Optimal taxation in a tournament setting

Master Thesis Policy Economics

ERASMUS UNIVERSITY ROTTERDAM

Erasmus School of Economics

Name student: Hubrecht Johannes Murre
Student ID number: 383606

Supervisor: dr. Aart Gerritsen
Second assessor:

Date final draft: April 26, 2020

Contents

1 Introduction

The system of income taxation often is a source of debate. Income taxation plays a significant role in the determination of the income of most people. Politicians often talk about how they are going to make the tax system more fair and simple. In the past decades, scientists have also joined the debate. Nowadays, a significant theoretical foundation about optimal taxation has been laid by multiple economists (Mankiw, Weinzierl, & Yagan, 2009; Piketty & Saez, 2013). However, there is still scope for further exploration of this area of research.

In this paper, I take a different view at optimal taxation. I make two important deviations from the "standard" literature. I want to investigate the question of optimal taxation in a tournament setting with a monopsonist firm.

The first deviation has to do with the type of economy we are in. I depart from from the simple view of wageworkers whose earnings ability directly determines their wage. This is replaced by a tournament setting. A tournament setting means that agents (in this paper there are two of them) are in a tournament battling for a prize. This resembles the case where colleagues are at the same time competitors. There is only one price: a promotion with salary increase attached to it. Contrary to the case in Mirrleesian optimal taxation models, earnings ability is now not the only factor determining how productive the agents are and, more importantly, who wins the promotion. Just like what is often the case in the real world, there are multiple other factors that can potentially affect the outcome: a good relationship with the manager, discrimination, virtue signaling or even just plain luck can disturb a proper view of who has the highest ability and hence should win the promotion. In my model I will fix all these factors together in one variable, a so-called “noise” factor.

The second deviation I take from standard literature is that the labor market is characterized by a monopsony. This means that the employer is the only employer in that sector of labor. Therefore, employees do not have the possibility to play companies off against each other to see which company comes with the best salary offer. In the case described in this paper it is the other way around: the company lets the workers play off against each other to see who is the most productive against the lowest possible salary. I take this step in line with the trend that big companies take over the market in some sectors. Particularly in the digital and service sectors this is the case, with big companies like Facebook, Google, Microsoft, Huawei, Amazon, eBay and Aliexpress. A monopsony setting is the way to model this phenomenon.

Given these two deviations, what I want investigate is how the government can use taxes in such a system to redistribute not only between workers, but also between the firm’s owners (who incur the profits) on the one hand and the workers on the other hand. The government has two instruments to use to intervene in the economy. The first one is a proportional labor tax rate that is applied to all labor income. The second one is a universal transfer. This transfer is received by the agents independently of whether they work or not. I want to find an expression for the tax rate and the universal transfer under which social welfare is at its maximum.

I will demonstrate that I find that the expression for the optimal tax rate is remarkably similar to the expression that is common in Mirrleesian optimal taxation models. The difference lies in the handling of capital ownership by the social welfare function of the government. This happens in two ways. First, the government must take into account the capital owners’ losses from compensating workers for the direct mechanical effect. This leads to a lower optimal tax rate than in the standard result in Mirrlees-like models. Second, provided that profits do not flow directly to the workers, the government can indirectly tax capital owners via the “compensation reflex” of the firm. This leads to a higher tax rate compared to the standard Mirrleesian outcome.

The remainder of this paper is constructed as follows. In the next section, I give an overview of relevant literature to make clear to which context this paper applies and to help the reader to place this paper in the discourse about optimal tax literature and tournament models. Subsequently, the model itself is presented and explained. In section 4, the model is solved in the form of expressions for the equilibrium wages. In section 5, the implications for optimal income taxation are derived, and three possible cases with respect to capital owners are introduced, which are analyzed in sections 6, 7 and 8. In section 9, the results are compared with existing literature. Some policy recommendations are presented in section 10. Section 11 concludes.

2 Theoretical framework

I will now look more closely at the relevant literature for this research. First, I will look at the existing literature about optimal taxation in general. Then, I will take a look at the literature about tournament models. After that, I will explore the literature about the shift of surplus from labor to capital. Lastly, I will demonstrate what literature says about the concept of monopsony.

2.1 Optimal taxation

Until now, in the literature about optimal tax theory, models were based on a Mirrleesian theory of optimal taxation (Mirrlees, 1971). In those models, the challenge is to get as close as possible to the first best optimum in income taxation theory: Tinbergen's talent tax, which says that the amount of taxation an individual must optimally pay is based on the individual's talent or capacity to earn income (Tinbergen, 1970). The idea behind this is that the amount of tax an individual must pay is fixed: there are no distortive behavioral responses in labor supply, because working less will not result in a smaller amount in taxes to pay.

The obvious problem is that the exact eearnings capacity of agents is unknown. It cannot be used as a tax base. Therefore, I take the actual income as a measure for the earnings capacity and use that as the tax base. However, an agent will react to this by adjusting the amount of labor he supplies: if he supplies less labor, his gross income is lower and hence he has to pay taxes over a smaller tax base. This is a distortive effect of income taxation:

the agent does not supply the optimal amount of work. So, how does this distortion occur? The agent decides to work an additional hour as long as the yields of this hour of work (additional income, joy, etc.) weigh up to the costs (time, energy, etc.). The marginal yields of an hour of work usually diminishes with the number of hours worked, while the marginal costs show an increasing trend (Boeri & Van Ours, 2013). The intersection of these two is the point where marginal yields of work equal marginal costs, and thus determines the number of hours the agent optimally works. However, income taxes absorb part of the yields of work and make the marginal yields of each hour worked go down. Thus, the yields of the last couple of hours the agent works do no longer weigh up to the costs of working these hours. Hence, the agent decides not to work those hours anymore: his labor supply goes down. And because working those hours would initially contribute to utility (because yields > costs), and subsequently to welfare, this decline in labor supply means a loss to society in terms of welfare, also called the "deadweight loss".

Why should income taxes then be levied? Because people have, to a certain extent, a preference for equality. This preference is reflected in the redistributional policies of our governments. Tax revenue can be used to redistribute from agents with high earnings to agents with low earnings. So, when it comes to optimal taxation, there is a tradeoff: the gains from redistribution versus the deadweight loss from the distortive effect in labor supply. In the standard Mirrleesian result, the optimal tax scheme is constituted of the optimal marginal tax rates at each income level. All different income levels together form an income distribution over the whole of society. This income distribution reflects the distribution of earnings-ability via a one-to-one relationship. The core tradeoff in these models yields that it is optimal to increase the marginal tax rate as long as the gains from redistribution outweight the deadweight loss from distortion (Mirrlees, 1971; Saez, 2001; Mankiw et al., 2009). There are multiple factors that influence the optimal marginal tax rate. Among them is the number of households that earn the marginal income. These reflect the households that will re-evaluate their labor supply at the **intensive margin** after a change in that marginal tax rate. The number of households that earn more than the marginal income will not alter their labor supply since their situation is not changed at the margin. However, they do pay a higher average tax rate, generating more tax revenue. A consequence of paying a higher tax rate can be that the agent quits working altogether, which is an effect on labor supply at the **extensive margin** (Saez, 2002).

