\
& + \lambda(n)\mu(n) + \psi(n) = 0, \\
\chi(n): \quad & \frac{\phi'}{n}\left((1-\mu(n))l(n) - \frac{\mu'(n)}{\mu(n)}\pi(n)\right) - \upsilon'(n) = 0, && (49)\\
\lambda(n): \quad & \mu(n)l(n) + \frac{\mu'(n)}{\mu(n)}\pi(n) - \pi'(n) = 0, && (50)\\
\eta: \quad & \int_{n_0}^{n_1}(nl(n) - \upsilon(n) - \phi(l(n)) - (1-\tau)\pi(n) - G)f(n)dn = 0, && (51)\\
\upsilon(n_0): \quad & \chi(n_0) = 0, && (52)\\
\upsilon(n_1): \quad & -\chi(n_1) = 0, && (53)\\
\pi(n_0): \quad & \lambda(n_0) + \xi = 0, && (54)\\
\pi(n_1): \quad & -\lambda(n_1) = 0, && (55)\\
\psi(n): \quad & \psi(n)l(n) = 0, \qquad \psi(n) \geq 0 \text{ and } l(n) \geq 0, && (56)\\
\xi: \quad & \xi\pi(n_0) = 0, \qquad \xi \geq 0 \text{ and } \pi(n_0) \geq 0. && (57)
\end{aligned}
$$

I assume the second-order conditions for the welfare maximization problem are satisfied and that earnings $z(n) = nl(n) - \pi(n)$ satisfy the monotonicity condition $z'(n) > 0$ if $l(n) > 0$.

V Derivation of the optimal marginal tax rate

This Appendix derives the optimal marginal tax rate in the general case where monopsony power $\mu(n)$ varies with ability. To that end, it is useful to first derive an expression for the multipliers $\chi(n)$ and $\lambda(n)$. Combining equations (46) and (53) gives

$$\chi(n) = \chi(n_1) - \int_n^{n_1} \chi'(m)dm = -\int_n^{n_1} (\eta - g(n))\, f(m)dm. \tag{58}$$

Evaluate equation (58) at $n = n_0$ and use the transversality condition (52) and the normalization $\int_{n_0}^{n_1} g(n)f(n)dn = 1$ to find

$$\int_{n_0}^{n_1} (\eta - g(n))\, f(n)dn = \eta - 1 = 0. \tag{59}$$

This is a standard result in optimal tax theory. When the tax system is optimized, the marginal cost of public funds equals one: see Jacobs (2018). Next, define by

$$\bar{g}(n) = \frac{\int_n^{n_1} g(m)f(m)dm}{1 - F(n)} \tag{60}$$

the average welfare weight of individuals with ability at least equal to n, so that $\chi(n) = -(1 - \bar{g}(n))(1-F(n))$. Because $\bar{g}(n_0) = 1$ and $g(n)$ is non-increasing in ability it follows that $\bar{g}(n) \leq 1$ and hence, $\chi(n) \leq 0$. To derive an expression for $\lambda(n)$, rewrite equation (47):

$$\lambda'(n) + \frac{\mu'(n)}{\mu(n)}\lambda(n) = (1-\tau)\Sigma f(n) - \frac{\mu'(n)}{\mu(n)} \frac{\phi'(l(n))}{n} \int_n^{n_1} (1 - g(m))f(m)dm, \tag{61}$$

where I used equation (58) to substitute out for $\chi(n)$. Equation (61) is a linear differential equation in $\lambda(n)$. Using the transversality condition (55), the solution is

$$\lambda(n) = -\frac{\bar{\mu}(n)}{\mu(n)}(1-\tau)\Sigma(1 - F(n)) + \int_n^{n_1} \frac{\mu'(m)}{\mu(n)} \frac{\phi'(l(m))}{m} \int_m^{n_1} (1 - g(s))f(s)dsdm, \tag{62}$$

where $\bar{\mu}(n)$ is the average monopsony power (i.e., one minus the bargaining power) of individuals with ability at least equal to n. To sign $\lambda(n)$, note that $\phi' \geq 0$ and monopsony power is non-increasing in ability. Consequently, $\lambda(n) \leq 0$. Equations (54) and (57) then imply $\xi \geq 0$. The assumption that firms do not earn profits from hiring the least productive workers (i.e., $\pi(n_0) = 0$) is therefore without loss of generality.

