UvA-DARE (Digital Academic Repository)

Non-linear taxation with monopsony power

Hummel, A.J.

Publication date
2019

Document Version
Submitted manuscript

Link to publication

Citation for published version (APA):
Hummel, A. J. (2019). *Non-linear taxation with monopsony power*.
https://www.albertjanhummel.com/wp-content/uploads/2018/09/Hummel_monopsony.pdf

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)

Download date:29 Sep 2022

Non-linear taxation with monopsony power*

Albert Jan Hummel†

October 30, 2019

Click here for the latest version

Abstract

How does monopsony power affect optimal income taxation and welfare? I study this question in a Mirrleesian framework where firms observe workers' abilities while the government does not. Monopsony power does not generate efficiency losses, but determines what share of the labor market surplus is translated into profits. Monopsony power increases the tax incidence borne by firms and mitigates (exacerbates) inequality in labor (capital) income. I derive intuitive expressions for the optimal marginal tax rate on labor income and the welfare effect of raising monopsony power. Monopsony power has an ambiguous effect on optimal tax rates and welfare. On the one hand, monopsony power enables the government to use firms as a screening device in order to extract information on hidden ability. This lowers optimal tax rates and raises welfare. On the other hand, monopsony power generates a distributional conflict over profits. This raises optimal tax rates and lowers welfare provided profits flow back to individuals with below-average welfare weights. I calibrate the model to the US economy and find that monopsony power raises optimal tax rates at low earnings levels and lowers optimal tax rates for middle- and high-income earners. Moreover, the welfare effect of eliminating monopsony power is very sizable and ranges between –3.4% and +6.8% of GDP depending on the government's redistributive preferences.

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 and Kevin Spiritus for helpful comments and suggestions. This paper has benefited greatly from discussions at the University of Amsterdam and Erasmus University Rotterdam.

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

1 Introduction

In recent years, labor market monopsony transformed from a textbook curiosum to a topic actively debated among economists and policymakers. In 2016, the Council of Economic Advisers published an issue brief expressing concerns about growing monopsony power (CEA (2016)) and in 2018, the Federal Trade Commission devoted a number of sessions to the topic during its hearings on competition and consumer protection (FTC (2018a,b)). Both the CEA report and FTC 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)).[1] The prevalence of monopsony power is argued to depress employment and output, thereby generating efficiency losses. This has led to a number of policy proposals aimed at limiting monopsony power in order to enhance efficiency. Examples include stricter enforcement of anti-trust laws and restricting the use of non-compete clauses to promote worker mobility (CEA (2016)).

Low wages due to monopsony power need not be a sign of large efficiency losses *per se*, but could also reflect workers' inability to capture a sizable share of the labor market surplus. For example, if firms engage in price discrimination it is in their best interest to set employment at its efficient level. Moreover, if firms can recruit or retain workers through other means than just the hourly wage (e.g., working hours, flexible schedules, work-from-home options, vacation time, etc.), they have an incentive to offer contracts which maximize the joint firm-worker surplus, so that more rents can possibly be extracted. In such cases, monopsony power has important implications for the distribution of income, without doing much harm to efficiency. This view is expressed by Alan Krueger 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))

What are the policy and welfare implications if monopsony power affects the distribution of income without generating efficiency losses? I study this question by extending the non-linear tax framework from Mirrlees (1971) with monopsony power. Crucially, in my model

[1] Using online vacancy data for the US, Azar et al. (2018) measure concentration using the Herfindahl-Hirschman index and find that 60% of local labor markets (defined by a combination of occupation and commuting zone) accounting for 20% of total employment are highly concentrated according to DOJ/FTC guidelines. Moreover, Azar et al. (2017) estimate that moving from the 25th to the 75th percentile in concentration is associated with a 17% decline in posted wages. Similarly, using administrative data for the US, Rinz (2018) estimates that moving from the 25th to the 75th percentile in concentration is associated with a 20% decline in earnings.

firms observe workers' abilities while the government does not. Firms offer workers a combination of earnings and labor effort to maximize profits, subject to promising workers their reservation utility. In the absence of taxes on labor earnings, the resulting equilibrium is efficient.[2] The main departure from Mirrlees (1971) is that I allow for the possibility that firms have monopsony power, which determines what share of the labor market surplus is translated into pure economic profits. These profits are taxed at an exogenous, maximum rate. After-tax profits flow back to workers who differ in their ability and to capitalists, who do not work and hold significantly more shares than workers on average. The model features inequality in labor income driven by differences in workers' abilities and inequality in capital income driven by differences in shareholdings. The government cares for redistribution: it attaches welfare weights to workers that are declining in ability and which, on average, exceed the welfare weight of capitalists. It observes only labor earnings but not workers' abilities or how pure economic profits are dissipated. Consequently, in addition to a linear tax on aggregate profits it levies 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 joint firm-worker surplus and monopsony power determines what share of the surplus accrues to firms. As a result, income taxes reduce profits unless labor markets are perfectly competitive (in which case profits are zero – irrespective of taxation). Second, monopsony power reduces inequality in labor income but increases inequality in capital income. Intuitively, monopsony power lowers the share of the labor market surplus that accrues to workers in the form of wages and raises the share that accrues to firms in the form of profits. As a result, any dispersion in labor (capital) income generated by differences in abilities (shareholdings) is mitigated (exacerbated) if firms capture a larger share of the surplus.

How should the government design tax policy in this environment? I derive an intuitive expression for the optimal marginal tax rate on labor income and show that income taxes not only serve to redistribute labor income, but also to redistribute capital income. Part of the incidence of income taxes falls on firms if they have monopsony power. As a result, monopsony power makes income taxes less effective in redistributing labor income, but more effective in redistributing capital income by lowering aggregate profits. Whether monopsony power raises or lowers optimal tax rates is *a priori* ambiguous. I derive a precise condition under which monopsony power raises the optimal tax rate at each point in the income distribution. This condition is always satisfied at low earnings levels. Intuitively, marginal tax rates at the bottom do not help to redistribute labor income, but are helpful to reduce inequality in cap-

[2] Hence, there are no unexploited gains from trade. This property stands in sharp contrast to the classic model of monopsony based on Robinson (1969) and the new monopsony models introduced in Manning (2003). In both types of models, the equilibrium in the absence of taxation is typically inefficient.

ital income by lowering aggregate profits. The latter becomes more important if monopsony power increases. I also show that monopsony power is more likely to raise optimal tax rates if profits flow back to individuals with below-average welfare weights. In this case, income taxes are particularly useful to redistribute capital income.

Monopsony power has an ambiguous effect on welfare. On the one hand, monopsony power increases inequality in capital income, which lowers welfare. On the other hand, monopsony power decreases inequality in labor income, which raises welfare. The reason why monopsony power might raise welfare is that it enables the government to use firms as a screening device. Intuitively, firms observe workers' abilities, while the government does not. Firms do not benefit from this information if labor markets are competitive. In this case, profits are zero and the full labor market surplus accrues to workers. However, if firms have monopsony power part of the surplus is not captured by workers but translated into profits. This leads to a compression of the labor income distribution, which lowers the need for using distortionary taxes to redistribute from high-abiltiy to low-ability workers. Monopsony power thus enables the government to exploit the informational advantage of firms, which alleviates the equity-efficiency trade-off that results from the assumption that ability is not observable to the government, cf. Mirrlees (1971). At the same time, however, monopsony power generates a distributional conflict over profits. Monopsony power raises aggregate profits, which reduces welfare if profits flow back to individuals with below-average welfare weights. Whether monopsony power raises or lowers welfare therefore crucially depends on how pure economic profits are dissipated.

In the baseline version of the model, I assume all workers hold an equal number of shares and inequality in capital income is driven solely by the assumption that capitalists hold more shares than workers. Moreover, in the baseline I assume all workers suffer to the same extent from monopsony power in the sense that in the absence of taxation, firms capture a constant (i.e., non ability-specific) share of the labor market surplus. In an extension I relax these assumptions and allow for the possibility that more productive workers own more shares and suffer less from monopsony. As before, I derive an expression for the optimal tax rate on labor income and for the welfare effect of raising monopsony power. Compared to the case where shareholdings and monopsony power do not vary with ability, optimal tax rates are higher and monopsony power is more likely to be welfare-reducing. Intuitively, inequality in total (labor and capital) income driven by differences in workers' abilities is exacerbated if more productive workers own more shares and suffer less from monopsony.

To assess the quantitative implications of monopsony power for optimal income taxation and welfare, I calibrate the baseline version of the model to match key moments of the US economy. Following Saez (2001), I infer the ability distribution from the empirical income distribution. Moreover, I calibrate the degree of monopsony power to be consistent with an estimate of the pure profit share as obtained by Barkai and Benzell (2018). The model is subsequently used to calculate optimal tax rates and the welfare effects of monopsony power for different specifications of the welfare function. I find that optimal tax rates are quite different if monopsony power is taken into account. For common specifications of the welfare function, monopsony power raises optimal tax rates at low levels of income and lowers optimal tax rates for middle- and high-income earners. Regarding the implications for welfare, the calibration exercise suggests that the costs of ignoring monopsony power when designing tax policy range between 0.2% and 0.4% of GDP. Moreover, the welfare effect eliminating monopsony power is very sizable and ranges from –3.4% to +6.8% of GDP depending on the redistributive preferences of the government.

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 (2018) analyze a model where labor supply responds on the intensive (hours, effort) margin, whereas Cahuc and Laroque (2014) focus on the extensive (participation) margin. Crucially, these studies assume that firms – like the government – do *not* observe workers' abilities (Hariton and Piaser (2007) and da Costa and Maestri (2018)) or their reservation wages (Cahuc and Laroque (2014)).[3] Monopsony power then leads to a downward distortion in employment, either in hours worked or the number of persons employed. To partly off-set this distortion, the government finds it optimal to subsidize employment. This requires *negative* marginal tax rates (levels of taxation) if labor supply responds along the intensive (extensive) margin. By contrast, in my model firms observe ability and as a result, 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 enables the government to use firms as a screening device. This is not possible in Hariton and Piaser (2007), Cahuc and Laroque (2014) and da Costa and Maestri (2018), since firms do not have an informational advantage compared to the government. A final difference is that I assume profit taxation is restricted. As a result, monopsony power generates a distributional conflict over profits. I show that in this case, income taxes can be used to indirectly tax profits.

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 too 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 does not affect welfare.[4] My paper differs from Kaplow (2019) in two important

[3] See Stantcheva (2014) and Bastani et al. (2015) for an analysis of optimal taxation if firms do not not observe ability, but also do not have monopsony power.

[4] If mark-ups vary across goods, market power does affect welfare. Kaplow (2019) shows that policies aimed at

ways. First, I do not assume firms charge a constant mark-up over labor costs but instead offer employees a combination of labor earnings and an effort requirement. As a result, the outcome in the absence of taxation is efficient and tax policy is exclusively aimed at redistribution – not to restore efficiency. Second, in my model monopsony power affects welfare despite the fact that there is only one good and hence, no variation in mark-ups. The reason is that tax policy cannot off-set the impact of monopsony power if firms do not charge a constant mark-up but instead offer combinations of earnings and effort requirements. As a result, monopsony power has welfare effects.

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 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 workers 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 workers' 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 positive profits. I assume a share of these profits is taxed, a share flows back to workers and a share goes to capitalists, who do not work and derive their income exclusively from holding shares. Consequently, the model features both inequality in labor income generated by differences in workers' abilities and inequality in capital income generated by differences in shareholdings. Both types of inequality play an important role in the remainder of the analysis.

2.1 Individuals

There are two types of individuals in the economy: workers and capitalists. Workers derive their income from providing labor effort and possibly from holding shares. The first (second) source of income is referred to as labor (capital) income. Workers are indexed by their ability $n \in [n_0, n_1]$, which measures how much output a worker produces per unit of effort.

reducing the *spread* in mark-ups are generally welfare-enhancing.

Ability is distributed according to a cumulative distribution function $F(n)$ with density $f(n)$. The mass of workers is normalized to one. Capitalists derive their income exclusively from holding shares. Hence, they only earn capital income. I abstract from heterogeneity among capitalists, and let their measure be denoted by $\kappa \geq 0$.