2.2 Tournament model

Mirrleesian tax models investigate a society where people only differ in earnings capacity, or earnings ability (Mirrlees, 1971; Saez, 2001; Mankiw et al., 2009). Earnings capacity is therefore the only variable determining one's income, and thus: differences in wage reflect differences in earnings capacity. This is quite an accurate description of reality for many labor markets; especially for low skilled labor sectors like production factories, assembly line jobs, picking fruits, farm work, construction and industrial services. There, basic economic theory applies: the wage reflects the marginal product of labor. This implicates that the way people are taxed is solely based on earnings capacity. However, there are labor markets for which the earnings capacity-based models are not so realistic. For example, looking at top level managers, a recent estimation shows that above a wage of €600.000 euros, there is no further increase in the quality of managers (Berg, Boutorat, Garretsen, Mounir, & Stoker, 2019). This means that, from that point in the wage distribution of top managers onwards, ability is not the (only) determinant of the wage. Besides that, factors like the relationship with the manager, discrimination, virtue signaling and plain luck can affect one's wage. In these cases, wages are determined differently, compared to low income wage workers. There is no one-to-one relationship between earnings ability and income. Consequently, optimal tax theory should take this into account. Indeed, a tax schedule that is based on a theory in which the earnings capacity determines the wage may not be optimal to apply to a situation where wages are determined in a different way. The different determination of wages as described above is common with high income earners like top level managers, top artists or top sports players. However, also in the workforce layers below these effects can play a role. We encounter a problem similar to that of the Tinbergen talent tax: the manager cannot see the true production ability of his employee; he decides over wages and promotions on the basis of what **he regards** as the productivity of the employee. The "tournament model" documented in the seminal paper of Lazear and Rosen (1981) describes accurately how wages are determined in those types of labor markets. In the standard model, two agents within the same company compete for a promotion. The promotion is awarded to the agent with highest productivity. However, productivity does not have a one-to-one relationship with production ability. All possible reasons that could harm this one-to-one relationship are taken together in a so-called "noise factor", that troubles the relationship between ability and productivity. Agents choose their effort level, based on what they think the effort level of their competitor will be. The higher their effort level, the higher their productivity level. However, they must deal with increasing costs of effort. Ultimately, the agent with the highest productivity (which is determined by production ability, effort and “noise”) wins the tournament and receives the promotion and corresponding wage increase. An overview of this theory can be found in Waldman (2007).

2.3 Why monopsony?

This tournament model is applied within a monopsonist labor market. This is another deviation from the Mirrleesian optimal taxation literature, which assumes perfect competition in the labor market, so that firms, on aggregate, do not make profits. However, recently, there has been a lot of debate concerning the shift of market power between firms and workers. Firms accrue more power in the labor market, even up to a (semi-)monopsonistic level in some cases; for example in the Dutch parcel delivery market (PostNL), the food delivery market (TakeAway.com), etcetera. Workers in those sectors can easily be replaced by other workers. More importantly, in a pure monopsony, workers do not have the possibility to work for another firm, since there is only one firm employing in their sector. This weakens their bargaining position, since they basically must choose between working for the monopsonist company or not working at all. This low bargaining power of the workers results in wages settling below marginal productivity, at the point where workers are just willing to work instead of quitting the job altogether. Piketty and Zucman (2014) and Karabarbounis and Neiman (2014) give an overview of the decreasing relative price of labor. Meanwhile, De Loecker and Eeckhout (2017) and Eggertsson, Robbins, and Wold (2018) explicitly point to firms increasing their profit shares: an important cause of the decrease of labor income compared to profits. Autor, Dorn, Katz, Patterson, and Van Reenen (2017) also write about this shift of surplus from labor to capital, which indicates that these firms have market power in the labor market: they are monopsonist firms. They can set wages at a low level, even below marginal productivity. The firms can do this because competition between multiple workers drives wages down to the level where they are just willing to take the job, instead of any possible outside option. The entire surplus of the production is incurred by the firm. The rise of this type of firms gives me reason to apply the tournament setting for wage determination to a monopsonistic case.

2.4 Similar research

To my knowledge, there is one paper that investigates a similar case as this paper does, namely the model of Persson and Sandmo (2005). They took the initial tournament model of Lazear and Rosen (1981) and implemented a simple tax system. Their analysis brought forth some interesting results that differ with the current optimal tax literature. One particularly remarkable result is that an increase in the tax rate could lead to an increase in income inequality. This is possible because in the tournament setting, the firm owner has an incentive to keep the difference in income between the winner and the loser of the tournament large enough. Usually, an income tax tries to achieve the opposite: more income equality. An increase in the income tax reduces the difference in net incomes between the winner and the loser, and thus mitigates the importance of winning the tournament. The firm owner responds to that by further increasing the gross difference between the rewards for the winner and the loser. Ultimately, the difference in net income stays constant, but the absolute income levels of both the winner and loser decrease because they must pay a larger share of their incomes in taxes. This implies that the **relative** net incomes of the winner and loser become more unequal.

Another result is that their setup seems to drive the equilibrium tax rate to 100%: due to the monopsonistic setting, workers do not incur any profits. They earn in expectation just their cost of working, which is the absolute minimum to get them to work anyway. It is the tournament system that causes them to end up with different gross incomes. Now, when the tax rate is 100%, nobody will work, wages are thus zero and the cost of effort is zero. This means that there is at least perfect equality, but it is not so realistic.

Persson and Sandmo (2005) made an important decision by disallowing people who do not work to receive a government transfer. Unemployment benefits or other social security systems do not exist in their analysis. In their model, they have a so-called “lumpsum transfer” that every worker receives from the government. But, if an agent chooses not to work, he does not receive this lumpsum transfer anymore. However, it hardly ever occurs that people who do not work also have no income. Many (developed) countries have a system of unemployment benefits. In such welfare states, even people who do not work have a certain income (e.g. Bijstand in the Netherlands (Rijksoverheid, 2020)) to meet the basic needs. The model used in this paper is therefore similar to the work of Persson and Sandmo (2005), but differs on an important aspect: contrary to their setup, I will assume that the lumpsum transfer from the government in the model is universal: it does not matter whether you work or not in order to receive it. Later on, I will compare my findings with the results of Persson and Sandmo (2005).

Another paper that combines the elements of a tournament model and taxation is the paper of Skåtun (2017), although his focus is primarily on risk-aversion.

3 The model

I will now walk through the model in detail. Remember that it is a monopsonist firm that creates a tournament setting for their workers. There are two identical workers who participate in the tournament[1]. Let's say they compete for a promotion. The winner gets the promotion and earns a salary W_1. The loser earns W_2. There is also a government who sets the tax regime. It sets a marginal tax rate t and a fixed transfer T. When defining Y_i as the after-tax income of worker i with $i = 1, 2$, I get

$$Y_i = (1-t)W_i + T$$

I exclude income effects and risk-aversion from the analysis. This makes the analysis less complex, while the general results are still valuable for my goal of finding the optimal tax scheme, since Saez (2001) argued that it is primarily the uncompensated effect (so, without income effects) that matters for the determination of optimal tax rates. Assume that the worker's utility function depends on his after-tax income and the effort level μ_i:

$$E[U(Y_i, \mu_i)] = E(Y_i) - C(\mu_i), \tag{1}$$

where $C(\mu_i)$ is the cost of effort with $C'(\mu_i) > 0$ and $C''(\mu_i) > 0$. This utility function is linear in income, which means that indeed there are no income effects and risk-aversion in the analysis.