To derive an expression for the marginal tax rate, consider the first-order condition for labor effort (48). Because $\phi' = 0$ and $\pi(n) = 0$ if $l(n) = 0$, the non-negativity constraint on labor effort is binding (i.e., $\psi(n) > 0$) if

$$nf(n) - \bar{\mu}(n)(1-\tau)\Sigma(1 - F(n)) + \int_n^{n_1} \mu'(m)\frac{\phi'(l(m))}{m} \int_m^{n_1} (1 - g(s))f(s)dsdm < 0, \tag{63}$$

where I imposed $\eta = 1$ and substituted out for $\lambda(n)$ using equation (62). The latter is true if the local Pareto parameter

$$a(n) = \frac{nf(n)}{1-F(n)} < \bar{\mu}(n)(1-\tau)\Sigma - \frac{\int_n^{n_1} \mu'(m)\frac{\phi'(l(m))}{m}\int_m^{n_1}(1-g(s))f(s)dsdm}{1-F(n)}. \tag{64}$$

Hence, at ability levels where condition (64) holds, optimal labor effort and earnings are zero: $l(n) = 0$ and $z(n) = nl(n) - \pi(n) = 0$. If monopsony power does not vary with ability (i.e., if $\mu(n) = \mu$), the right-hand side simplifies to $\mu(1-\tau)\Sigma$. At ability levels where condition (64) does not hold, labor effort and earnings are positive. Substituting $\psi(n) = 0$, $\eta = 1$ and the first-order condition for profit maximization $n(1-T') = \phi'$ in equation (48) gives

$$T'(z(n))nf(n) = -\frac{\chi(n)}{n}\left((1-\mu(n))(\phi' + \phi''l(n)) - \phi''\frac{\mu'(n)}{\mu(n)}\pi(n)\right) - \mu(n)\lambda(n). \tag{65}$$

Substituting $\chi(n)$ and $\lambda(n)$ from equations (58) and (62), equation (65) can be written as

$$T'(z(n))nf(n) = (1-\bar{g}(n))(1-F(n))\frac{\phi'}{n}\left[(1-\mu(n))\left(1+\frac{\phi''l(n)}{\phi'}\right) - \pi(n)\frac{\phi''}{\phi'}\frac{\mu'(n)}{\mu(n)}\right]$$
$$+\bar{\mu}(n)(1-\tau)\Sigma(1-F(n)) - \int_n^{n_1}\mu'(m)\frac{\phi'(l(m))}{m}\left(\int_m^{n_1}(1-g(s))f(s)ds\right)dm. \tag{66}$$

Next, use the condition $n(1-T') = \phi'$ and denote by $\varepsilon(n) = \frac{\phi'}{\phi''l(n)}$ the elasticity of labor supply. Upon dividing equation (66) by $nf(n)$ and rearranging, we obtain equation (18) from Proposition 3:

$$T'(z(n)) = \frac{1-F(n)}{nf(n)}\left[\bar{\mu}(n)(1-\tau)\Sigma + (1-\mu(n))(1-T'(z(n)))\left(1+\frac{1}{\varepsilon(n)}\right)(1-\bar{g}(n))\right. \tag{67}$$
$$\left. - \frac{\mu'(n)\pi(n)(1-T'(z(n)))(1-\bar{g}(n))}{\mu(n)\varepsilon(n)l(n)} - \frac{\int_n^{n_1}\mu'(m)(1-T'(z(m)))\left(\int_m^{n_1}(1-g(s))f(s)ds\right)dm}{1-F(n)}\right].$$

If monopsony power does not vary with ability (i.e., $\mu'(n) = 0$), the last two terms cancel. Substituting $\mu(n) = \bar{\mu}(n) = \mu$ gives equation (11) from Proposition 1.

From equation (67) it follows immediately that the optimal marginal tax rate is zero at the top: $T'(z(n_1)) = 0$. To show that the optimal marginal tax rate is generally positive, note that monopsony power is non-increasing: $\mu'(n) \leq 0$. Moreover, $\bar{g}(n) \leq 1$ and from the profit-maximization condition (4) it follows that the marginal tax rate cannot exceed one at an interior solution. Hence, all terms on the right-hand side of equation (67) are non-negative.