Individuals derive utility from consumption c and disutility from providing labor effort l. Their preferences are described by a quasi-linear and separable utility function

$$u(c, l) = c - \phi(l), \tag{1}$$

where $\phi(\cdot)$ is strictly increasing, strictly convex and I normalize $\phi(0) = 0$. By assumption, capitalists do not exert labor effort. I abstract from non-participation among workers and denote by $l(n)$ the labor effort exerted by a worker with ability n. In exchange for her services, a worker receives labor income $z(n)$, which is subject to a labor income tax $T(\cdot)$. Like capitalists, workers may also generate income from shareholdings. I assume all individuals hold shares in an aggregate portfolio, so their capital income is proportional to the economy's aggregate profits. Let $\pi(n) = nl(n) - z(n)$ denote the profits firms generate from hiring a worker with ability n. Aggregate profits are then given by

$$\pi^* = \int_{n_0}^{n_1} \pi(n) f(n) dn. \tag{2}$$

Profits are taxed linearly at an exogenous rate $\tau \in [0, 1]$ and after-tax profits flow back to workers and capitalists according to their shareholdings. Let $\sigma(k)$ denote the shares held by a capitalist and $\sigma(n)$ the shares held by a worker with ability n. To make sure high-ability workers are not worse off, I assume more productive workers do not own fewer shares: $\sigma'(n) \geq 0$. Moreover, because capitalists derive their income exclusively from firm-ownership, I assume they hold significantly more shares than workers on average: $\sigma(k) \gg \int_{n_0}^{n_1} \sigma(n) f(n) dn$. The aggregate measure of shares is normalized to one:

$$\int_{n_0}^{n_1} \sigma(n) f(n) dn + \sigma^*(k) = 1, \tag{3}$$

where $\sigma^*(k) = \kappa\sigma(k)$ denotes the aggregate shares held by capitalists. We can now write the utility of a worker with ability n and a capitalist as

$$v(n) = z(n) - T(z(n)) + \sigma(n)(1 - \tau)\pi^* - \phi(l(n)), \tag{4}$$

$$v(k) = \sigma(k)(1 - \tau)\pi^*. \tag{5}$$

Workers consume their labor income net of taxes and their income from capital. Moreover, they experience disutility from providing labor effort. Capitalists do not work and consume their income from holding shares.

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. How firms exercise monopsony power over their workers is explained below. I make the important assumption that each firm observes the ability of the workers with whom it is matched. To a (potential) employee with ability n, a firm offers a bundle $(l(n), z(n))$ which consists of an effort (or hours) requirement $l(n)$ and labor earnings $z(n)$. Firms choose the bundles to maximize profits, subject to the requirement that they need to promise each worker at least her reservation utility, or outside option $\underline{v}(n)$. The latter is taken as given by firms and allowed to vary with ability. If a firm is matched to a worker with ability n, it solves

$$\max_{l(n),z(n)} \quad \pi(n) = nl(n) - z(n), \tag{6}$$
$$\text{s.t.} \quad z(n) - T(z(n)) + (1-\tau)\sigma(n)\pi^* - \phi(l(n)) \geq \underline{v}(n).$$

I assume the tax function $T(\cdot)$ is such that the first-order conditions are both necessary and sufficient. Because individuals' income from holding shares is proportional to the economy's aggregate profits, any impact of labor effort and earnings on π^* can be ignored. Combining the first-order conditions with respect to labor effort and earnings then gives

$$n = \frac{\phi'(l(n))}{1 - T'(z(n))}. \tag{7}$$

This condition is intuitive. In the optimum, firms offer bundles which equate a worker's productivity (on the left-hand side) to her willingness to substitute between labor effort and earnings (on the right-hand side). From equation (7) it is clear that without taxes on labor income, there is no distortion in labor supply and the equilibrium is efficient. The reason 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

Up to this point the model is similar to Mirrlees (1971), who in addition assumes workers can always choose their preferred number of hours at an hourly wage equal to their productivity.[5]

[5]The model from Mirrlees (1971) is obtained if the outside option of a worker with ability n is given by

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

Labor earnings then satisfy $z(n) = nl(n)$ and profits are driven to zero: $\pi(n) = 0$. In this case, the full labor market surplus is captured by workers. I generalize this framework by allowing for the possibility that firms have monopsony power, which determines what share of the labor market surplus is translated into profits. To keep the optimal tax problem studied in Section 3 tractable, I choose a specific way to operationalize monopsony power. In particular, monopsony power 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*

$$\pi(n) = \mu(n) \int_{n_0}^{n} l(m)dm. \tag{8}$$

Together with the definition $\pi(n) = nl(n) - z(n)$, equations (7) and (8) pin down profits, earnings and labor effort for a given profile of monopsony power $\mu(n)$. Clearly, profits are zero if labor markets are competitive (i.e., if $\mu(n) = 0$). In that case, the model coincides with Mirrlees (1971) and the full labor market surplus accrues to workers. Conversely, if labor markets are fully monopsonistic (i.e., if $\mu(n) = 1$), the full labor market surplus accrues to firms as workers are put on their participation constraint. To see this, suppose a worker's only outside option is non-employment. In that case, her reservation utility is given by $\underline{v}(n) = \sigma(n)(1 - \tau)\pi^* - T(0)$, where $\sigma(n)(1 - \tau)\pi^*$ denotes income from capital and $-T(0)$ is a non-employment benefit.[6] Substituting for $\underline{v}(n)$ in the profit maximization problem (6), the corresponding Lagrangian is given by

$$\mathcal{L}(n) = nl(n) - z(n) + \gamma(n)\Big[z(n) - T(z(n)) - \phi(l(n)) + T(0)\Big], \tag{9}$$

where $\gamma(n)$ is the Lagrange multiplier. Next, differentiate the objective (9) with respect to ability and apply the Envelope theorem to find $\mathcal{L}'(n) = \pi'(n) = l(n)$. Integrating this relationship gives an expression for profits if firms have full monopsony power:

$$\pi(n) = \pi(n_0) + \int_{n_0}^{n} l(m)dm. \tag{10}$$

Under the assumption that firms do not make profits from hiring the least productive workers (i.e., $\pi(n_0) = 0$), equation (10) coincides with the equation for profits (8) from Definition 1 if monopsony power $\mu(n) = 1$.[7] Hence, if firms have full monopsony power, workers are

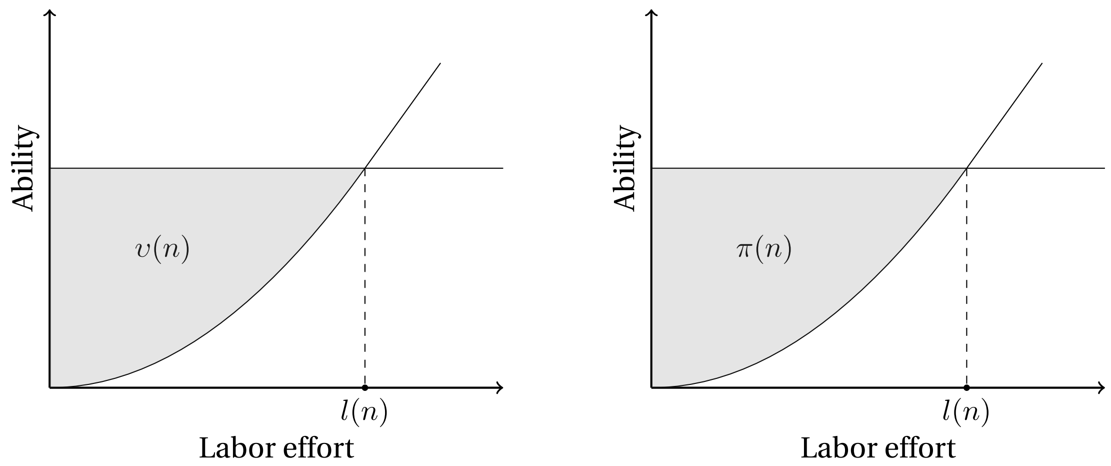
[6] The non-employment benefit $-T(0)$ is the negative of the tax burden a worker must pay when she earns zero labor income. As can be seen from equation (5), the non-employment benefit is not paid to capitalists, who also earn zero labor income. A possible micro-foundation is that the government can identify who is a worker and who is a capitalist or that the non-employment benefit is subject to an asset test.

[7] As is formally demonstrated in Appendix D, from an optimal tax perspective the assumption that firms do not earn profits from hiring the least productive worker (i.e., $\pi(n_0) = 0$) is without loss of generality.

put on their participation constraint and the full labor market surplus accrues to firms.

If taxes on labor income are linear, monopsony power $\mu(n) \in [0,1]$ captures what share of the labor market surplus accrues to firms if they 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 then be interpreted as the bargaining power of firms and workers, respectively. To make sure high-ability workers are not worse off, I assume more productive workers do not have lower bargaining power (i.e., do not suffer more from monopsony): $\mu'(n) \leq 0$.

Figure 1 graphically illustrates how monopsony power affects the payoffs of workers and firms. Here, I assume income taxes are absent and workers do not hold shares: $T(\cdot) = \sigma(n) = 0$. The horizontal line plots a worker'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 only 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 a worker's utility $v(n)$: 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

[8]Strictly speaking, the payoffs no longer coincide with those from the weighted Kalai-Smorodinsky solution if taxes on labor income are non-linear. In this case, labor effort generally depends on the degree of monopsony power as it is no longer pinned down solely by the first-order condition (7). As stated above, the reason for choosing to operationalize monopsony power in this specific way is to guarantee that the optimal tax problem remains tractable.

[9]This equilibrium would also occur 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.

2.4 Government

The government has a preference for redistribution. Its objective is

$$\mathcal{W} = \int_{n_0}^{n_1} \omega(n)\upsilon(n)f(n)dn + \kappa\omega(k)\upsilon(k), \tag{11}$$

where $\omega(n) \geq 0$ and $\omega(k) \geq 0$ are a set of Pareto weights for workers and capitalists. I assume $\omega(n)$ is non-increasing, which implies the government wishes to redistribute from high-ability to low-ability workers. In addition, I assume the Pareto weight for capitalists does not exceed the average Pareto weight for workers: $\omega(k) \leq \int_{n_0}^{n_1} \omega(n)f(n)dn$. The government thus wishes to redistribute on average from capitalists to workers, for example because capitalists hold significantly more shares.

As in Mirrlees (1971), I assume the government does not observe workers' 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 at an exogenous rate $\tau \in [0, 1]$. The government's budget constraint reads

$$\int_{n_0}^{n_1} \Big[T(z(n)) + \tau\pi(n)\Big] f(n)dn = G, \tag{12}$$

where $G \geq 0$ denotes some exogenous government spending. Because shareholdings are non-decreasing in ability and because capitalists hold more shares than workers on average, taxing pure economic profits is a very efficient way to redistribute capital income. Therefore, one can interpret the exogenous rate τ as the *maximum* share of 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 political constraints (e.g., due to firm lobbying), the government's inability to distinguish between normal and above-normal profits, or the existence of tax havens and profit-shifting opportunities.

2.5 Equilibrium

An equilibrium is formally defined as follows.

Definition 2. *An **equilibrium** consists of levels of labor effort $l(n)$, earnings $z(n)$ and profits $\pi(n) = nl(n) - z(n)$ 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, firms maximize profits: (7),

(ii) for all n, profits satisfy (8),

(iii) the government runs a balanced budget: (12).

Definition 2 describes the equilibrium for a given profile of monopsony power and a given set of tax instruments. Before turning to the optimal tax problem in Section 3, it is useful to highlight two implications of monopsony power. First, monopsony power increases the incidence of the tax burden 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). 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.[10]

Second, monopsony power decreases inequality in labor income generated by differences in workers' abilities but increases inequality in capital income generated by differences in shareholdings. Intuitively, monopsony power determines what share of the labor market surplus accrues to workers in the form of wages and what share accrues to firms in the form of profits. An increase in monopsony power therefore 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. How monopsony power affects both inequality in labor or capital income and the incidence of labor income taxes turns out to be crucial for understanding its implications for policy and welfare.

3 Optimal income taxation and welfare

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

[10] What share of the tax burden is borne by workers and firms is ultimately an empirical question. In a recent paper, Saez et al. (2019) find that a payroll tax cut in Sweden raised profits without affecting net-of-tax wages. This finding is consistent with a model of monopsony, where the incidence of taxes falls at least partly on firms. Rothstein (2010) also finds that a large share of the EITC-payments for low-income single mothers in the US is captured by employers through lower wages.

3.1 Optimal tax formulas

To solve the optimal tax problem, I follow the approach pioneered by Mirrlees (1971) and let the government choose the allocation that maximizes social welfare (11) subject to resource and incentive constraints. The optimal tax problem is formally defined in Appendix A. The next proposition states the first main result of this paper.