The worker's productivity is given by $q_i = \mu_i + \varepsilon_i$, where ε_i is a stochastical noise factor with a mean of 0, a variance of σ^2 and with $Cov(\varepsilon_i, \varepsilon_j) = 0$. With this productivity, both workers simultaneously choose their effort level. Naturally, they cannot choose their noise level since this is a random factor. After they have committed to an effort level, the random factor

[1]Note that this model is to a large extent similar to the models it is based on, namely those from Lazear and Rosen (1981) and Persson and Sandmo (2005). It only differs in the way the fixed transfer is being handled. However, it is still presented wholly in order to make it easier for the reader to understand the model.

becomes known and subsequently their productivity levels. The worker with the highest productivity receives the promotion. The expected utility of the workers can be rewritten as

$$P_i[(1-t)W_1 + T_0] + (1-P_i)[(1-t)W_2 + T] - C(\mu_i)$$

Here, $P_i(\mu_i, \mu_j, \varepsilon_i, \varepsilon_j)$ is the function that defines the probability that worker i wins the tournament. This can be rewritten as

$$P_i \equiv prob(q_i > q_j) \equiv prob(\mu_i - \mu_j > \varepsilon_j - \varepsilon_i) \equiv G(\mu_i - \mu_j)$$

where $G(\cdot)$ is the cumulative distribution function (CDF) of $\varepsilon_j - \varepsilon_i$. That is, $\varepsilon_j - \varepsilon_i$ is the difference of two zero-mean stochastic variables, which itself is a new zero-mean stochastic variable. That variable can take a range of values with corresponding probability. The CDF represents the cumulative probability over this range of values.

3.1 Workers

Both workers maximize their expected utility, taking the effort level of the other worker as given. Marginal benefits of exerting effort must equal the marginal costs:

$$[W_1(1-t) - W_2(1-t)]\frac{dP_i}{d\mu_i} = C'(\mu_i) \tag{2}$$

Now, the following is true:

$$\frac{dP_i}{d\mu_i} \equiv \frac{dG(\mu_i - \mu_j)}{d\mu_i} = g(\mu_i - \mu_j)$$

where $g(\cdot) \equiv G'(\cdot)$ is the probability density function (PDF) of $\varepsilon_j - \varepsilon_i$. I have already assumed that the workers are identical. So, after they have made their decisions, they will exert the same amount of effort, and therefore have equal probability of winning the tournament. Hence, $\mu_i = \mu_j = \mu$, $P_i = P_j = \frac{1}{2}$ and $g(\mu_i - \mu_j) = g(0)$. Equation 2 can then be rewritten to

$$[W_1(1-t) - W_2(1-t)]g(0) = C'(\mu) \tag{3}$$

The expression $g(0)$ is the density of $\varepsilon_j - \varepsilon_i$ for identical individuals. It measures the importance of the noise factor in deciding the tournament winner. When there is much noise, $Var(\varepsilon_i)$ and $Var(\varepsilon_j)$ are large and so is $Var(\varepsilon_j - \varepsilon_i) \equiv Var(\varepsilon_i) + Var(\varepsilon_j)$. This stretches the PDF over a large range of values, thereby lowering the corresponding probability density of these values, which implies $g(0)$ goes down. Exerting more effort then does not so much increase the likelihood of winning: you can work much harder, but as long as disturbing factors have a large impact on how the firm views your productivity, it does not matter much. In order to still motivate the workers to exert effort to win the tournament, the wage gap $W_1 - W_2$ must be large. On the contrary, when there is little noise and $g(0)$ is large, exerting a little bit more effort increases the likelihood of winning to a great extent: the firm can see very clearly which worker works harder, with few things that blur the view of the firm. In that case the wage gap does not have to be so large to motivate workers.

3.2 The firm

The objective of the firm is to maximize profits. The firm wants to make its workers exert as much effort as possible, because more effort increases productivity, and productivity adds to their profits. As I have already shown in the previous section, workers are motivated to exert effort through a large enough salary gap between the net incomes of the winner and the loser. Therefore, the firm takes the tax system into account when setting the payments W_1 and W_2, to make sure that the difference $W_1 - W_2$ stays large enough. To analyze the firm's anticipation to different tax rates, I constitute a profit function:

$$\pi = 2\mu + \varepsilon_i + \varepsilon_j - (W_1 + W_2) \tag{4}$$

While the noise factors do influence the firm's profits, they are random, exogenous and independent of the firm's behavior[2]. Therefore, since the noise factors both have a zero mean, the firm's profits are in expectation equal to

$$\pi = 2\mu - (W_1 + W_2)$$

This is the function the firm tries to maximize. Since the firm is a monopsonist, it does not have to compete with other firms for the workers to work for them. The only other option for the workers is not to work. Since there is no shortage of workers, workers don not have bargaining power. So, the firm only has to make sure that the worker participates. As long as the firm offers a wage that compensates the worker for the cost of exerting effort for work $C(\mu)$, the worker will work for the firm.

[2]I thereby assume that the earlier mentioned misassessments by the firm of the workers' productivities are randomly and symmetrically distributed around the true productivities of the workers, and that the firm does not misassess intentionally, so that this aspect of noise does not belong to the "firm's behavior".

If the workers' productivity is larger than his cost of effort, this creates a surplus of labor $q - C(\mu)$. Given that productivity is linear with effort, and cost of effort are convex, this difference $q - C(\mu)$ is largest at the point where the slopes of both are parallel. But, although there is a surplus, the wage does not have to be higher than the cost of effort to motivate workers to participate. Hence the whole surplus is captured by the firm[3]. This suppresses the workers' expected utility to the point where it is equal to the outside option. In this outside option, I assume that the worker does not work. The worker then does not make costs for exerting effort, and his income only comprises the universal fixed transfer T. Therefore, the following must hold:

$$\frac{1}{2}[W_1(1-t)+T]+\frac{1}{2}[W_2(1-t)+T]-C(\mu)=T$$

I assume that the labor market of the entire country is larger than only this tournament. In particular, the labor market is so large that one tax-paying worker more or less does not impact the magnitude of the fixed government transfer. Hence, when the workers would decide either to alter their labor supply or to stop working completely, this has no effect on the value of this outside option. However, within the sector of the labor market where the firm operates, this firm is the only employer and I further assume that, in the short term, employees cannot switch to another sector of the labor market, so that indeed they have no other outside option than to quit working altogether.

I now continue by simplifying the previous expression to obtain the participation constraint:

$$\frac{1}{2}(1-t)W_1+\frac{1}{2}(1-t)W_2-C(\mu)\geq 0 \tag{5}$$

Again, since the firm is a monopsonist, the workers have no market power. Therefore, they do not grab any share of the surplus, meaning this participation constraint will bind in equilibrium.

4 Equilibrium

In this section, I will determine the equilibrium wage setting of the model. From the firm's perspective, the tax rate and the worker's response to wage setting is given. The firm can only set the wages, for which the expressions are derived below using the Lagrangian optimization method.