VI Impact of monopsony power on optimal marginal tax rates

To derive an expression for the optimal marginal tax rate if the utility function is iso-elastic (i.e., $\phi(l) = l^{1+1/\varepsilon}/(1+1/\varepsilon)$), substitute $\varepsilon(n) = \varepsilon$ in equation (11) and use the definition of $a(n)$. Rearranging gives the result from Corollary 1:

$$T'(z(n)) = \frac{\mu(1-\tau)\Sigma + (1-\mu)(1+1/\varepsilon)(1-\bar{g}(n))}{a(n) + (1-\mu)(1+1/\varepsilon)(1-\bar{g}(n))}. \tag{68}$$

This is a closed-form solution for the optimal marginal tax rate. To determine how the latter varies with monopsony power, differentiate equation (68) with respect to μ to find

$$\frac{\partial T'(z(n))}{\partial \mu} = \frac{a(n)((1-\tau)\Sigma - \left(1+\frac{1}{\varepsilon}\right)(1-\bar{g}(n))) + (1-\tau)\Sigma\left(1+\frac{1}{\varepsilon}\right)(1-\bar{g}(n))}{(a(n) + (1-\mu)\left(1+\frac{1}{\varepsilon}\right)(1-\bar{g}(n)))^2}. \tag{69}$$

Equation (69) is positive if and only if the numerator is positive. Because $\bar{g}(n_0) = 1$, this is always the case at the bottom of the income distribution if $(1-\tau)\Sigma > 0$. At higher ability levels, the impact of monposony power on optimal tax rates is generally ambiguous. To see why, note that $\bar{g}(n) < 1$ for all $n > n_0$ if the government wishes to reduce inequality generated by differences in ability. To derive the result from the corollary, divide the numerator in equation (69) by $a(n)(1-\tau)\Sigma(1-\bar{g}(n)) > 0$. The resulting expression is positive if and only if

$$((1-\tau)\Sigma))^{-1} < ((1+1/\varepsilon)(1-\bar{g}(n)))^{-1} + a(n)^{-1}. \tag{70}$$

VII Welfare effect of raising monopsony power

This Appendix analyzes the welfare effect of a proportional increase in monopsony power by α percent, starting from a situation where monopsony power might vary with ability. Hence, after the increase monopsony power is $\hat{\mu}(n) = \mu(n)(1+\alpha)$. Welfare is then given by

$$\begin{aligned}\mathscr{L}(\alpha) = \int_{n_0}^{n_1} \Bigg[&\Big((g(n)-\eta)\upsilon(n) + (1-\Sigma-\eta)(1-\tau)\pi(n) + \eta(nl(n) - \phi(l(n)) - G)\Big)f(n) \\ &+ \chi'(n)\upsilon(n) + \chi(n)\frac{\phi'(l(n))}{n}\left((1-\mu(n)(1+\alpha))l(n) - \frac{\mu'(n)}{\mu(n)}\pi(n)\right) + \lambda'(n)\pi(n) \\ &+ \lambda(n)\left(\mu(n)(1+\alpha)l(n) + \frac{\mu'(n)}{\mu(n)}\pi(n)\right) + \psi(n)l(n)\Bigg]dn + \chi(n_0)\upsilon(n_0) - \chi(n_1)\upsilon(n_1) \\ &+ \lambda(n_0)\pi(n_0) - \lambda(n_1)\pi(n_1) + \xi\pi(n_0),\end{aligned} \tag{71}$$

which is the optimized Lagrangian (45) evaluated at $\hat{\mu}(n) = \mu(n)(1+\alpha)$. Here I used the fact that the increase in monopsony power is proportional, which implies

$$\frac{\hat{\mu}'(n)}{\hat{\mu}(n)} = \frac{\mu'(n)(1+\alpha)}{\mu(n)(1+\alpha)} = \frac{\mu'(n)}{\mu(n)}. \tag{72}$$

By the Envelope theorem, the welfare effect is

$$\frac{\partial \mathcal{W}(\alpha)}{\partial \alpha} = \frac{\partial \mathscr{L}(\alpha)}{\partial \alpha} = \int_{n_0}^{n_1} \left(-\chi(n)\frac{\phi'}{n} + \lambda(n) \right) \mu(n) l(n) dn. \tag{73}$$

Next, use equations (58) and (62) to substitute out for $\chi(n)$ and $\lambda(n)$. This leads to equation (19) as stated in Proposition 3:

$$\frac{\partial \mathcal{W}(\alpha)}{\partial \alpha} = \int_{n_0}^{n_1} \left[\frac{\phi'(l(n))}{n} \int_{n}^{n_1} (1-g(m)) f(m) dm - (1-\tau)\Sigma \int_{n}^{n_1} \frac{\mu(m)}{\mu(n)} f(m) dm \right.$$
$$\left. + \int_{n}^{n_1} \frac{\mu'(m)}{\mu(n)} \frac{\phi'(l(m))}{m} \left(\int_{m}^{n_1} (1-g(s)) f(s) ds \right) dm \right] \mu(n) l(n) dn. \tag{74}$$