Proposition 1. *Consider the case where monopsony power and shareholdings do not vary with ability: $\mu(n) = \mu$ and $\sigma(n) = \sigma$ for all n. Then, at the optimal allocation,*

(i) the average welfare weight of workers equals one:

$$\int_{n_0}^{n_1} g(n)f(n)dn = 1, \tag{13}$$

(ii) for each n the marginal tax rate satisfies

$$T'(z(n)) = \mu(1-\tau)\sigma^*(k)(1-g(k))\left(\frac{1-F(n)}{nf(n)}\right) + (1-\mu)(1-T'(z(n)))\left(1+\frac{1}{\varepsilon(n)}\right)(1-\bar{g}(n))\left(\frac{1-F(n)}{nf(n)}\right), \tag{14}$$

which is generally positive and zero at the top: $T'(z(n_1)) = 0$.

Here, $g(k) = \omega(k)/\eta$ and $g(n) = \omega(n)/\eta$ denote the welfare weight of a capitalist and a worker with ability n, where η is the multiplier on the aggregate resource constraint. Moreover, $\bar{g}(n)$ is the average welfare weight for workers with ability at least equal to n and $\varepsilon(n) = \frac{\phi'(l(n))}{\phi''(l(n))l(n)}$ denotes the elasticity of labor supply.

Proof. See Appendix D. □

Equation (13) is a standard result in optimal tax theory. It states that when the tax system is optimized, the marginal cost of public funds equals one (Jacobs (2018)). Intuitively, the government should be indifferent between giving all workers an additional unit of net income (on the left-hand side) and receiving a unit of public resources (on the right-hand side). It can always guarantee this is the case by lowering the tax burden on labor income for all workers by the same amount and raising the non-employment benefit to make sure firms earn zero profits from hiring the least productive workers. The assumption on the Pareto weights $\omega(k) \leq \int_{n_0}^{n_1} \omega(n)f(n)dn$ then implies that the welfare weight of capitalists is below one: $g(k) \leq 1$. Hence, at the margin the government values a unit of net income for capitalists less than a unit of public resources.

Condition (ii) gives an expression for the optimal marginal tax rate at earnings level $z(n)$. According to equation (14), optimal marginal tax rates are generally positive, except at the top.[11] At each earnings level $z(n)$, the optimal marginal tax rate equals a weighted average between two components, where the weights reflect the degree of monopsony power. To understand this result, first consider the case where firms have full monopsony power: $\mu = 1$. The optimal marginal tax rate then satisfies

$$T'(z(n)) = (1-\tau)\sigma^*(k)(1-g(k))\left(\frac{1-F(n)}{nf(n)}\right). \tag{15}$$

If labor markets are fully monopsonistic, income taxes are used exclusively to redistribute capital income and not to redistribute labor income. This is because all workers are put on their identical participation constraint.[12] An increase in labor income taxes must then be compensated for by higher earnings as otherwise workers prefer non-employment. Hence, the full incidence of the tax burden is borne by firms. 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 $z(n)$. The mass of workers for whom this is the case equals $1 - F(n)$, which shows up in the numerator of equation (15). 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 welfare weight of capitalists is below the average welfare weight of workers: $\tau < 1$ and $g(k) < 1$. At the same time, however, the marginal tax rate generates distortions in labor effort: see equation (7). The associated efficiency costs are proportional to the density $f(n)$, which determines for how many workers labor effort is distorted and shows up in the denominator. Equation (15) thus states that if firms have full monopsony power, the optimal marginal tax rate balances the costs of distorting labor effort against the benefits of indirectly taxing profits.

The second component in the optimal tax formula (14) 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

$$\frac{T'(z(n))}{1-T'(z(n))} = \left(1+\frac{1}{\varepsilon(n)}\right)(1-\bar{g}(n))\left(\frac{1-F(n)}{nf(n)}\right). \tag{16}$$

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 from high-ability to low-ability workers. The term $1 - \bar{g}(n)$ summarizes the redistributive preferences of the government. It captures the average distributional gain of increasing the tax burden for all workers with ability at least equal to n. The total distributional benefits are

[11] 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 not only serves to redistribute labor income from high- to low-ability workers, but also to indirectly tax profits.

[12] This means all workers are equally well off. It does *not* mean all workers earn the same income. Firms find it optimal to demand more labor effort from more productive workers. To compensate them (i.e., to ensure the participation constraint holds), firms must promise higher labor earnings to these workers.

given by $(1-\bar{g}(n))(1-F(n))$. These benefits, in turn, are weighed against the distortionary costs. The latter are increasing in the labor supply elasticity $\varepsilon(n)$ and the density $f(n)$ of workers for whom the marginal tax rate distorts labor effort.

According to equation (14), the higher the degree of monopsony power, the more labor income taxes 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 tax rates is *a priori* ambiguous. This insight is formalized in the next Corollary.

Corollary 1. *Suppose the utility function is iso-elastic:* $u(c,l)=c-\frac{l^{1+1/\varepsilon}}{1+1/\varepsilon}$ *and hence,* $\varepsilon(n)=\varepsilon$. *The optimal marginal tax rate then satisfies the closed-form relationship*

$$T'(z(n))=\frac{\mu(1-\tau)\sigma^*(k)(1-g(k))+(1-\mu)(1+1/\varepsilon)(1-\bar{g}(n))}{a(n)+(1-\mu)(1+1/\varepsilon)(1-\bar{g}(n))}, \tag{17}$$

where $a(n)=nf(n)/(1-F(n))$ *is the local Pareto parameter of the ability distribution. Moreover, if* $a(n)(1-\tau)\sigma^*(k)(1-g(k))(1-\bar{g}(n))>0$, *an increase in monopsony power raises the optimal marginal tax rate at earnings level* $z(n)$ *if and only if the following condition holds:*

$$((1-\tau)\sigma^*(k)(1-g(k)))^{-1}<((1+1/\varepsilon)(1-\bar{g}(n)))^{-1}+a(n)^{-1}. \tag{18}$$

Proof. Substitute $\varepsilon(n)=\varepsilon$ in equation (14) and use the definition of $a(n)$. Rearranging gives equation (17). To determine how monopsony power affects the optimal tax rate, differentiate equation (17) with respect to μ:

$$\frac{\partial T'(z(n))}{\partial \mu}= \tag{19}$$
$$\frac{a(n)((1-\tau)\sigma^*(k)(1-g(k))-\left(1+\frac{1}{\varepsilon}\right)(1-\bar{g}(n)))+(1-\tau)\sigma^*(k)(1-g(k))\left(1+\frac{1}{\varepsilon}\right)(1-\bar{g}(n))}{(a(n)+(1-\mu)\left(1+\frac{1}{\varepsilon}\right)(1-\bar{g}(n)))^2}.$$

The latter is positive if and only if the numerator is positive. To simplify the exposition, divide the numerator by $a(n)(1-\tau)\sigma^*(k)(1-g(k))(1+1/\varepsilon)(1-\bar{g}(n))$ and set the resulting expression strictly larger than zero. Rearranging leads to the inequality (18). □

Equation (17) gives a closed-form expression for the optimal marginal tax rate. It follows from rearranging equation (14) and plays an important role in the numerical analysis in Section 4. Equation (18), in turn, gives a precise condition under which an increase in monopsony power raises the optimal marginal tax rate at each point in the income distribution. The impact of monopsony power on optimal tax rates is ambiguous. On the one hand, income taxes can be used to indirectly tax profits if firms have monopsony power. This raises the optimal tax rate. On the other hand, income taxes become less effective in redistributing labor income because workers do not bear the full incidence of income taxes if firms have monopsony power. This lowers the optimal tax rate. According to equation (18), the first of these effects is more likely to dominate if profit taxation is severely restricted (i.e., τ low) and if profits flow back to individuals with low welfare weights (i.e., $g(k)$ low). Conversely, the second effect is more likely to dominate if redistributing from high-ability to low-ability workers is very important (i.e., $\bar{g}(n)$ low) or if the local Pareto parameter $a(n)$ is high.[13]

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 workers equals one (i.e., $\bar{g}(n_0) = 1$: see equation (13)), condition (18) is always satisfied at the bottom of the income distribution. Hence, monopsony power unambiguously raises optimal marginal tax rates at low levels of income. Intuitively, marginal tax rates at the bottom do not help to redistribute labor income from high-ability to low-ability workers but are helpful to indirectly tax profits. The latter becomes more important if monopsony power increases. At higher levels of income, redistributing labor income from individuals above that level to individuals below that level becomes 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 is therefore more likely to reduce optimal tax rates at higher income levels. I explore the quantitative implications of monopsonony power for the pattern of optimal income taxes in Section 4.

3.2 Welfare effects of monopsony power

I now turn to analyze how a change in monopsony power affects welfare. To that end, consider the following result.

Proposition 2. *Suppose monopsony power and shareholdings do not vary with ability and suppose income taxes $T(\cdot)$ are optimally set. Consider a proportional increase in μ, so that monopsony power raises to $\mu(1+\alpha)$. The associated impact on welfare is determined by*

$$\frac{\partial \mathcal{W}(\alpha)}{\partial \alpha} = \eta \int_{n_0}^{n_1} \mu \Big[(1 - T'(z(n)))(1 - \bar{g}(n)) - (1-\tau)\sigma^*(k)(1 - g(k)) \Big] (1 - F(n)) l(n) dn, \quad (20)$$

which can either be positive, zero or negative.

Proof. See Appendix E. □

[13]The reason why the local Pareto parameter $a(n)$ shows up in equation (18) is quite mechanical. In the second component of equation (14), 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 the marginal tax rate. This is the case if the local Pareto parameter is low. Therefore, a lower Pareto parameter makes it easier for condition (18) to be satisfied.

According to equation (20), monopsony power has an ambiguous effect on welfare. On the one hand, monopsony power enables the government to use firms as a screening device – as captured by the first term. This is because firms observe workers' abilities, while the government does not. As a result, the government cannot redistribute income through lump-sum taxes on ability. High-ability individuals, in turn, benefit more from the government's inability to observe workers' abilities. In mechanism-design terminology, they earn information rents. Part of these information rents are captured by firms if they have monopsony power. Monopsony power then decreases inequality in labor income generated by differences in workers' abilities. This lowers the need for using distortionary taxes on labor income to redistribute from high-ability to low-ability workers, which *ceteris paribus* is welfare-enhancing. On the other hand, monopsony power has a negative welfare effect as it generates a distributional conflict over profits – as captured by the second term. Monopsony power raises aggregate profits and thereby inequality in capital income. The increase in profits lowers welfare if profits cannot be taxed at a confiscatory rate and after-tax profits flow back to individuals with below-average welfare weights.

To gain further intuition why monopsony power might be welfare-enhancing, suppose profit taxation is unrestricted: $\tau = 1$. According to equation (20), welfare is then unambiguously higher with monopsonistic than with competitive labor markets. Intuitively, if labor markets are competitive, profits are zero and the government does not benefit from the fact that firms observe ability. However, profits are positive if firms have monopsony power. The combination of monopsony power and profit taxes allows the government to exploit the informational advantage of firms and extract information on hidden ability. This raises welfare as it alleviates the equity-efficiency trade-off that results from the assumption that the government does not observe ability (cf. Mirrlees (1971)).

The trade-off between equity and efficiency disappears altogether if profit taxation is unrestricted and firms have full monopsony power: $\tau = \mu = 1$. In this case, firms extract the full labor market surplus and all workers are put on their identical participation constraint: see Figure 1b. Hence, all workers are equally well off. Moreover, the government can use profit taxes to redistribute income from capitalists to workers. It is then optimal for the government to levy a confiscatory tax on profits and make the workers' outside option of not working as attractive as possible. It can do so by providing a basic income $-T(0)$.[14] Importantly, this basic income should not be taxed away if individuals earn higher income. Put differently, optimal marginal tax rates are zero: see equation (14) and set $\tau = \mu = 1$. The reason is that taxing labor income only generates distortions and does not lead to more redistribution.

Returning to the case with an intermediate degree of monopsony power and a possible restriction on profit taxation (i.e., the case from Proposition 2), a few remarks are in place.

[14] The basic income can also be paid to capitalists if profits are taxed at a confiscatory rate.