[3]This could also reflect the case for government-directed firms, for example national railway firms, where profits are not allowed and consequently do not flow to the workers.

4.1 Maximizing profits

The firm maximizes profits subject to the participation constraint for the workers:

$$\max_{W1,W2} \quad 2\mu - (W_1 + W_2)$$
$$s.t. \qquad (1-t)W_1 + (1-t)W_2 - 2C(\mu) = 0$$

The Lagrangian is then

$$\mathcal{L} = \quad 2\mu - (W_1 + W_2) + \lambda[(1-t)W_1 + (1-t)W_2 - 2C(\mu)]$$

I calculate the First Order Conditions (FOCs) with respect to the variables the firm can choose, namely W_1 and W_2:

$$\frac{\partial \mathcal{L}}{\partial W_1} = 2\frac{\partial \mu}{\partial W_1} - 1 + \lambda\left[(1-t) - 2C'(\mu)\frac{\partial \mu}{\partial W_1}\right] = 0$$

$$\frac{\partial \mathcal{L}}{\partial W_2} = 2\frac{\partial \mu}{\partial W_2} - 1 + \lambda\left[(1-t) - 2C'(\mu)\frac{\partial \mu}{\partial W_2}\right] = 0$$

This tells us that if the firm increases W_1 by a unit, their profits are lowered by 1, while the winner gets $(1-t)$ more wage, which the firm values at λ. The firm attaches value to the winner's wage increase, because it relaxes the participation constraint: it creates room for the firm to set lower wages; for as long as the firm maintains the raw increase in W_1, it can then decrease both wages equally until the point where the average wage again makes the participation constraint bind.

The wage increase of the winner creates a behavioral response from the workers. They will both work harder, giving the firm $2\frac{\partial \mu}{\partial W_2}$ in additional profits. The $\frac{\partial \mu}{\partial W_2}$ term tells us how one extra unit of gross wage for the winner is translated into extra effort. Finally, the workers bear additional costs of effort, which are simply the additional exerted effort times $C'(\mu)$, which measures the change in costs of effort for a given change in effort.

For an increase in W_2, the explanation is similar, but in this case the effort response goes in the exact opposite direction. This is because the **difference** between the wages (hence, $W_1 - W_2$) matters for the effort exerted by the workers. If the firm increases both W_1 and W_2 with the same amount, keeping $W_1 - W_2$ constant, nobody will work harder: the workers will not exert more effort. The firm has then only thrown away some share of its surplus to the workers.

This exercise can be modified a bit, by letting the firm give 1 unit of income extra to the winner and take the same amount away from the loser. The winner then increases his income with $1 \cdot (1-t) = 1-t$ on net, and the loser sees his income decrease with the same amount. Again, the workers will alter their behavior in response to this wage regime change. Because the wage gap between the winner and loser is widened, winning the tournament becomes even more rewarding. The workers want therefore to exert even more effort to win. Assuming that the firm will indeed do this, I rearrange the FOCs to obtain:

$$1 - 2\frac{\partial \mu}{\partial W_1} = \lambda\left[(1-t) - 2C'(\mu)\frac{\partial \mu}{\partial W_1}\right] \tag{6}$$

$$1 + 2\frac{\partial \mu}{\partial W_1} = \lambda\left[(1-t) + 2C'(\mu)\frac{\partial \mu}{\partial W_1}\right] \tag{7}$$

Dividing everything by 2 and taking the differences gives

$$2\frac{\partial \mu}{\partial W_1} = 2\frac{\partial \mu}{\partial W_1}\lambda C'(\mu) \tag{8}$$

Here, it can be seen that the firm equalizes two things. On the left hand side, there are the additional profits from the extra effort exerted by the workers due to the wage intervention. On the right hand side, the firm's valuation of the additional costs of effort for the workers is visible.

4.2 Expressions for the wages

I now proceed to the expressions for the wage rates. I add up equations 6 and 7 to obtain

$$\lambda = \frac{1}{1-t}$$

Implementing this into equation 8 and rearranging yields

$$C'(\mu) = 1 - t \tag{9}$$

Since an increase in effort increases productivity, and thus profits, the firm wants the workers to exert as much effort as possible. However, the marginal costs of effort are increasing in effort (remember that $C''(\mu) > 0$). That means that it is not optimal to let effort increase unboundedly, because the firm needs to compensate the workers for every increase in their

costs of effort in order to keep them on their participation constraint. This can already be seen from equation 8. So, the firm optimally sets wages such that the workers choose their effort level efficiently: at the point where the marginal productivity of effort equates the marginal costs of effort. That point is defined in equation 9.
Now, from equation 3, I also know that the net wage difference $(1-t)W_1-(1-t)W_2$, scaled by the noise factor $g(0)$, equals the increase in the cost of effort of the worker $C'(\mu)$. The firm knows that when workers choose their effort, equation 9 will hold and sets W_1 and W_2 accordingly. I therefore combine equation 3 with 9 to obtain:

$$W_1 - W_2 = \frac{1}{g(0)} \tag{10}$$

This equation determines the equilibrium spread in wages: wages for the winner and loser should be far apart when $g(0)$ is small (remember: this is when there is a lot of variance in the "noise-factors"), and there should be just a small differential between the wages if $g(0)$ is large, e.g. when there is little "noise".

I take equation 5 and modify it by a bit:

$$W_1 + W_2 = \frac{2C(\mu)}{1-t} \tag{11}$$

This equation defines the absolute level of the wages. When the workers work hard, their cost of effort is high, implying high wages to compensate for that in order to keep them on their participation constraints. When they do not work hard, the firm does not have to offer much compensation for this effort and sets wages on a low level.

I continue to derive the formulas for the wages by adding $10+11$ (to isolate W_1) or taking the difference of $10-11$ (to isolate W_2):

$$W_1 = \frac{1}{1-t}C(\mu) + \frac{1}{2g(0)}$$

$$W_2 = \frac{1}{1-t}C(\mu) - \frac{1}{2g(0)}$$

Notice that wages are symmetrically distributed around the average $\frac{1}{1-t}C(\mu)$. The spread is determined by $\frac{1}{2g(0)}$.

5 Optimal taxation

I will now look at the optimal tax system. This is done step by step. First, I postulate a social welfare function (SWF). Subsequently, the Government Budget Constraint (GBC) is introduced. Then I set up the Lagrangian and derive expressions for the optimal t and T.

This procedure will be repeated for 3 different situations, which differ in their handling of capital owners. There is the situation with **foreign** capital owners, with **local** capital owners and the situation where the capital owners are the **workers** themselves.

In the first situation, the profit earners of the firm, e.g. the firm owners, stakeholders etcetera are not included. This could be the case when, for example, the firm owners or stakeholders are foreigners. Redistributing to and from the profit earners is impossible. That means that profits go untaxed, and that there is no profit earner eligible for receiving the general transfer T.