This expression simplifies considerably if monopsony power does not vary with ability. The term in the second line of equation (74) cancels. Substituting $\mu(n) = \mu$ gives

$$\frac{\partial \mathcal{W}(\alpha)}{\partial \alpha} = \tag{75}$$
$$\int_{n_0}^{n_1} \left[\frac{\mu}{1-\mu} \underbrace{(1-\mu)\frac{\phi'(l(n))l(n)}{n}}_{=\upsilon'(n)} \int_{n}^{n_1} (1-g(m)) f(m) dm - (1-\tau)\Sigma \underbrace{\mu l(n)}_{=\pi'(n)} \int_{n}^{n_1} f(m) dm \right] dn.$$

Apply integration by parts with boundary conditions $\bar{g}(n_0) = 1$ and $\pi(n_0) = 0$:

$$\frac{\partial \mathcal{W}(\alpha)}{\partial \alpha} = \int_{n_0}^{n_1} \left[\frac{\mu}{1-\mu} (1-g(n))\upsilon(n) - (1-\tau)\Sigma\pi(n) \right] f(n) dn. \tag{76}$$

The latter can be simplified further after defining

$$\Sigma^{\upsilon} = -\text{Cov}[\boldsymbol{\upsilon}, \boldsymbol{\gamma}] \geq 0, \tag{77}$$
$$\Sigma^{k} = -\text{Cov}[\boldsymbol{\sigma}(1-\tau)\bar{\pi}, \boldsymbol{\gamma}] = \Sigma(1-\tau)\bar{\pi} \geq 0. \tag{78}$$

The first measures the negative covariance between labor market payoffs $\upsilon(n)$ and welfare weights $\gamma(n,\sigma)$. The second measures the negative covariance between capital income $\sigma(1-\tau)\bar{\pi}$ and welfare weights. It is proportional to the covariance between shareholdings and welfare weights introduced before. Substituting these terms in (76) gives

$$\frac{\partial \mathcal{W}(\alpha)}{\partial \alpha} = \frac{\mu}{1-\mu}\Sigma^{\upsilon} - \Sigma^{k}. \tag{79}$$

From this relationship, it immediately follows that if the tax system is optimized, an increase in monopsony power raises welfare if and only if (cf. Proposition 2)

$$\mu\Sigma^{\upsilon} > (1-\mu)\Sigma^{k}. \tag{80}$$

As stated in the main text, it is possible to derive an expression for the welfare effect of raising monopsony power in terms of exogenous variables if the utility function is iso-elastic: $\phi(l) = l^{1+1/\varepsilon}/(1+1/\varepsilon)$. To see this, recall that Corollary 1 gives a closed-form expression for the marginal tax rate:

$$T'(z(n)) = \frac{\mu(1-\tau)\Sigma + (1-\mu)(1+1/\varepsilon)(1-\bar{g}(n))}{a(n) + (1-\mu)(1+1/\varepsilon)(1-\bar{g}(n))}, \tag{81}$$

provided $a(n) \geq \mu(1-\tau)\Sigma$. Labor effort can then be determined from equation (4):

$$l(n) = n^{\varepsilon}\left(\frac{a(n) - \mu(1-\tau)\Sigma}{a(n) + (1-\mu)(1+1/\varepsilon)(1-\bar{g}(n))}\right)^{\varepsilon} \tag{82}$$

and $l(n) = 0$ if $a(n) < \mu(1-\tau)\Sigma$. Denote by $n' \geq n_0$ the participation threshold, which is the highest ability level where the non-negativity constraint on labor effort $l(n) \geq 0$ binds. Substituting the above in equation (74) and setting $\mu(n) = \mu$ and hence, $\mu'(n) = 0$ gives

$$\begin{aligned}\frac{\partial \mathcal{W}(\alpha)}{\partial \alpha} = \int_{n'}^{n_1} \mu\bigg[\left(\frac{a(n) - \mu(1-\tau)\Sigma}{a(n) + (1-\mu)(1+1/\varepsilon)(1-\bar{g}(n))}\right)(1-\bar{g}(n)) - (1-\tau)\Sigma\bigg] \\ \times (1-F(n))n^{\varepsilon}\left(\frac{a(n) - \mu(1-\tau)\Sigma}{a(n) + (1-\mu)(1+1/\varepsilon)(1-\bar{g}(n))}\right)^{\varepsilon} dn,\end{aligned} \tag{83}$$

which is expressed solely in terms of exogenous variables.