First, equation (20) gives the welfare effects of raising monopsony power – provided income taxes are optimally set. Consequently, the result can only be used to assess the welfare effects of raising monopsony power at the *current* tax system under the additional assumption that the latter properly reflects the government's redistributive preferences. Second, equation (20) is not written in terms of the model's primitives as it depends on the entire profile of labor effort and marginal tax rates, which are both endogenous. However, if one assumes the labor supply elasticity does not depend on labor effort (and hence, $\varepsilon(n) = \varepsilon$ for all n), it is possible to express the result in terms of exogenous variables: see Appendix E. Third, Proposition 2 considers a *proportional* increase in monopsony power rather than an increase by a constant amount. The reason for doing so is that the result can be generalized to the case where monopsony power varies with ability, which is studied next.[15]

3.3 Ability-specific monopsony power and shareholdings

Up to this point, the results have been derived under the assumption that monopsony power and shareholdings do not vary with ability. Both assumptions can be relaxed. To guarantee that individuals with higher ability are not worse off, I assume more productive workers do not have a lower bargaining power (i.e., do not suffer more from monopsony) and do not own fewer shares: $\mu'(n) \leq 0$ and $\sigma'(n) \geq 0$. The next Proposition summarizes the implications for optimal income taxation and the welfare effect of raising monopsony power.

Proposition 3. *Suppose monopsony power might decrease and shareholdings might increase*

[15]A slight disadvantage is that – by definition – a proportional increase in monopsony power has no welfare effect if $\mu = 0$ (i.e., starting from a setting where labor markets are competitive). The result from Proposition (2) remains valid in a slightly modified form if monopsony power increases by a constant amount from μ to $\mu + \delta$. Following the same steps as in Appendix E, the welfare effect is then given by

$$\frac{\partial \mathcal{W}(\delta)}{\partial \delta} = \eta \int_{n_0}^{n_1} \left[(1 - T'(z(n)))(1 - \bar{g}(n)) - (1 - \tau)\sigma^*(k)(1 - g(k)) \right] (1 - F(n)) l(n) dn,$$

which is identical to equation (20), except that μ does not appear on the right-hand side.

with ability: $\mu'(n) \leq 0$ *and* $\sigma'(n) \geq 0$. *Then, the optimal marginal tax rate satisfies*

$$
\begin{aligned}
T'(z(n)) = \bar{\mu}(n)(1-\tau)\sigma^*(k)(1-g(k)) \left(\frac{1-F(n)}{nf(n)}\right) \qquad (21) \\
+ (1-\mu(n))(1-T'(z(n))) \left(1+\frac{1}{\varepsilon(n)}\right)(1-\bar{g}(n)) \left(\frac{1-F(n)}{nf(n)}\right) \\
- \left(\frac{1-F(n)}{nf(n)}\right) \Bigg[\left(\frac{\mu'(n)}{\mu(n)}\right)\left(\frac{\pi(n)}{\varepsilon(n)l(n)}\right)(1-T'(z(n)))(1-\bar{g}(n)) \\
+ \int_n^{n_1} \mu'(m)(1-T'(z(m)))(1-\bar{g}(m)) \left(\frac{1-F(m)}{1-F(n)}\right) dm \\
- \bar{\mu}(n)(1-\tau)\int_{n_0}^{n_1} \sigma'(m)(1-\bar{g}(m))(1-F(m))dm \Bigg],
\end{aligned}
$$

which is generally positive and zero at the top: $T'(z(n_1)) = 0$. *Here,* $\bar{\mu}(n)$ *denotes the average monopsony power for workers with ability at least equal to* n. *Moreover, the welfare effect of a proportional increase in monopsony power is given by*

$$
\begin{aligned}
\frac{\partial \mathcal{W}(\alpha)}{\partial \alpha} = \eta \int_{n_0}^{n_1} \Bigg[\mu(n)(1-T'(z(n)))(1-\bar{g}(n)) - \bar{\mu}(n)(1-\tau)\sigma^*(k)(1-g(k)) \qquad (22) \\
+ \int_n^{n_1} \mu'(m)(1-T'(z(m)))(1-\bar{g}(m)) \left(\frac{1-F(m)}{1-F(n)}\right) dm \\
- \bar{\mu}(n)(1-\tau)\int_{n_0}^{n_1} \sigma'(m)(1-\bar{g}(m))(1-F(m))dm \Bigg](1-F(n))l(n)dn.
\end{aligned}
$$

Proof. See Appendix D and E. □

Equations (21) and (22) generalize the findings from Propositions 1 and 2, respectively. Compared to the case where monopsony power and shareholdings are constant, inequality in labor and capital income generated by differences in ability is higher if more productive workers suffer less from monopsony and hold more shares. This explains why *ceteris paribus* optimal marginal tax rates are higher. In the optimal tax formula (21), three additional effects show up. First, a reduction in monopsony power *at a particular ability level* implies more productive workers are better off as they suffer less from monopsony. Second, a reduction in monopsony power *at higher ability levels* lowers the profits firms generate from hiring more productive workers. Hence, high-ability workers manage to capture a larger share of the labor market surplus. Third, if shareholdings are increasing in ability a larger share of aggregate profits flows back to more productive workers. All these effects raise the distributional benefits of income taxes and hence, raise the optimal marginal tax rate.

The welfare effect of raising monopsony power is lower compared to the case where shareholdings and monopsony power do not vary with ability. The reason is that inequality generated by differences in ability is exacerbated if more productive workers suffer less from monopsony and hold more shores. Compared to the result from Proposition 2, two additional effects show up in equation (22), which are both negative. First, as stated before, more productive workers capture a larger share of the labor market surplus if they suffer less from monopsony. Second, a larger share of aggregate profits flows back to more productive workers if they own more shares. Both effects reduce the positive welfare effect of raising monopsony power (which consists of reducing inequality generated by differences in ability: see Proposition 2) compared to the case where shareholdings and monopsony power are constant. Monopsony power is therefore more likely to be welfare-reducing.

4 Numerical illustration

This Section quantitatively explores the implications of monopsony power in the baseline version of the model where monopsony power and shareholdings do not vary with ability. After presenting the calibration (Section 4.1) and different specifications for 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 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 (Saez and Stantcheva (2018)).[16]

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

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:

$$u(c,l) = c - \frac{l^{1+1/\varepsilon}}{1+1/\varepsilon}, \tag{23}$$

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:

$$T(z(n)) = -g + tz(n). \tag{24}$$

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 also Saez (2001). This gives $g = \$4,590$ and $t = 33.1\%$ with an R^2 of approximately 0.94. Figure 2 plots the actual and fitted values for incomes up to $500,000.

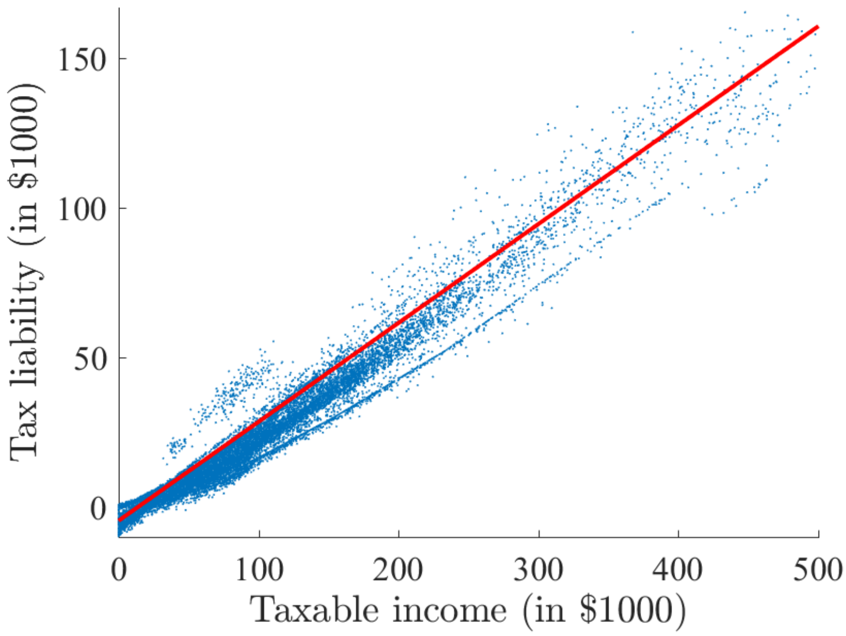


Figure 2: Current tax schedule

4.1.3 Equilibrium

With the above specification for the utility function and the tax schedule, it is straightforward to derive the equilibrium (cf. Definition 2). Labor effort follows from equation (7):

$$l(n) = (1-t)^{\varepsilon} n^{\varepsilon}. \tag{25}$$

Labor earnings, in turn, are obtained by substituting labor effort in equation (8) and using the definition $\pi(n) = nl(n) - z(n)$. This gives

$$z(n) = \left(1 - \frac{\mu}{1+\varepsilon}\right)(1-t)^{\varepsilon} n^{1+\varepsilon} + \left(\frac{\mu}{1+\varepsilon}\right) z(n_0). \tag{26}$$

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

$$\pi(n) = \left(\frac{\mu}{1+\varepsilon-\mu}\right)(z(n) - z(n_0)). \tag{27}$$

Equations (26) and (27) give a mapping from (observable) labor income to (unobservable) ability and pure profits, respectively. To calculate ability and profits, the only component which still needs to be determined is the degree of monopsony power μ.

4.1.4 Monopsony power

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*. In recent work, Barkai (2016) and Barkai and Benzell (2018) 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. Their measure of profits thus corresponds to pure economic (or above-normal) profits, as is the case in my model. For the most recent year 2015, Barkai and Benzell (2018) calculate that the ratio of aggregate profits to aggregate wages (i.e., the ratio of the profit share to the labor share) is approximately 24.2%. Using their estimate, the value for monopsony power μ can be calculated by integrating equation (27) over the ability distribution and dividing by aggregate labor income $z^* = \int_{n_0}^{n_1} z(n)f(n)dn$. This gives

$$\left(\frac{\pi^*}{z^*}\right) = \left(\frac{\mu}{1+\varepsilon-\mu}\right)\left(1 - \left(\frac{z(n_0)}{z^*}\right)\right) \quad \Leftrightarrow \quad \mu = (1+\varepsilon)\left[\frac{(\pi^*/z^*)}{1 + (\pi^*/z^*) - (z(n_0)/z^*)}\right]. \tag{28}$$

Substituting out for the elasticity of labor supply and the ratio of profits to wages gives a value for monopsony power of approximately $\mu = 0.26$.[17]

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

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 (26) 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 left panel of Figure 3. To facilitate the interpretation, the right panel plots the distribution of labor earnings and the implied kernel density.

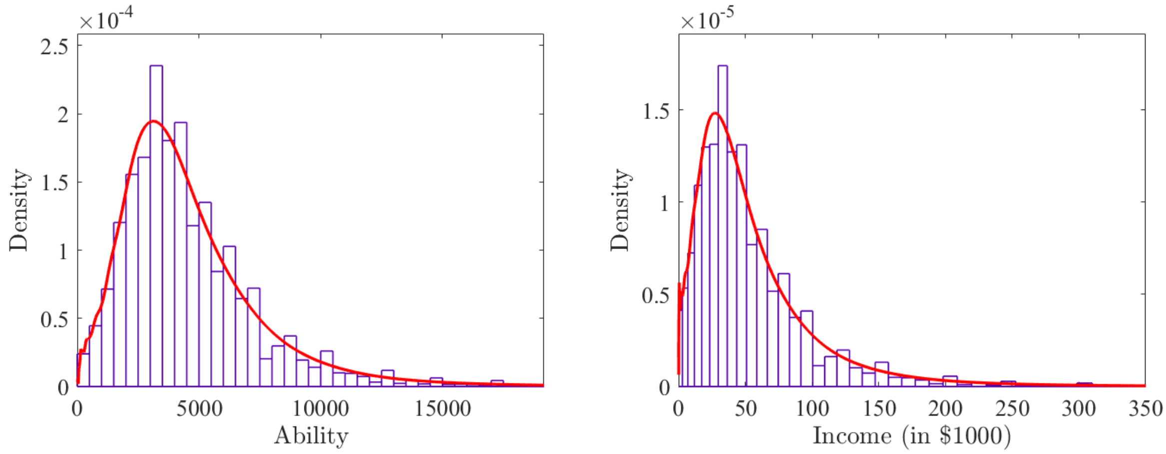


Figure 3: Distribution of ability and income

I make two adjustments to the density as plotted in the left panel of Figure 3. First, 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.[18] 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. Second, I exclude ability levels at the bottom where a very low value of the local Pareto parameter $a(n)$ implies that optimal tax rates according to equation (17) exceed 100%.[19] The latter is never feasible as it violates the first-order condition for profit maximization (7). In the baseline calibration, this means that individuals with earnings below $11,350 are excluded.