However, if the profit earners are inhabitants under the same government, they usually have a non-zero, positive welfare weight. This is the case in the second model, where the capital owners benefit from the firms' profits. This can, for example, represent the situation where stockholders have taken ownership of the firm and obtain dividends from it. Usually, firms pay (part of) their profits to the shareholders in dividends. The higher the profits, the more room there is for dividend payments, and the more the shareholders (who are in fact capital owners) earn from their stocks. It is possible to include these capital owners in the SWF. That is what is carried out in this section. Then I take the same steps of analysis as in the previous section.

In the third case, I assume that the two workers are the ones who earn the profits. This could reflect the case where they are both owner-directors of the firm, or when the firm is a partnership. In these cases, the profits ultimately belong to the workers. Although they would likely act very cooperative and not merely as competitors in such a case, the tournament model can then still be applied. After all, they will still exert effort as long as it positively impacts their expected net income, and there could still be cases when one of the partners ultimately obtains a larger share of the wages.

5.1 The social welfare function

The SWF is the function that defines the preferences of the government with respect to welfare. Usually, the government maximizes a weighed sum of utilities of the inhabitants. They do not care about the ex-ante expected utilities, but about ex-post utilities. Ex-post utilities means utilities after the tournament has been played, so when the winner and loser with corresponding gross wages are known. The government can, to some extent, have a preference for equality of these ex-post utilities. It is possible to fulfill the preference for equality through income redistribution. The instruments the government can use for this are embedded in their income taxation policy. So, there is a policy-directed relationship between the preferences for social welfare and the income taxation scheme. This relationship was used by Jacobs, Jongen, and Zoutman (2017) to determine the true preferences for social welfare of political parties in the Netherlands, based on the parties' proposals for an income taxation scheme.

If the government has preferences for equality of outcomes, it cares about the ex post utility levels of the agents. Then the government acts as if it wants to maximize a social welfare function $\phi(U_1, U_2)$ that is symmetric for U_1 and U_2 and has $\frac{d\phi}{dU_1} > 0$ and $\frac{d\phi}{dU_2} > 0$. It also has the property that $\frac{d\phi}{dU_1} \leq \frac{d\phi}{dU_2}$. This means that the welfare weight attached to the winner is smaller than the welfare weight attached to the loser: according to the government, social welfare increases stronger when the income of the loser (i.e. the lower income) increases than when the income of the winner (i.e. the higher income) increases.

Let us label the social welfare weights $\alpha_1 = \frac{d\phi}{dU_1}$ for the winner and $\alpha_2 = \frac{d\phi}{dU_2}$ for the loser. When there are third-party capital owners involved, the government can value their welfare via the social welfare weight α_π.

5.2 The government budget constraint

Just like in most other optimal taxation models (see e.g. Mirrlees (1971)), the income tax distorts the labor supply of the workers. Thus, it is optimal to redistribute until the point is reached where the welfare gains of redistribution no longer outweigh the costs of the distortion in labor supply. I will now derive where that point is located in this model. Therefore, first the budget of the government that is available for redistribution needs to be defined. The universal tax rate is levied upon the wages of the winner and the loser. With the money collected with this tax, the universal government transfer can be paid.

6 Optimal taxation with foreign capital owners

The following formula defines the SWF that I will use:

$$\phi = \alpha_1 U_1 + \alpha_2 U_2 = (\alpha_1 W_1 + \alpha_2 W_2)(1-t) + (\alpha_1 + \alpha_2)[T - C(\mu)]$$

Without loss of generalisation, $\alpha_1 + \alpha_2 = 2$: using this does not alter the reasoning, analysis or outcomes. Social welfare in this simple form can be increased by imposing taxes that redistribute between U_1 and U_2.

Concerning the GBC, the following equation characterizes the constraint:

$$t \cdot (W_1 + W_2) = 2T \tag{12}$$

I can now combine the SWF and the GBC to obtain the Lagrangian that is going to be used to calculate the optimal tax rate t^* and the corresponding universal transfer T^*, which together determine the tax scheme. The expression for this Lagrangian is as follows:

$$\mathcal{L} = (1-t)(\alpha_1 W_1 + \alpha_2 W_2) + (\alpha_1 + \alpha_2)[T - C(\mu)] + \gamma[t \cdot (W_1 + W_2) - 2T]$$

6.1 The optimal tax scheme

The government wants to set the optimal tax scheme. Therefore, I calculate the derivatives of the Lagrangian with respect to T and t:

$$\frac{d\mathcal{L}}{dT} = (\alpha_1 + \alpha_2) - 2\gamma = 0 \quad \Longrightarrow \quad \gamma = \bar{\alpha} = \frac{\alpha_1 + \alpha_2}{2} = 1 \tag{13}$$

$$\begin{aligned} \frac{d\mathcal{L}}{dt} = &- (\alpha_1 W_1 + \alpha_2 W_2) + \gamma(W_1 + W_2) + (1-t)\left(\alpha_1 \frac{dW_1}{dt} + \alpha_2 \frac{dW_2}{dt}\right) \\ &- (\alpha_1 + \alpha_2) C'(\mu) \frac{d\mu}{dt} + \gamma t \left(\frac{dW_1}{dt} + \frac{dW_2}{dt}\right) = 0 \end{aligned} \tag{14}$$

This can be rewritten in order to obtain more usable expressions. To this end, I use the definition of the average wage: $\bar{W} \equiv (W_1 + W_2)/2$. Notice that the derivatives of the wages with respect to the tax rate t is equal for both W_1 an W_2. Therefore, I can write $\frac{dW_1}{dt} = \frac{dW_2}{dt} = \frac{d\bar{W}}{dt}$. Furthermore, from expression 13, it follows that $\gamma = 1$. Lastly, I can use the result from equation 9: $C'(\mu) = 1 - t$. With these helpful notes equation 14 can be rewritten:

$$2\bar{W} - \alpha_1 W_1 - \alpha_2 W_2 + (1-t)(\alpha_1 + \alpha_2)\left(\frac{d\bar{W}}{dt} - \frac{d\mu}{dt}\right) + 2t\frac{d\bar{W}}{dt} = 0$$

Notice that this expression contains the derivative of the average wage with respect to the tax rate, which depicts how the wage level responds to changes in the tax rate. This derivative is specified in more detail:

$$\frac{d\bar{W}}{dt} = \frac{(C'(\mu) \cdot (1-t)\frac{d\mu}{dt} - C(\mu) \cdot -1}{(1-t)^2} = \frac{d\mu}{dt} + \frac{C(\mu)}{(1-t)^2} = \frac{d\mu}{dt} + \frac{\bar{W}}{1-t} \tag{15}$$

From this equation it can already be seen that the impact of a tax rate change on the wage goes via both a behavioral effect (depicted by $\frac{d\mu}{dt}$) and a mechanical effect (depicted by $\frac{\bar{W}}{1-t}$). Some more explanation will be given in section 6.2. I substitute this in the previous expression to obtain

$$2\bar{W} - \alpha_1 W_1 - \alpha_2 W_2 + (\alpha_1 + \alpha_2)\bar{W} + 2t\frac{d\bar{W}}{dt} = 0$$

As already shown, it is possible to impose that $\alpha_1 + \alpha_2 = 2$ without loss of generality. I use this to rearrange the above expression:

$$\frac{t}{1-t}\left(\frac{-d\bar{W}}{dt}\frac{1-t}{\bar{W}}\right) = 1 - \left(\frac{1}{2}\alpha_1\frac{W_1}{\bar{W}} + \frac{1}{2}\alpha_2\frac{W_2}{\bar{W}} - 1\right)$$