To derive the result from Corollary 2, note that equation (80) gives a necessary and sufficient condition to determine if an increase in monopsony power raises welfare. Next, write

$$\begin{aligned}\Sigma^{v} &= \int_{n_0}^{n_1}(1-g(n))v(n)f(n)dn = \int_{n_0}^{n_1}(1-g(n))(z(n) - T(z(n)) - \phi(l(n)))f(n)dn \\ &= \int_{n_0}^{n_1}(1-g(n))(z(n) - T(z(n)))f(n)dn - \int_{n_0}^{n_1}(1-g(n))\phi(l(n))f(n)dn \\ &= -\text{Cov}[\boldsymbol{z} - T(\boldsymbol{z}), \boldsymbol{\gamma}] - \int_{n_0}^{n_1}(1-g(n))\phi(l(n))f(n)dn \\ &= \Sigma^{\ell} - \int_{n_0}^{n_1}(1-g(n))\phi(l(n))f(n)dn.\end{aligned} \tag{84}$$

Because $g(n)$ is weakly decreasing in ability and averages to one, the second term on the last line of equation (84) is non-negative if labor effort is weakly increasing in ability. Therefore, $\Sigma^{\ell} \geq \Sigma^{v}$ if $l'(n) \geq 0$. In that case, an increase in monopsony power raises welfare only if

$$\mu\Sigma^{\ell} > (1-\mu)\Sigma^{k}. \tag{85}$$

Unlike equation (80), this condition is only necessary and not sufficient.