[18] If we denote by $H(z(n))$ the labor income distribution and by $h(z(n))$ the corresponding density, monotonicity of labor earnings (see Appendix B) implies $F(n) = H(z(n))$ for all n and hence, $f(n) = h(z(n))z'(n)$. The local Pareto parameter of the ability distribution $a(n) = nf(n)/(1 - F(n))$ and income distribution $\hat{a}(z(n)) = z(n)h(z(n))/(1 - H(z(n)))$ are then related through $a(n) = \hat{a}(z(n))e_{zn}$, 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 (26).

[19] This happens if $a(n) < \mu(1 - \tau)\sigma^*(k)(1 - g(k))$: see equation (17). To make sure this does not occur for any welfare weight of capitalists $g(k)$, I exclude ability levels where $a(n) < \mu(1 - \tau)\sigma^*(k)$.

4.1.6 Capitalists and profit taxation

The final parameters to be calibrated are those related to capitalists and the taxation of profits. In the model, capitalists are individuals who derive their income exclusively from holding shares. Saez and Stantcheva (2018) show that capital (as opposed to labor) is the primary source of income only for individuals at the very top of the *total* income distribution. I therefore set κ to be consistent with a fraction of capitalists equal to 0.01. Saez and Stantcheva (2018) also report that the top 1% in the capital income distribution (which overlaps significantly with the top 1% in the total income distribution) earns approximately 63% of total capital income. This suggests a value of $\sigma^*(k) = 0.63$.

Regarding the taxation of profits, in the model τ is the *maximum* share of above-normal profits that can be taxed. A natural upper bound is therefore $\tau = 1$, but in practice a confiscatory tax on above-normal profits is unlikely to be feasible. In the baseline calibration, I therefore set τ consistent with the current corporate tax rate of 21% (levied at the firm level) and a capital gains tax rate of 20% (levied at the individual level). This implies capital income cannot be taxed at a rate higher than $\tau = 36.8\%$. For a given value of τ, the government's budget constraint can then be used to calculate the revenue requirement. This gives $G = \$21,740$, which in the calibrated economy corresponds to approximately 26.5% of aggregate output. The calibration strategy is summarized in Table 1.

Variable	Target	Source	Value
μ	Aggregate profits over wages (0.24)	Barkai and Benzell (2018)	0.26
ε	Elasticity of labor supply (0.33)	Chetty (2012)	0.33
κ	Fraction of capitalists (0.01)	Saez and Stantcheva (2018)	0.01
$\sigma^*(k)$	Share capital income by top 1% (0.63)	Saez and Stantcheva (2018)	0.63
τ	Corporate tax rate (0.21) Capital gains tax rate (0.20)	OECD IRS	0.37
G	Government budget constraint	Equilibrium condition	\$21,740
$T(z)$	Tax liability	CPS 2018	Figure 2
$F(n)$	Income distribution	CPS 2018	Figure 3

Table 1: Calibration

4.2 Welfare weights

How taxes should be set and how monopsony power affects optimal income taxation and welfare critically depends on the welfare weight of capitalists $g(k)$ and how the government values redistributing income from high-ability to low-ability workers (as summarized by $\bar{g}(n)$). Below I consider a number of different specifications for the welfare function, which are summarized as follows.[20]

Definition 3. *The following specifications for the welfare function are considered.*

(i) ***Case B****. Monopsony power is bad for welfare:* $\bar{g}(n) = 1$ *for all* n *and* $g(k) = 0$.

(ii) ***Case G****. Monopsony power is good for welfare:* $\bar{g}(n) = 0$ *almost everywhere and* $g(k) = 1$.

(iii) ***Case M****. Welfare weights satisfy* $g(n) \propto n^{-1}$ *and* $g(k) = g(n')$ *for* n' *such that* $\upsilon(n') = \upsilon(k)$.

Case B considers the scenario where monopsony power is most detrimental to welfare. The government attaches a zero welfare weight to capitalists and a welfare weight of one to all workers. An increase in profits at the expense of wages then unambiguously lowers welfare. Moreover, any reduction in labor income inequality due to monopsony power is not valued by the government since all workers have the same welfare weight. Case G considers the opposite scenario where monopsony power is unambiguously welfare-enhancing. Capitalists have a welfare weight of one, which is the same as workers on average: see equation (13).[21] In this case, an increase in profits is not harmful to welfare. Moreover, among the group of workers the government has essentially Rawlsian preferences: it only cares about the least well-off. Monopsony power then raises welfare as it reduces inequality generated by differences in ability. Finally, case M considers a scenario where the government has moderate preferences for redistribution. Welfare weights are inversely proportional to ability and workers and capitalists who are equally well off get the same welfare weight. [22]

4.3 Optimal marginal tax rates

4.3.1 Optimal marginal tax rates in the calibrated economy

Figure 4 plots the optimal marginal tax rates for different specifications of the welfare function. To facilitate the comparison, in all cases optimal marginal tax rates are plotted against *current* labor earnings. In the first scenario (case B), income taxes only serve to redistribute

[20] In each case, I also assume there is a mass of $\nu = 0.05$ individuals (non-participants) who earn zero labor *and* capital income. They have a higher welfare weight than workers on average and are entitled to a benefit, but their utility is bounded from above by that of the least productive workers. Under these assumptions, the optimal marginal tax rate is positive at the bottom – even if labor markets are competitive. Assuming there is a mass of non-participants avoids technical difficulties associated with steeply increasing marginal tax rates at low levels of earnings. These difficulties arise because without a mass point at the bottom of the income distribution, the least-skilled workers should face a marginal tax rate of zero if labor markets are competitive: see Seade (1977). This is only a very local result, as the low value of the local Pareto parameter $a(n)$ at the bottom implies the optimal marginal tax rate immediately jumps to a high value. Such a jump often leads to a violation of the monotonicity condition: see Appendix B.

[21] Strictly speaking, capitalists have a welfare weight which equals the average welfare weight of *all* other individuals (i.e., workers *and* non-participants). Formally, the optimality condition $\bar{g}(n_0) = 1$ from equation (13) is modified to $(\frac{1}{1+\nu})\bar{g}(n_0) + (\frac{\nu}{1+\nu})g(\nu) = 1$, where $g(\nu) \geq \bar{g}(n_0)$ is the welfare weight of non-participants.

[22] With these preferences, the welfare weight of capitalists equals $g(k) = 0.06$. In addition, the government is indifferent between giving \$2 to a worker with ability n and \$1 to another worker whose ability equals twice that amount. In terms income, at the current tax system a doubling in ability corresponds to an increase in labor earnings by a factor roughly $2^{1+\varepsilon} \approx 2.51$: see equation (26). Finally, the welfare weight of non-participants is equated to that of the least productive workers, who are equally well off. This gives $g(\nu) = 2.70$.

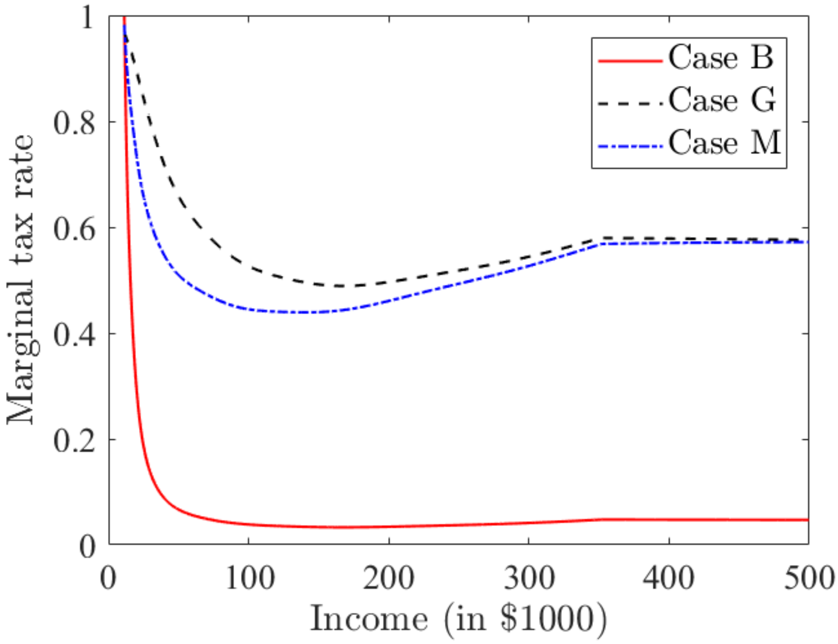


Figure 4: Optimal marginal tax rates

capital income by indirectly taxing profits. To see this, substitute $\bar{g}(n) = 1$ and $g(k) = 0$ in equation (17) to find

$$T'(z(n)) = \frac{\mu(1-\tau)\sigma^*(k)}{a(n)}, \tag{29}$$

which equals zero if there is no restriction on profit taxation (i.e., if $\tau = 1$). As can be seen from the figure, optimal tax rates are generally very low except at low earnings levels (where the local Pareto parameter $a(n)$ is low: see Figure 6 in Appendix F). Intuitively, a high marginal tax rate at the bottom distorts the labor effort of only a few workers, while raising the tax burden for all workers. The latter is beneficial because the incidence of the tax burden falls partly on firms if they have monopsony power (i.e., if $\mu > 0$). Afterwards optimal tax rates decrease rapidly in earnings, to values below 5% at income levels above the mean (approximately $67,000 in the calibrated economy). Hence, using marginal tax rates on labor income to indirectly tax profits appears most relevant at low levels of labor earnings.

Optimal tax rates are much higher if income taxes do not serve to redistribute from capitalists to workers, but only to redistribute labor income from high-ability to low-ability workers (case G). With these preferences, optimal tax rates satisfy

$$T'(z(n)) = \frac{1-\mu}{a(n)\varepsilon/(1+\varepsilon)+1-\mu}, \tag{30}$$

which is obtained by substituting $\bar{g}(n) = 0$ and $g(k) = 1$ in equation (17). 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 F. Optimal tax rates start out very high, then decrease monotonically to approximately 49% (at around \$165,000 in current earnings) after which they increase to a constant top rate of around 58% at the point where the Pareto tail is pasted (i.e., \$350,000 in current earnings).

Optimal tax rates lie in between those obtained in the previous cases if the government has more moderate preferences for redistribution (case M). In this case, the government cares both about redistributing from capitalists to workers and about redistributing from high-ability to low-ability workers. Compared to the case where the government has Rawlsian preferences among workers (case G), optimal tax rates are considerably lower especially at low earnings. This is because the government not only cares about workers at the very bottom, but also attaches a sizable weight to low- and middle-class workers. For example, with these preferences the government attaches a weight of approximately $\bar{g}(n) = 0.71$ to individuals with earnings at least equal to \$27,000, compared to a weight of zero for a Rawlsian government. Only at high income levels are the differences between cases G and M very small. This is because in both cases, the government attaches a low welfare weight both capitalists and workers with high incomes.

The above illustrates that how tax policy should be designed is, unsurprisingly, sensitive to the choice of the welfare function. The more substantive insight is that only marginal tax rates at low levels of income appear effective as a way to indirectly tax profits – see case B. This insight is confirmed if we compare the optimal tax system with monopsony power to the optimal tax system with competitive labor markets, which is done next.

4.3.2 Comparison to the competitive benchmark

I now turn to compare the optimal tax rates in the calibrated economy with those that would be obtained if labor markets are competitive, as in Saez (2001). To that end, I set monopsony power $\mu = 0$ and calculate optimal tax rates using equation (17) for different specifications of the welfare function, cf. Definition 3. For simplicity, I do not recalibrate the ability distribution to match the empirical income distribution assuming labor markets are competitive. As it turns out, doing so has a negligible effect on the outcomes.[23]

Figure 5 plots the optimal tax rates in the benchmark economy and the ones under perfect competition. Here, I assume the government values both redistributing from high- to low-ability workers and from capitalists to workers (case M). For individuals with earnings currently below \$25,000, optimal marginal tax rates are higher if firms have monopsony power. This confirms the finding from Corollary 1 that monopsony power tends to raise optimal tax rates at low levels of income. It is also in line with the result plotted in Figure 4, where it was

[23] This can be seen from Figure 6 in Appendix F, which shows that the local Pareto parameter in the calibrated economy is almost identical to the one that is obtained if the ability distribution is calibrated assuming $\mu = 0$.

shown that optimal tax rates are high only at the bottom if the sole purpose of income taxes is to indirectly tax profits (case B). For income levels above \$25,000, optimal tax rates with monopsony power are lower. The difference is fairly stable at earnings above the mean and varies between 5 and 6 percentage points. The reason why optimal tax rates are lower for middle- and high-income earners is that monopsony power makes income taxes less effective in redistributing labor income, as part of the incidence falls on firms.