It can be defined that

$$e \equiv \frac{-d\bar{W}}{dt}\frac{1-t}{\bar{W}} \tag{16}$$

where e is the elasticity of taxable income with respect to the net-of-tax rate $(1-t)$. This elasticity measures how the average wage responds to the net-of-tax rate. Say for example that the government increases the tax rate. Then the net-of-tax rate decreases. If then the firm decreases the wages in response, this elasticity is positive. With this definition, I can rewrite the above expression further:

$$\frac{t}{1-t} \cdot e = 1 - \left(\frac{1}{2}\alpha_1\frac{W_1}{\bar{W}} + \frac{1}{2}\alpha_2\frac{W_2}{\bar{W}} - 1\right)$$

I can also use the definition of the covariance to obtain more insight in the expression. Remember that the covariance is defined as follows:

$$cov(X,Y) = E[(X - E(X))(Y - E(Y))] = E(XY) - E(X)E(Y) \tag{17}$$

Here, $E(\cdot)$ denotes the expected value. Applying this to my model, and using again the fact that $\bar{\alpha} = 1$, the following is obtained:

$$cov\left(\alpha_i, \frac{W_i}{\bar{W}}\right) = E\left(\alpha_i \frac{W_i}{\bar{W}}\right) - E(\alpha_i)E\left(\frac{W_i}{\bar{W}}\right) = \left(\frac{1}{2}\alpha_1 \frac{W_1}{\bar{W}} + \frac{1}{2}\alpha_2 \frac{W_2}{\bar{W}}\right) - 1 \cdot 1 \leq 0, \quad i = 1, 2$$

which is the covariance between the welfare weights and the normalized wage, and captures the degree of inequality aversion of the government. After implementing this covariance, the expression for the optimal tax looks like this:

$$\frac{t^*}{1-t^*} = \frac{1 - cov\left(\alpha_i, \frac{W_i}{\bar{W}}\right)}{e}$$

This tells us that the optimal tax rate increases with the degree of inequality aversion. So, when does the covariance become more negative? This is fueled by two reasons. The first one is when the government becomes more egalitarian: then it values the welfare of the loser relatively more compared to the welfare of the winner and thus assigns a relatively large welfare weight α_2 and a very small α_1 to the respective income earners.

The second reason for the covariance to be small is when the difference between the wages of the winner and the loser is very big. So, when there is more pretax inequality, the government has a bigger job to do when it comes to redistribution. Both mechanisms make the covariance more negative, implying a higher optimal tax rate.

6.2 The elasticity of the wage with respect to the net-of-tax rate

The tax rate decreases with the wage-elasticity with respect to the net-of-tax rate: the more negative the wage responds to changes in the net-of-tax rate, the lower the tax rate. It is important to know whether this elasticity is positive, because a negative elasticity would imply a tax rate below zero or above unity. In order to determine the sign of the elasticity, I take the following steps. It can be derived that

$$e = -\frac{d\bar{W}}{dt}\frac{1-t}{\bar{W}} = -\left(\frac{d\mu}{dt} + \frac{\bar{W}}{1-t}\right)\frac{1-t}{\bar{W}} = -\frac{d\mu}{dt}\frac{1-t}{\bar{W}} - 1 \tag{18}$$

Now there are 4 cases.

If there is no behavioral response from the workers to a change in the net-of-tax rate, i.e. when $-\frac{d\mu}{dt} = 0$, e is negative and equal to -1. The -1 implies that for every additional unit of net income that the workers obtain due to a tax rate change, the firm can lower the wages with the same amount. And vice versa: for every additional unit of net income that the workers must give up due to a tax rate change, the firm must compensate the workers with the same amount to keep the participation constraint binding. The government can raise more taxes for redistribution with zero consequences for the productivity of the workers. This leaves the optimal tax rate unboundedly high.

If there is a positive behavioral response from the workers to a change in the tax rate, which is when workers are going to work more when the tax rate increases and vice versa, i.e. when $\frac{d\mu}{dt} > 0$ or $-\frac{d\mu}{dt} < 0$, we have to do with $e < -1$. Economically, this is a case that can be explained through the income effect: workers are going to work harder when they must pay more taxes to compensate for the loss in income. But, since I have excluded income effects, this situation exists only theoretically in my model. The firm responds by increasing wages to compensate the workers for the additional costs of effort they bear to meet the participation constraint. Also in this case, it is optimal for the government to set the tax rate unboundedly high, since a higher tax rate now not only gives more tax revenue for redistribution, but at the same time increases the tax base.

If there is a small negative behavioral response from the workers to a change in the tax rate, i.e. when $0 < -\frac{d\mu}{dt} < \frac{\bar{W}}{1-t}$. In that case, $-1 < e < 0$. Such a negative behavioral response characterizes the **substitution effect**. This means that there is some deadweight loss when increasing taxes: workers choose to work less because the surplus of their productivity partially flows to the government via taxes. So, the firm can offer lower wages, since it has to compensate the workers for a less high cost of effort. But, at the same time, the tax increase directly hits the net income of the workers. This is the **mechanical effect**. The firm must in turn offer higher wages to maintain the workers' participation. In this case, the latter effect is bigger. So, while there is some economic loss, this effect is not sufficiently large to alter the optimal tax scheme: the redistribution gains still outweigh the deadweight loss, and hence, the optimal tax rate is still unboundedly high.

However, it could be that the substitution effect is large enough, i.e. when $-\frac{d\mu}{dt} > \frac{\bar{W}}{1-t}$. Then, the workers are going to work significantly less when the tax rate increases, which in turn implies that $e > 0$: the wages decrease even when the net-of-tax rate decreases because workers are so much less productive. The optimal tax rate in this situation is somewhere between zero and unity.

6.3 Determining the sign of the elasticity

It is possible to use another method to investigate whether e is positive or negative. To this end, I rewrite $\frac{d\mu}{dt}$ to be able to express e in terms of the cost of effort $C(\mu)$ and its derivatives and then make an additional assumption about the functional form of $C(\mu)$. To help reformulating $\frac{d\mu}{dt}$, I again use the fact from equation 9 that $C'(\mu) = 1 - t$. I take the total derivative and write

$$dC'(\mu) = d(1-t) \quad \Longrightarrow \quad C''(\mu)d\mu = -dt \quad \Longrightarrow \quad \frac{d\mu}{dt} = -\frac{1}{C''(\mu)} \tag{19}$$

This makes it possible to rewrite e in terms of the cost of effort function and its derivatives, again using the result from equation 9:

$$e = -\frac{d\bar{W}}{dt}\frac{1-t}{\bar{W}} = -\left(\frac{d\mu}{dt} + \frac{\bar{W}}{1-t}\right)\frac{1-t}{\bar{W}} = -\frac{-1}{C''(\mu)}\frac{1-t}{\frac{C(\mu)}{1-t}} - 1 = \frac{[C'(\mu)]^2}{C''(\mu)C(\mu)} - 1 \tag{20}$$