VIII Additional figures

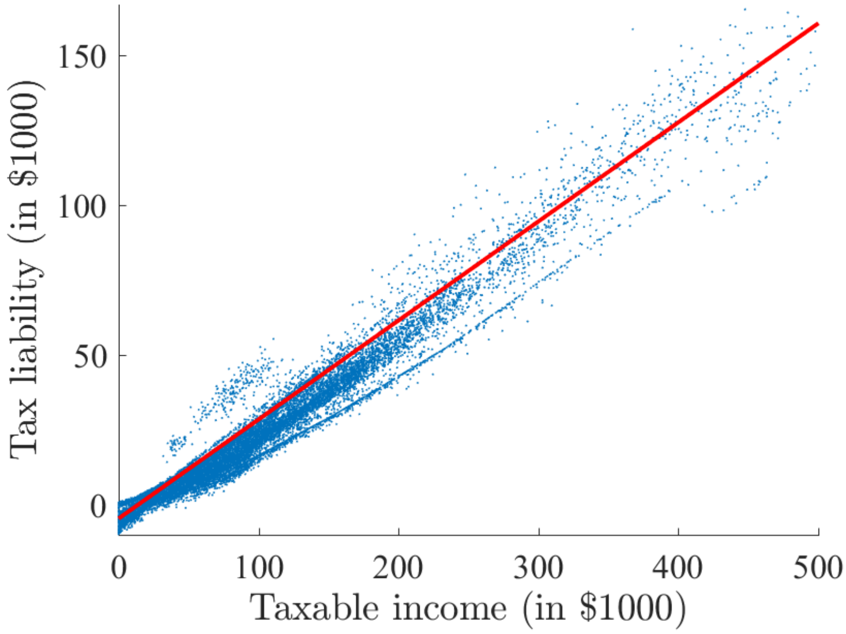


Figure 4: Current tax schedule

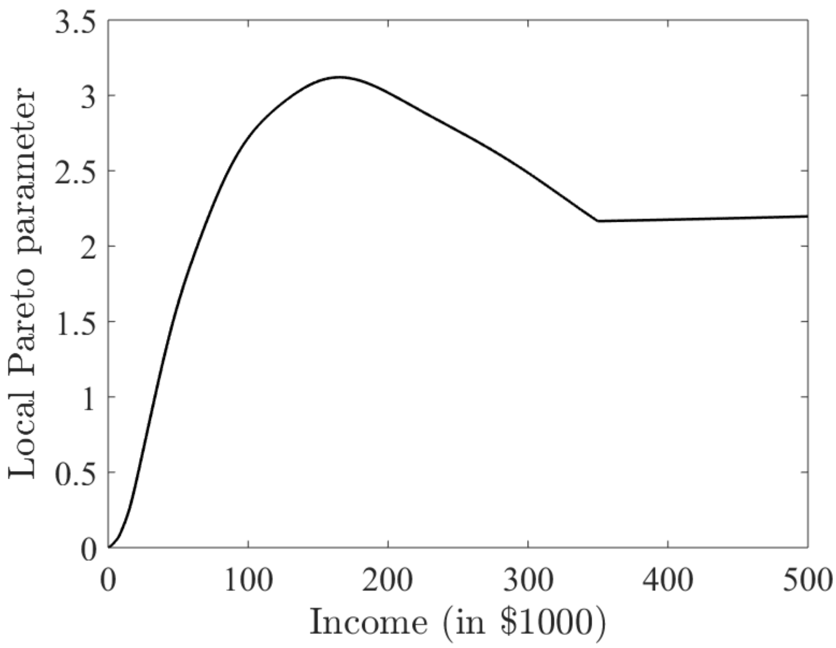


Figure 6: Local Pareto parameter

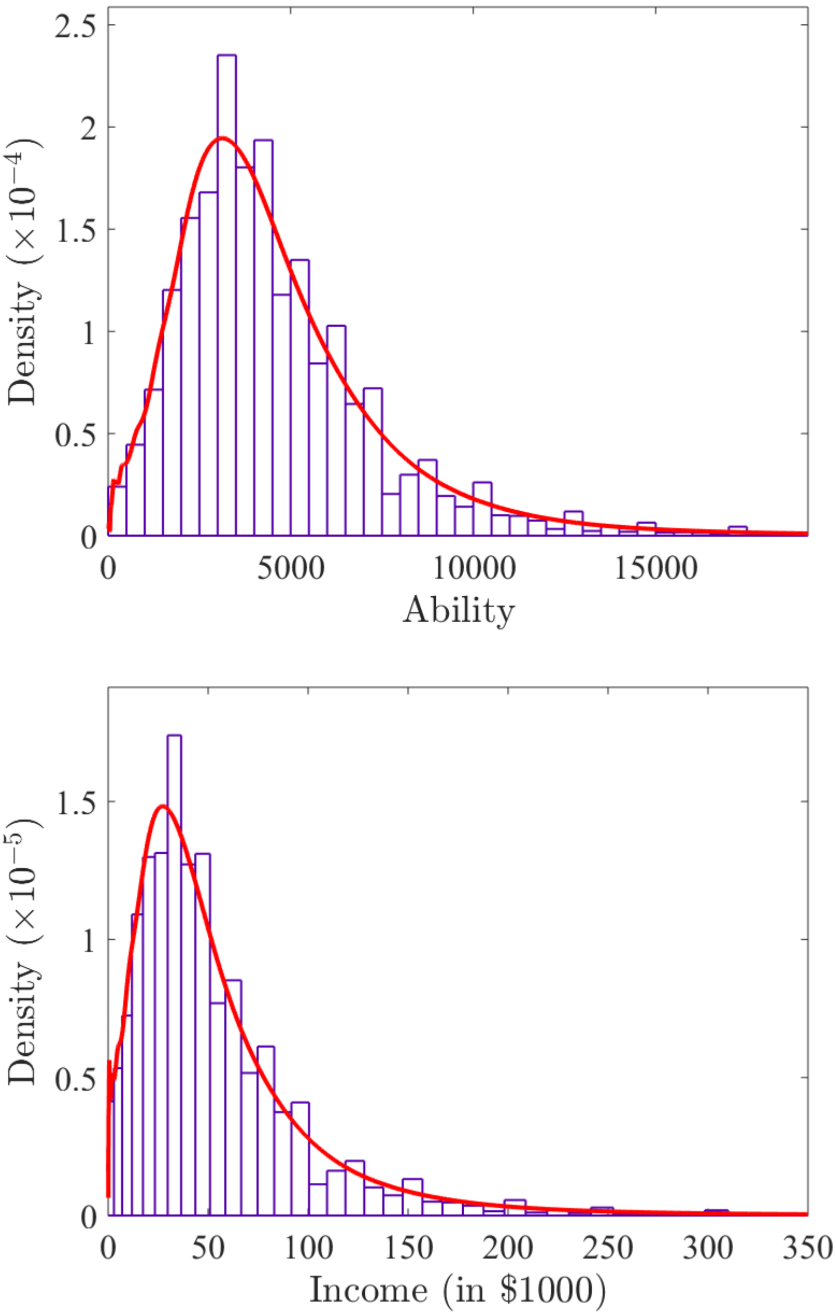


Figure 5: Distribution of ability and income