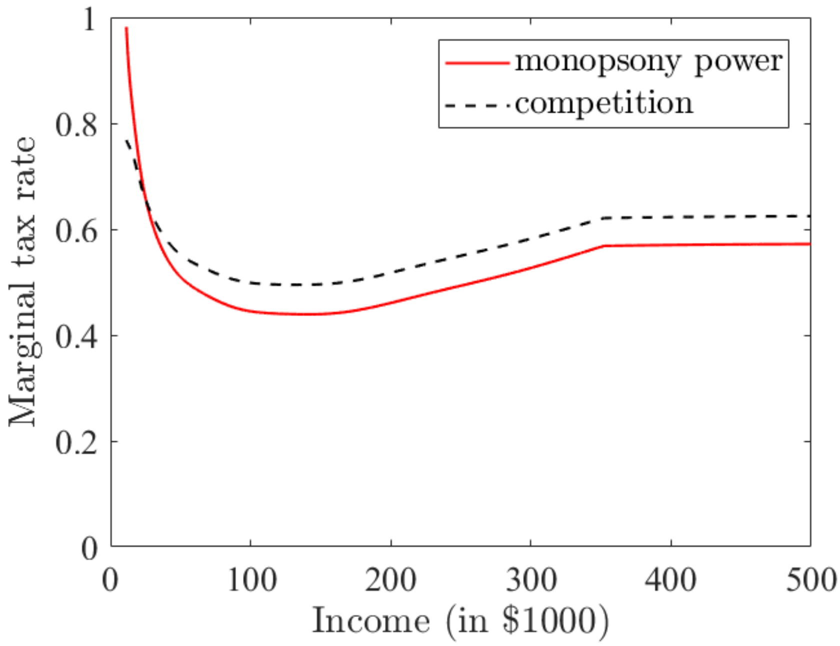


Figure 5: Comparison optimal marginal tax rates (case M)

Figures 7 and 8 in Appendix F show the comparison between optimal tax rates in the calibrated economy with those under perfect competition for the other specifications of the welfare function. For a government that only wishes to redistribute away from capitalists (case B), monopsony power unambiguously raises optimal tax rates. In fact, optimal tax rates are zero if labor markets are competitive: see equation (29) and set $\mu = 0$. Intuitively, income taxes cannot be used to indirectly tax profits if labor markets are competitive and the government does not value redistribution from high-ability to low-ability workers. Conversely, for a government that only wishes to redistribute from high-ability to low-ability workers (case G), monopsony power unambiguously lowers optimal tax rates. As explained above, this is because monopsony power makes income taxes less effective in redistributing labor income as part of the incidence falls on firms. Consequently, at each income level the optimal tax rate is lower than would be the case if labor markets are competitive.

Clearly, how monopsony power affects optimal tax rates depends on the specification of the welfare function, see also Corollary 1. The main insight here is that for a government that wishes to redistribute from high-ability to low-ability workers *and* from capitalists to workers, monopsony power tends to raise optimal tax rates at low earnings levels and lowers optimal tax rates for middle- and high-income earners (cf. Figure 5).

4.4 Implications for welfare

To assess the quantitative implications of monopsony power for welfare, 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 under the optimal tax system (see Proposition 1) with the one that is obtained if a "naive" planner wrongfully sets tax policy *as if* labor markets are competitive (i.e., assuming $\mu = 0$). Second, I calculate how much the government is willing to pay/receive for completely eliminating monopsony power. To that end, I compare the optimal allocation in the calibrated economy with monopsony power and the optimal allocation under perfect competition. The first exercise gives an indication how important it is to take 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, for example through competition policy.

	Case B	Case G	Case M
Misoptimization	\$168 (0.2)	\$343 (0.4)	\$133 (0.2)
No monopsony power	\$5,716 (6.8)	–\$2,852 (–3.4)	\$1,588 (1.9)

Table 2: Consumption equivalents (% of GDP)

Table 2 shows the results for different specifications of the welfare function. To facilitate the interpretation, in each case I express the welfare effects in consumption equivalents that are given to *all* individuals. In addition to the dollar amounts, the consumption equivalents are also expressed as a fraction of output (i.e., earnings plus profits) in the calibrated economy at the current tax system. The welfare costs of misoptimization (i.e., setting tax policy as if labor markets are competitive) range between \$133 and \$343 per capita, or between 0.2% and 0.4% of current output. Hence, the government is indifferent between correcting the sub-optimal tax schedule and increasing all individuals' net income by this amount. The welfare costs of ignoring monopsony power are largest if the government only wishes to redistribute from high-ability to low-ability workers (case G). In this case, the difference in optimal marginal tax rates with and without monopsony power is also sizable: see Figure 8 in Appendix F. For income levels above the mean, optimal tax rates are between 7 and 8 percentage points lower if monopsony power is taken into account.

The second row in Table 2 shows the welfare effect of changing the degree of monopsony power from its value in the calibrated economy to zero, i.e., to a setting where labor markets are competitive. The welfare effect can be positive or negative depending on the redistributive preferences of the government. If the government only wishes to redistribute from capitalists to workers (case B), monopsony power is very costly as it raises profits. With these preferences, the government is willing to give up \$5,716 per capita (or 6.8% of output) to go from a situation with the current degree of monopsony power to a situation where labor markets are competitive. Conversely, if the government only wishes to redistribute from high-ability to low-ability workers (case G), monopsony power is welfare-enhancing as it reduces inequality generated by differences in workers' abilities. With these preferences, the government would be willing to pay \$2,852 per capita (or 3.4% of output) to *prevent* firms losing their monopsony power. If the government wishes to redistribute both from capitalists to workers and from high-ability to low-ability workers, the welfare effect of monopsony power is theoretically ambiguous: see Proposition 2. In the current calibration, the effect is negative. A government with moderate preferences for redistribution (case M) would be willing to pay \$1,588 per capita (1.9% of output) to get rid of monopsony power.

The above exercise suggests that correcting a sub-optimal tax code by taking monopsony power into account leads to modest welfare gains, whereas eliminating monopsony power leads to sizable welfare effects. How large these welfare effects are depends critically on the redistributive preferences of the government, in particular how much it cares about redistributing income from capitalists to workers and from high-ability to low-ability workers.

5 Conclusion

This paper extends the non-linear tax framework from Mirrlees (1971) with monopsony power and studies the implications for optimal income taxation and welfare. In my model, firms observe workers' abilities while the government does not. Monopsony power does not generate efficiency losses, but determines what share of the labor market surplus is translated into pure economic profits. These profits flow back as capital income to workers who differ in their ability and to capitalists, who do not work and derive their income exclusively from holding shares. The government wishes to redistribute from capitalists to workers and from high-ability to low-ability workers. In addition to an exogenous tax on profits, the government optimizes a non-linear tax on labor income.

I derive an intuitive expression for the optimal marginal tax rate on labor income and show that income taxes are not only used to redistribute labor income, but also to redistribute capital income. Intuitively, part of the incidence of income taxes falls on firms if they have monopsony power. As a result, monopsony power makes income taxes less effective in reducing inequality in after-tax labor income, but more effective in reducing inequality in capital income by lowering aggregate profits. The impact of monopsony power on optimal tax rates is generally ambiguous. I calibrate the model to the US economy and show that monopsony power tends to raise optimal tax rates at low earnings levels and to lower optimal tax rates for middle- and high-income earners.

Monopsony power has an ambiguous effect on welfare. On the one hand, it enables the government to use firms as a screening device in order to exploit their informational advantage. Intuitively, firms capture part of the labor market surplus if they have monopsony power. As a result, monopsony power decreases inequality in labor income generated by differences in ability. This raises welfare as it reduces the need for using distortionary taxation to redistribute from high-ability to low-ability workers. On the other hand, monopsony power generates a distributional conflict over profits. This lowers welfare provided profits flow back to individuals with below-average welfare weights. Which of these forces dominates critically depends on how much the government values redistribution from capitalists to workers and from high-ability to low-ability workers. In the calibrated economy, the welfare effect of eliminating monopsony power is sizable and ranges between –3.4% and +6.8% of GDP depending on the government's redistributive preferences.

The most critical assumption in my analysis is that firms observe workers' abilities, while the government does not. If neither firms nor the government observes ability, monopsony power has vastly different implications for optimal income taxation and welfare.[24] While the assumption made in this paper that firms perfectly observe ability and the government does not is clearly a strong one, it seems natural to think that firms know more about their workers' abilities than the government. Firms spend significant resources to assess potential applicants and conduct performance evaluations once workers are hired. Whenever firms have monopsony power and an informational advantage compared to the government, there is scope for the government to use firms as a screening device. I therefore expect my results to be applicable more generally. Moreover, my analysis also goes through if the government observes ability but is restricted in conditioning taxes only on labor income (and not on ability directly, or characteristics correlated with it).

The analysis from this paper can be extended in at least two directions. First, I have abstracted from efficiency costs associated with monopsony power. These costs may be very significant and affected by tax policy as well.[25] Second, I have treated monopsony power as exogenously determined. However, in reality monopsony power is unlikely to be policy-invariant. Extending the analysis to include distortions associated with monopsony power and a potential role for the government to affect monopsony power (e.g., through competition policy) seems highly policy-relevant.

[24] In particular, optimal marginal tax rates are negative (see Hariton and Piaser (2007) and da Costa and Maestri (2018)) and monopsony power cannot raise welfare. See also the discussion in Section 1.

[25] See Berger et al. (2019) for a recent attempt to quantify the efficiency costs due to monopsony power.

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. (2016). Declining labor and capital shares. Stigler Center Working Paper No. 2, Chicago: Stigler Center.

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

Bastani, S., T. Blumkin, and L. Micheletto (2015). Optimal wage redistribution in the presence of adverse selection in the labor market. *Journal of Public Economics 131*, 41–57.

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.

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 (2018). Optimal Mirrleesian taxation in non-competitive labor markets. *Economic Theory*. Forthcoming.

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.

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.

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

Robinson, J. (1969). *The economics of imperfect competition.* London: Palgrave Macmillan UK.

Rothstein, J. (2010). Is the EITC as good as an NIT? Conditional cash transfers and tax incidence. *American Economic Journal: Economic Policy 2*(1), 177–208.

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.

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

Stantcheva, S. (2014). Optimal income taxation with adverse selection in the labour market. *Review of Economic Studies 81*(3), 1296–1329.

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.

A 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 social welfare (11) subject to resource and incentive constraints. The allocation variables are labor effort $l(n)$ and utility $v(n)$ for each worker, and the profits $\pi(n)$ firms make from hiring a worker with ability n. For a given tax on profits, the utility of capitalists $v(k)$ can then be determined from equation (5). Moreover, the consumption of a worker with ability n is obtained from the relationship $c(n) = v(n) + \phi(l(n))$. To derive the resource constraint, substitute $T(z(n)) = z(n) - c(n) + \sigma(n)(1-\tau)\pi^*$ and $\pi(n) = nl(n) - z(n)$ in the government's budget constraint (12). Expressed in terms of the allocation variables, the latter can be written as

$$\int_{n_0}^{n_1} nl(n)f(n)dn = \int_{n_0}^{n_1} \Big[v(n) + \phi(l(n)) \Big] f(n)dn + \sigma^*(k)(1-\tau)\int_{n_0}^{n_1} \pi(n)f(n)dn + G. \quad (31)$$

Equation (31) states that aggregate output equals the sum of consumption by workers (first term), by capitalists (second term) and by the government (third term).

In addition to the resource constraint, the allocation must also satisfy incentive constraints. To derive the first of these, write workers' utility as $v(n) = z(n) - T(z(n)) + \sigma(n)(1-\tau)\int_{n_0}^{n_1}\pi(m)f(m)dm - \phi(l(n))$. Differentiating with respect to ability gives

$$v'(n) = (1 - T'(z(n)))z'(n) + \sigma'(n)(1-\tau)\int_{n_0}^{n_1} \pi(m)f(m)dm - \phi'(l(n))l'(n). \quad (32)$$

Next, impose profit maximization (7) and use the relationship $\pi(n) = nl(n) - z(n)$. Condition (32) can then be written as

$$v'(n) = \frac{\phi'(l(n))}{n}\Big[l(n) - \pi'(n) \Big] + \sigma'(n)(1-\tau)\int_{n_0}^{n_1} \pi(m)f(m)dm. \quad (33)$$

This condition is similar to the incentive constraint from the Mirrlees (1971) problem. Both conditions coincide if labor markets are competitive (in which case $\pi(n) = \pi'(n) = 0$ for all n). If firms have monopsony power, the standard incentive constraint is modified in two ways. First, equation (33) states that workers' utility increases less quickly in ability if firms generate more profits from hiring more productive workers (i.e, if $\pi'(n) > 0$). Second, it states that utility increases more quickly in ability if more productive workers own more shares (i.e., if $\sigma'(n) > 0$) and if aggregate after-tax profits are high.