Now I am ready to make a more specific assumption concerning the expression for $C(\mu)$. I can use the following isoelastic function where $\eta > 0$:

$$C(\mu) = \frac{\mu^{1+\frac{1}{\eta}}}{1+\frac{1}{\eta}} \quad \Longrightarrow \quad C'(\mu) = \mu^{\frac{1}{\eta}} \quad \Longrightarrow \quad C''(\mu) = \frac{1}{\eta}\mu^{\frac{1}{\eta}-1}$$

I reformulate equation 20 a bit:

$$e = \frac{C'(\mu)}{\mu C''(\mu)}\frac{\mu C'(\mu)}{C(\mu)} - 1$$

I draw the result up by implementing the expressions for $C(\mu)$ and its derivatives:

$$\frac{C'(\mu)}{\mu C''(\mu)} = \frac{\mu^{\frac{1}{\eta}}}{\mu \cdot \frac{1}{\eta}\mu^{\frac{1}{\eta}-1}} = \eta$$

$$\frac{\mu C'(\mu)}{C(\mu)} = \frac{\mu \cdot \mu^{\frac{1}{\eta}}}{\frac{\mu^{\frac{1}{\eta}+1}}{\frac{1}{\eta}+1}} = \frac{1}{\eta} + 1$$

Then I use this to simplify the expression for e and draw a conclusion concerning its sign:

$$e = \underbrace{\frac{C'(\mu)}{\mu C''(\mu)}}_{=\eta}\underbrace{\frac{\mu C'(\mu)}{C(\mu)}}_{=\frac{1}{\eta}+1} - 1 = (1+\eta) - 1 = \eta > 0$$

So, if the cost of effort function has the above isoelastic form, the elasticity of the wage with respect to the net-of-tax rate is positive. This in turn implies an optimal tax rate $0 < t^* < 1$.

6.4 The revenue maximizing tax rate

Now the possible values that the optimal tax rate can take are determined a bit more specific. To this end, I calculate the revenue maximizing tax rate. The revenue maximizing tax rate is located at the top of the Laffer curve, which is the curve that denotes the tax revenue as a function of the tax rate (see e.g. Laffer (2004)). Total tax revenue is characterised by the function $\tau = t(W_1 + W_2) = 2t\bar{W}$. Maximizing this with respect to t leads to the following steps:

$$\frac{d\tau}{dt} = 2\bar{W} + 2t\frac{d\bar{W}}{dt} = 0 \quad \Longrightarrow \quad \frac{\bar{W}}{\frac{d\bar{W}}{dt}} = -t \quad \Longrightarrow \quad \frac{1}{\frac{d\bar{W}}{dt}\frac{1-t}{\bar{W}}} = -\frac{t}{1-t} \quad \Longrightarrow \quad \frac{t}{1-t} = \frac{1}{e}$$

This means that even when there are no redistributive preferences (i.e. when $cov\Big(\alpha_i, \frac{W_i}{\bar{W}}\Big) = 0$), the optimal tax rate maximizes revenue. This is because the firm compensates the workers through offering higher wages when taxes increase. Why is that? Remember that the ex ante expected utility of the workers is just equal to the value of the outside option T, because the firm catches all the surplus from labor productivity. A tax increase makes workers worse off: their expected utility falls below T. The firm wants the workers to keep participating and therefore increases their gross wages until the level where their expected utility is back at the initial level, the so-called “compensation reflex”. Hence, a tax increase is fully at the expense of the firm. And since the government attaches no value to the utility of the firm owner, increasing taxes until the top of the Laffer curve is a “free lunch”: the workers are compensated for any income loss while redistributional gains are obtained by raising the tax revenue. So, it is now clear that in the optimum, the tax rate has a lower boundary at

$$\frac{t}{1-t} = \frac{1}{e} \implies t = \frac{1}{1+e}$$

When the government does have preferences to redistribute from the winner to the loser, the optimal tax rate could go even beyond the top of the Laffer curve. But, beyond this revenue maximizing rate, tax revenue declines. Also, labor supply distortions become even stronger, increasing the deadweight loss. But, this loss is completely incurred by the firm, so it does not affect social welfare. While tax revenue declines, still some redistribution takes place. This happens through the fact that the wage difference is constant, as can be seen from formula 10 making the winner pay a larger share of the tax revenue increase. However, the “compensation reflex” from the firm affects both workers equally. Notice that this is only valid under the assumption that the value of T remains unaffected by the tax revenue that the government extracts from this tournament. If the value of T would decline with the tax revenue, the level of utility would be impacted. In that situation, the relevant tradeoff for the government would be between the ex post equality in utility levels versus the ex ante absolute level of expected utility (which is already equal for both workers!). Notice that this would be Pareto-inefficient: starting from a tax rate beyond the top of the Laffer curve, it is then possible to reduce the tax rate, which makes nobody worse off and generates more tax revenue for redistribution purposes through T. Essentially, if the government would nonetheless choose increase the tax rate to beyond the top of the Laffer curve, the government attaches a negative social welfare value to the winner (Lorenz & Sachs, 2016; Jacobs et al., 2017).

The above discussion is applicable in the case where the government attaches no value to the firm owner. But, what if the government actually cares about the utility of profit earners?

7 Optimal taxation with local capital owners

To include the capital owners in the SWF, I first must establish the income and the utility function of the profit earners. I assume that profit earners do not incur costs of effort. They do receive the general government transfer T, since this transfer is universal: it is independent of whether you have a job, and independent of whether you own shares. The income of the capital owners is correlated with the firms' profits. Therefore, I use the profits to describe their income. Profits are characterised by equation 4. The income of capital owners cannot go untaxed. Therefore, the tax rate is also applied to the profits. With this information, the income of capital owners can be described as follows:

$$Y_\pi = (1-t)\pi + T = (1-t)(2\mu + \varepsilon_i + \varepsilon_j - W_1 - W_2) + T$$

The utility function is determined the same way as in the workers' case, via equation 1:

$$U_\pi = (1-t)(2\mu + \varepsilon_1 + \varepsilon_2 - W_1 - W_2) + T$$

The SWF can be established in the following way:

$$\begin{aligned}\phi = \quad & \alpha_1 U_1 + \alpha_2 U_2 + \alpha_\pi U_\pi = (1-t)(\alpha_1 W_1 + \alpha_2 W_2) - (1-t)\alpha_\pi (W_1 + W_2) + (\alpha_1 + \alpha_2 + \alpha_\pi)T \\ & - (\alpha_1 + \alpha_2)C(\mu) + (1-t)\alpha_\pi(2\mu + \varepsilon_1 + \varepsilon_2)\end{aligned}$$

The GBC looks like this:

$$t \cdot (W_1 + W_2) + t \cdot (2\mu + \varepsilon_1 + \varepsilon_2 - W_1 - W_2) = t \cdot (2\mu + \varepsilon_1 + \varepsilon_2) = 3T$$

For simplicity, I use that $\frac{\varepsilon_1+\varepsilon_2}{2} \equiv \bar{\varepsilon}$. Then, the Lagrangian is composed:

$$\begin{aligned}\mathcal{L} = \quad & (1-t)(\alpha_1 W_1 + \alpha_2 W_2) - (1-t)\alpha_\pi(W_1+W_2) + (\alpha_1+\alpha_2+\alpha_\pi)T - (\alpha_1+\alpha_2)C(\mu) \\ & + (1-t)\alpha_\pi(2\mu+2\bar{\varepsilon}) + \gamma[t\cdot(2\mu+\varepsilon_1+\varepsilon_2) - 3T]\end{aligned}$$

Without loss of generality, I now use that $\alpha_0+\alpha_1+\alpha_2 = 3$, such that $\frac{\alpha_1+\alpha_2+\alpha_\pi}{3} = \bar{\alpha} = 1$.