To derive the second incentive constraint, differentiate the condition for profits (8) with respect to ability to find

$$\pi'(n) = \mu(n)l(n) + \frac{\mu'(n)}{\mu(n)}\pi(n). \tag{34}$$

Intuitively, profits increase more rapidly in worker productivity the higher is monopsony power and labor effort. Conversely, profits increase less quickly in worker productivity if labor markets for more productive workers are less monopsonistic (i.e., if $\mu'(n) < 0$). Combining equations (33) and (34) gives

$$\upsilon'(n) = \frac{\phi'(l(n))}{n}\left[(1-\mu(n))l(n) - \frac{\mu'(n)}{\mu(n)}\pi(n)\right] + \sigma'(n)(1-\tau)\int_{n_0}^{n_1} \pi(m)f(m)dm. \tag{35}$$

Recall that by assumption, more productive workers do not suffer more from monopsony and do not own fewer shares: $\mu'(n) \leq 0$ and $\sigma'(n) \geq 0$. Equation (35) then implies that utility is weakly increasing in ability: $\upsilon'(n) \geq 0$. The only instance where utility is constant across workers is if firms have full monopsony power (i.e., if $\mu(n) = 1$ for all n) and if income from capital does not vary with ability (i.e., if $\sigma'(n) = 0$). In that case, all workers 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 $l(n)$ for each worker type n to maximize social welfare (11), subject to the resource constraint (31) and the incentive constraints (34) – (35). The final restriction we need to impose on the allocation variables is that the profits from hiring the least productive workers are non-negative: $\pi(n_0) \geq 0$. This constraint guarantees that firms are willing to hire workers of all ability levels.[26] It is shown in Appendix D that the 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 (8). The optimal tax problem can now be formulated as a standard optimal

[26] To see why, note that the general solution to the differential equation (34) is

$$\pi(n) = \mu(n)\left[\frac{\pi(n_0)}{\mu(n_0)} + \int_{n_0}^{n} l(m)dm\right], \tag{36}$$

which simplifies to equation (8) 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.

control problem where $v(n)$ 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 C.

To make sure that the optimal allocation (as implicitly characterized in Appendix C) 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 monotonically increasing in ability: $z'(n) > 0$. This condition serves two purposes. First, it guarantees that workers 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 conditions for profit maximization are satisfied – see Appendix B for details.

B Monotonicity condition

This Appendix demonstrates the equivalence between the monotonicity condition $z'(n) > 0$ and the requirement that the second order-conditions for the profit maximization problem (6) are satisfied. To do so, note that the constraint in the firm's maximization problem (6) is always binding. If not, firms can raise profits by increasing l. Invert the constraint with respect to labor effort to write $l = \hat{l}(z, v(n))$. The profit maximization problem then reads

$$\max_{z} \quad n\hat{l}(z, v(n)) - z. \tag{37}$$

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

$$\frac{n(1 - T'(z))}{\phi'(\hat{l}(z, v(n)))} - 1 = 0. \tag{38}$$

The second-order condition is strictly satisfied if the left-hand side of equation (38) is decreasing in earnings z. The latter is true if and only if

$$-\phi''(l) - n^2 T''(z) < 0, \tag{39}$$

where I used the first-order condition (38) and substituted out for $\hat{l}(z, v(n)) = l$. Because $\phi(\cdot)$ is strictly convex, condition (39) is satisfied as long as the tax function is not too concave.

To determine how earnings z vary with ability, rewrite equation (38) and define

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

Next, apply the implicit function theorem and use equation (38) and the property $\hat{l}_v = -1/\phi'$

to find

$$z'(n) = -\frac{L_n(z,n)}{L_z(z,n)} = \frac{\phi'(l) + \frac{\phi''(l)}{\phi'(l)} n\upsilon'(n)}{\phi''(l) + n^2 T''(z)}. \tag{41}$$

From the incentive constraint (35), $\upsilon'(n) \geq 0$ as long as monopsony power is non-increasing and shareholdings are non-decreasing in ability. The numerator in (41) 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 (39) is satisfied. Therefore, if the allocation satisfies the monotonicity condition $z'(n) > 0$, it follows that the first-order condition for profit maximization (38) is both necessary and sufficient.

C Lagrangian and first-order conditions

Written in terms of the allocation variables, the government's optimization problem reads

$$\max_{[\upsilon(n),\pi(n),l(n)]_{n_0}^{n_1}} \mathcal{W} = \int_{n_0}^{n_1} \Big[\omega(n)\upsilon(n) + \omega(k)\sigma^*(k)(1-\tau)\pi(n)\Big] f(n)dn, \tag{42}$$

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

$$\forall n: \quad \upsilon'(n) = \frac{\phi'(l(n))}{n}\left[(1-\mu(n))l(n) - \frac{\mu'(n)}{\mu(n)}\pi(n)\right] + \sigma'(n)(1-\tau)\int_{n_0}^{n_1}\pi(m)f(m)dm,$$

$$\forall n: \quad \pi'(n) = \mu(n)l(n) + \frac{\mu'(n)}{\mu(n)}\pi(n),$$

$$\pi(n_0) \geq 0.$$

The Lagrangian associated with the above maximization problem is

$$\begin{aligned}
\mathscr{L} = \int_{n_0}^{n_1} \Bigg[&\Big((\omega(n)-\eta)\upsilon(n) + (\omega(k)-\eta)\sigma^*(k)(1-\tau)\pi(n) + \eta(nl(n) - \phi(l(n)) - G)\Big) f(n) \\
&+ \chi(n)\left[\frac{\phi'(l(n))}{n}\left((1-\mu(n))l(n) - \frac{\mu'(n)}{\mu(n)}\pi(n)\right) + (1-\tau)\sigma'(n)\int_{n_0}^{n_1}\pi(m)f(m)dm\right] \\
&+ \chi'(n)\upsilon(n) + \lambda(n)\left(\mu(n)l(n) + \frac{\mu'(n)}{\mu(n)}\pi(n)\right) + \lambda'(n)\pi(n)\Bigg] dn \\
&+ \chi(n_0)\upsilon(n_0) - \chi(n_1)\upsilon(n_1) + \lambda(n_0)\pi(n_0) - \lambda(n_1)\pi(n_1) + \xi\pi(n_0).
\end{aligned} \tag{43}$$

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

$$\upsilon(n): \quad (\omega(n)-\eta) f(n) + \chi'(n) = 0, \tag{44}$$

$$\pi(n): \quad (\omega(k)-\eta)\sigma^*(k)(1-\tau)f(n) - \frac{\mu'(n)}{\mu(n)}\left(\chi(n)\frac{\phi'}{n} - \lambda(n)\right) \tag{45}$$

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

I assume the system (44) – (54) admits an interior solution, where earnings $z(n) = nl(n) - \pi(n)$ satisfy the monotonicity condition $z'(n) > 0$.

D Derivation optimal tax formulas

This Appendix derives optimal tax formulas for the general case where monopsony power $\mu(n)$ and shareholdings $\sigma(n)$ might vary with ability. To derive the first result from Proposition 3, use equations (44) and (51) to obtain an expression for $\chi(n)$:

$$
\begin{aligned}
\chi(n) &= \chi(n_1) - \int_n^{n_1} \chi'(m)dm && (55) \\
&= -\int_n^{n_1} (\eta - \omega(n)) f(m)dm.
\end{aligned}
$$

Evaluate equation (55) at $n = n_0$ and use the transversality condition (50) to find

$$
\int_{n_0}^{n_1} (\eta - \omega(n)) f(n)dn = 0, \tag{56}
$$

which coincides with equation (13) after dividing by η and using the definition for the welfare weight $g(n) = \omega(n)/\eta$.

To proceed, note that $\omega(n)$ is non-increasing in ability. Condition (44) and the transversality conditions (50) – (51) then imply that $\chi(n) \leq 0$ for all n and strictly negative almost everywhere when the Pareto weights are strictly decreasing. Next, rewrite equation (45) to find

$$\lambda'(n) + \frac{\mu'(n)}{\mu(n)}\lambda(n) = \eta\Bigg[(1-\tau)\left(\sigma^*(k)(1-g(k)) + \int_{n_0}^{n_1}\sigma'(m)(1-\bar{g}(m))(1-F(m))dm\right)f(n) - \frac{\mu'(n)}{\mu(n)}(1-T'(z(n)))(1-\bar{g}(n))(1-F(n))\Bigg], \quad (57)$$

where I used equations (55) and (7) to substitute out for $\chi(n)$ and ϕ'/n, respectively. Moreover, I used the definitions for the welfare weights $g(n) = \omega(n)/\eta$, $g(k) = \omega(k)/\eta$ and

$$\bar{g}(n) = \frac{\int_n^{n_1} g(m)f(m)dm}{1-F(n)}, \quad (58)$$

denotes the average welfare weight of workers with ability *at least* equal to n. Equation (56) implies that $\bar{g}(n_0) = 1$ and because $\omega(n)$ is non-increasing, $\bar{g}(n) \leq 1$ for all n.

Equation (57) is a linear differential equation in $\lambda(n)$. Using the transversality condition (53), the solution is

$$\lambda(n) = \frac{\eta}{\mu(n)}\Bigg[\int_n^{n_1}\mu'(m)(1-T'(z(m)))(1-\bar{g}(m))(1-F(m))dm - \bar{\mu}(n)(1-\tau)(1-F(n))\left(\sigma^*(k)(1-g(k)) + \int_{n_0}^{n_1}\sigma'(m)(1-\bar{g}(m))(1-F(m))dm\right)\Bigg], \quad (59)$$

where $\bar{\mu}(n)$ is the average monopsony power for workers with ability at least equal to n. To sign $\lambda(n)$, note that equation (7) implies $T'(z(n)) < 1$ at any interior solution. Moreover, monopsony power is non-increasing and shareholdings are non-decreasing. Finally, the welfare weights satisfy $\bar{g}(n) \leq 1$ and $g(k) \leq \bar{g}(n_0) = 1$. Consequently, $\lambda(n) \leq 0$. From the first-order conditions (52) and (54) it follows that $\xi \geq 0$. Hence, without loss of generality, firms do not earn profits from hiring the least productive workers: $\pi(n_0) = 0$.

To derive an expression for the optimal marginal tax rate, consider the first-order condition (46) and use the condition for profit maximization $n(1-T') = \phi'$. Rearranging gives

$$T'(z(n))nf(n) = -\frac{\mu(n)\lambda(n)}{\eta} - \frac{\chi(n)}{\eta n}\left((1-\mu(n))(\phi' + \phi''l(n)) - \phi''\frac{\mu'(n)}{\mu(n)}\pi(n)\right). \quad (60)$$

Substituting $\chi(n)$ and $\lambda(n)$ from equations (55) and (59), equation (60) can be written as

$$T'(z(n))nf(n) = \quad (61)$$

$$\bar{\mu}(n)(1-F(n))(1-\tau)\sigma^*(k)(1-g(k))$$

$$+\,\bar{\mu}(n)(1-\tau)(1-F(n))\int_{n_0}^{n_1}\sigma'(m)(1-\bar{g}(m))(1-F(m))dm$$

$$
- \int_{n}^{n_1} \mu'(m)(1 - T'(z(m)))(1 - \bar{g}(m))(1 - F(m))dm
$$

$$
+ (1 - \bar{g}(n))(1 - F(n)) \frac{\phi'}{n} \left[(1 - \mu(n)) \left(1 + \frac{\phi'' l(n)}{\phi'} \right) - \pi(n) \frac{\phi''}{\phi'} \frac{\mu'(n)}{\mu(n)} \right].
$$

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 (61) by $nf(n)$ and rearranging, we obtain equation (21) from Proposition 3:

$$
\begin{aligned}
T'(z(n)) = {} & \bar{\mu}(n)(1 - \tau)\sigma^*(k)(1 - g(k)) \left(\frac{1 - F(n)}{nf(n)} \right) \\
& + (1 - \mu(n))(1 - T'(z(n))) \left(1 + \frac{1}{\varepsilon(n)} \right) (1 - \bar{g}(n)) \left(\frac{1 - F(n)}{nf(n)} \right) \\
& - \left(\frac{1 - F(n)}{nf(n)} \right) \left[\left(\frac{\mu'(n)}{\mu(n)} \right) \left(\frac{\pi(n)}{\varepsilon(n) l(n)} \right) (1 - T'(z(n)))(1 - \bar{g}(n)) \right. \\
& \qquad + \int_{n}^{n_1} \mu'(m)(1 - T'(z(m)))(1 - \bar{g}(m)) \left(\frac{1 - F(m)}{1 - F(n)} \right) dm \\
& \qquad \left. - \bar{\mu}(n)(1 - \tau) \int_{n_0}^{n_1} \sigma'(m)(1 - \bar{g}(m))(1 - F(m))dm \right].
\end{aligned} \tag{62}
$$

If monopsony power and shareholdings do not vary with ability (i.e., $\mu'(n) = \sigma'(n) = 0$), the terms in the last three lines of equation (62) cancel. Substituting $\mu(n) = \bar{\mu}(n) = \mu$ gives equation (14) from Proposition 1.

From equation (62) 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 and shareholdings are non-decreasing: $\mu'(n) \leq 0$ and $\sigma'(n) \geq 0$. Moreover, $\chi(n) \leq 0$ implies $\bar{g}(n) \leq 1$, by assumption $g(k) \leq 1$ and from the profit-maximization condition (7) it follows that the marginal tax rate cannot exceed one at an interior solution. Hence, all terms on the right-hand side of equation (62) are non-negative. The optimal marginal tax rate is therefore generally positive.

E 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

$$\mathscr{L}(\alpha) = \int_{n_0}^{n_1} \Bigg[\Big((\omega(n)-\eta)v(n) + (\omega(k)-\eta)\sigma^*(k)(1-\tau)\pi(n) + \eta(nl(n) - \phi(l(n)) - G) \Big) f(n)$$
$$+ \chi(n) \left[\frac{\phi'(l(n))}{n} \left((1-\mu(n)(1+\alpha))l(n) - \frac{\mu'(n)}{\mu(n)}\pi(n) \right) + (1-\tau)\sigma'(n) \int_{n_0}^{n_1} \pi(m)f(m)dm \right]$$
$$+ \chi'(n)v(n) + \lambda(n) \left(\mu(n)(1+\alpha)l(n) + \frac{\mu'(n)}{\mu(n)}\pi(n) \right) + \lambda'(n)\pi(n) \Bigg] dn$$
$$+ \chi(n_0)v(n_0) - \chi(n_1)v(n_1) + \lambda(n_0)\pi(n_0) - \lambda(n_1)\pi(n_1) + \xi\pi(n_0), \tag{63}$$

which is the optimized Lagrangian (43) evaluated at $\hat{\mu}(n) = \mu(n)(1+\alpha)$. Here I used the fact that the increase in monopsony power is proportional, which implies

$$\frac{\hat{\mu}'(n)}{\hat{\mu}(n)} = \frac{\mu'(n)(1+\alpha)}{\mu(n)(1+\alpha)} = \frac{\mu'(n)}{\mu(n)}. \tag{64}$$

By the Envelope theorem, the welfare effect is

$$\frac{\partial \mathcal{W}(\alpha)}{\partial \alpha} = \frac{\partial \mathscr{L}(\alpha)}{\partial \alpha} = \int_{n_0}^{n_1} \left(-\chi(n)\frac{\phi'}{n} + \lambda(n) \right) \mu(n)l(n)dn. \tag{65}$$

Next, use the profit-maximization condition (7) and equations (55) and (59) to substitute out for $\chi(n)$ and $\lambda(n)$. This leads to equation (22) as stated in Proposition 3:

$$\frac{\partial \mathcal{W}(\alpha)}{\partial \alpha} = \eta \int_{n_0}^{n_1} \Bigg[\mu(n)(1-T'(z(n)))(1-\bar{g}(n)) - \bar{\mu}(n)(1-\tau)\sigma^*(k)(1-g(k)) \tag{66}$$
$$+ \int_{n}^{n_1} \mu'(m)(1-T'(z(m)))(1-\bar{g}(m)) \left(\frac{1-F(m)}{1-F(n)} \right) dm$$
$$- \bar{\mu}(n)(1-\tau) \int_{n_0}^{n_1} \sigma'(m)(1-\bar{g}(m))(1-F(m))dm \Bigg] (1-F(n))l(n)dn.$$

This condition simplifies considerably if one imposes that monopsony power and shareholdings do not vary with ability. In that case, the last two lines cancel and $\mu(n) = \bar{\mu}(n) = \mu$. Substituting this in equation (66) and collecting terms gives the result from Proposition 2.

As stated in the main text, equation (20) depends on marginal tax rates and labor effort, which are both endogenous. However, under the additional assumption that the utility function is iso-elastic (i.e., $u(c,l) = c - \frac{l^{1+1/\varepsilon}}{1+1/\varepsilon}$) it is possible to express the result from Proposition 2 in terms of exogenous variables only. To see this, note that Corollary 1 gives a closed-form expression for the marginal tax rate:

$$T'(z(n)) = \frac{\mu(1-\tau)\sigma^*(k)(1-g(k)) + (1-\mu)(1+1/\varepsilon)(1-\bar{g}(n))}{a(n) + (1-\mu)(1+1/\varepsilon)(1-\bar{g}(n))}. \tag{67}$$

Labor effort can then be determined from equation (7):

$$l(n) = n^{\varepsilon} \left(\frac{a(n) - \mu(1-\tau)\sigma^*(k)(1-g(k))}{a(n) + (1-\mu)(1+1/\varepsilon)(1-\bar{g}(n))} \right)^{\varepsilon}. \tag{68}$$

Substituting the above in equation (20) gives

$$\frac{\partial \mathcal{W}(\alpha)}{\partial \alpha} = \eta\mu \int_{n_0}^{n_1} \left[\left(\frac{a(n) - \mu(1-\tau)\sigma^*(k)(1-g(k))}{a(n)/(1-\bar{g}(n)) + (1-\mu)(1+1/\varepsilon)} \right) - (1-\tau)\sigma^*(k)(1-g(k)) \right]$$
$$\times (1-F(n))n^{\varepsilon} \left(\frac{a(n) - \mu(1-\tau)\sigma^*(k)(1-g(k))}{a(n) + (1-\mu)(1+1/\varepsilon)(1-\bar{g}(n))} \right)^{\varepsilon} dn, \tag{69}$$

which is expressed solely in terms of exogenous variables if one substitutes out for the welfare weights $g(k) = \omega(k)/\eta$, $g(n) = \omega(n)/\eta$ and uses that equation (13) implies $\eta = \int_{n_0}^{n_1} \omega(n) f(n) dn$.

F Additional graphs

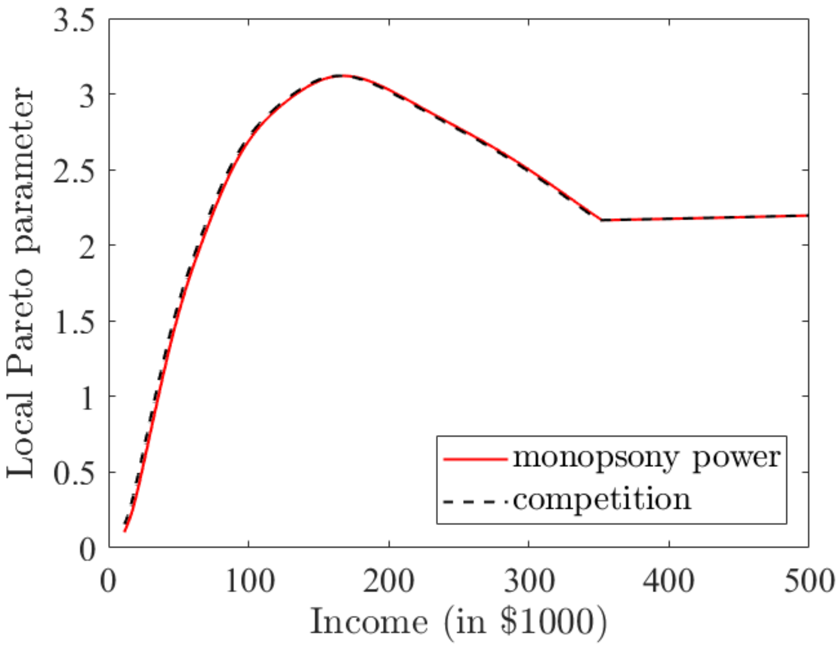


Figure 6: Local Pareto parameter

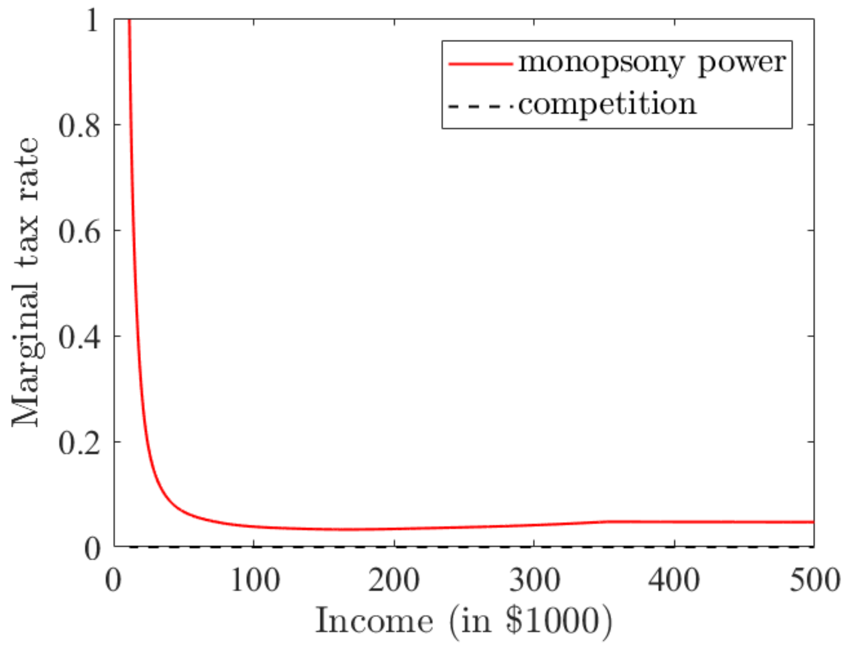


Figure 7: Comparison optimal marginal tax rates (case B)

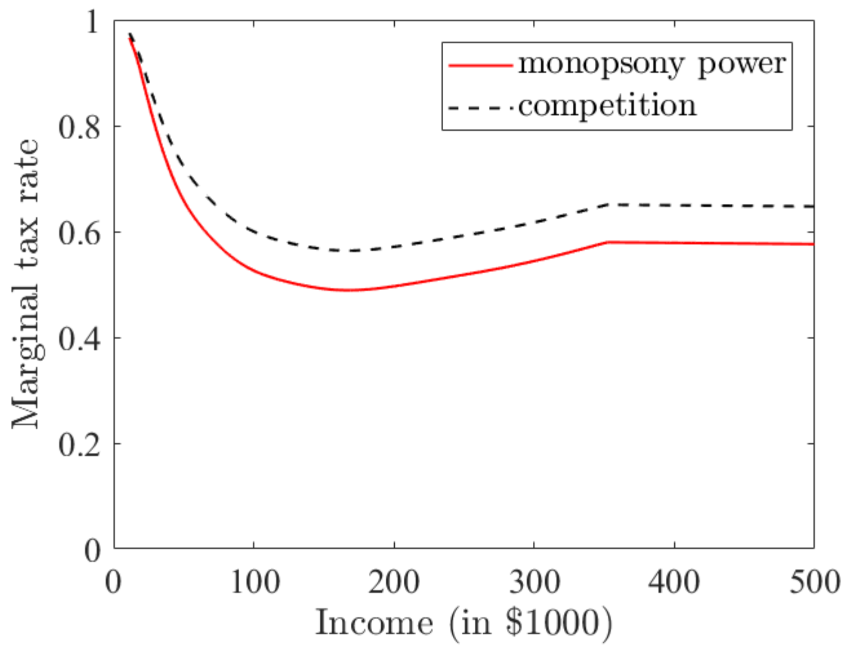


Figure 8: Comparison optimal marginal tax rates (case G)