7.1 The optimal tax scheme

I proceed with the calculation of the derivatives with respect to T and t, in order to obtain the expressions for the optimal T^* and t^*:

$$\frac{d\,\mathcal{L}}{dT} = \alpha_1+\alpha_2+\alpha_3 - 3\gamma = 0 \qquad \Longrightarrow \qquad \gamma = \bar{\alpha} = 1$$

$$\begin{aligned}\frac{d\,\mathcal{L}}{dt} = & -(\alpha_1 W_1+\alpha_2 W_2) + \alpha_\pi(W_1+W_2) - \alpha_\pi(2\mu+2\bar{\varepsilon}) + \gamma(2\mu+2\bar{\varepsilon}) + (1-t)\left(\alpha_1\frac{dW_1}{dt}+\alpha_2\frac{dW_2}{dt}\right) \\ & -(1-t)\alpha_\pi\left(\frac{dW_1}{dt}+\frac{dW_2}{dt}\right) - (\alpha_1+\alpha_2)C'(\mu)\frac{d\mu}{dt} + (1-t)\alpha_\pi 2\frac{d\mu}{dt} + \gamma t 2\frac{d\mu}{dt} = 0\end{aligned}$$

Some rules can be applied to simplify this. Use the fact that $\frac{W_1+W_2}{2} = \bar{W}$, that $\frac{dW_1}{dt} = \frac{dW_2}{dt} = \frac{d\bar{W}}{dt}$, $\gamma = 1$ and $C'(\mu) = 1-t$ and then write

$$\begin{aligned}& -(\alpha_1 W_1+\alpha_2 W_2) + \alpha_\pi(W_1+W_2) - \alpha_\pi(2\mu+2\bar{\varepsilon}) + 2\mu + 2\bar{\varepsilon} + (1-t)(\alpha_1+\alpha_2)\frac{d\bar{W}}{dt} \\ & -(1-t)2\alpha_\pi\frac{d\bar{W}}{dt} - (1-t)(\alpha_1+\alpha_2)\frac{d\mu}{dt} + (1-t)2\alpha_\pi\frac{d\mu}{dt} + 2t\frac{d\mu}{dt} = 0\end{aligned}$$

Now use $\frac{d\bar{W}}{dt} = \frac{d\mu}{dt} + \frac{\bar{W}}{1-t}$, as can be seen from formula 15:

$$\begin{aligned}& -(\alpha_1 W_1+\alpha_2 W_2) + \alpha_\pi(W_1+W_2) + (1-\alpha_\pi)(2\mu+2\bar{\varepsilon}) + (1-t)(\alpha_1+\alpha_2)\frac{d\mu}{dt} + (\alpha_1+\alpha_2)\bar{W} \\ & -(1-t)2\alpha_\pi\frac{d\mu}{dt} - 2\alpha_\pi\bar{W} - (1-t)(\alpha_1+\alpha_2)\frac{d\mu}{dt} + (1-t)2\alpha_\pi\frac{d\mu}{dt} + 2t\frac{d\mu}{dt} = 0\end{aligned}$$

Simplifying leads to

$$-(\alpha_1 W_1+\alpha_2 W_2) + (1-\alpha_\pi)(2\mu+2\bar{\varepsilon}) + (\alpha_1+\alpha_2)\bar{W} + 2t\frac{d\mu}{dt} = 0$$

Remember the definition for the elasticity of the wage with respect to th net-of-tax rate from equation 16. From formula 15 and definition 16, one can obtain that

$$-\frac{d\mu}{dt} = (e+1)\frac{\bar{W}}{1-t} \tag{21}$$

Remember also the definition for the covariance from equation 17. This definition can still be used to show that

$$cov\left(\alpha_i, \frac{W_i}{\bar{W}}\right) = \frac{1}{2}\alpha_1\frac{W_1}{\bar{W}} + \frac{1}{2}\alpha_2\frac{W_2}{\bar{W}} - \frac{\alpha_1+\alpha_2}{2}\cdot 1, \quad i = 1,2$$

With the help of these definitions, I can rewrite to

$$\frac{\mu+\bar{\varepsilon}}{\bar{W}}(1-\alpha_\pi) - cov\left(\alpha_i, \frac{W_i}{\bar{W}}\right) = \frac{t}{1-t}(e+1)$$

This leads to following formula for the optimal tax rate

$$\frac{t^*}{1-t^*} = \frac{\frac{\mu+\bar{\varepsilon}}{\bar{W}}(1-\alpha_\pi) - cov\left(\alpha_i, \frac{W_i}{\bar{W}}\right)}{e+1}$$

Some remarks are required regarding this formula. First, note that it is the ratio of effort versus the wage level that influences the tax rate. Remember that effort directly translates into profits for the firm. This implies that if the firm makes high profits relative to the wage level, the optimal tax rate is high, and vice versa. The $(1-\alpha_\pi)$ term denotes the concern of the government about the welfare of the capital owners. The higher their welfare weight, the lower the tax rate, since a tax increase goes at the expense of the capital owners, like I have discussed in section 6.4. Furthermore, it is still the degree of inequality aversion that influences the optimal tax rate: the more the government cares about the losers' income, and the lower the relative income of the loser is, the higher the tax rate. Lastly, $(e+1)$ appears in the numerator. Concerning what I have discussed in section 6.2, the situation has become more simple: again using formula 15, we have that

$$e+1 = \frac{-d\bar{W}}{dt}\frac{1-t}{\bar{W}} + 1 = \frac{-d\mu}{dt}\frac{1-t}{\bar{W}} - \frac{\bar{W}}{1-t}\frac{1-t}{\bar{W}} + 1 = \frac{-d\mu}{dt}\frac{1-t}{\bar{W}}$$

So, as long as there is a substitution effect, i.e. when $-\frac{d\mu}{dt} > 0$, then $e+1>0$, implying a nonzero, positive tax rate.

7.2 The revenue maximizing tax rate

The question can be posed whether the top of the Laffer curve (i.e., where tax revenue is at its maximum) serves as a lower bound again. To answer this, the formula for total tax revenue τ is postulated:

$$\tau = t(W_1+W_2) + t(2\mu + 2\bar{\varepsilon} - W_1 - W_2) = t(2\mu+2\bar{\varepsilon})$$

This can be maximized by calculating the derivative with respect to t. Remember that ε_1, ε_2 and thus $\bar{\varepsilon}$ are independent:

$$\frac{d\tau}{dt} = (2\mu + 2\bar{\varepsilon}) + 2t\frac{d\mu}{dt} = 0$$

This can be rewritten to

$$t = \frac{\mu + \bar{\varepsilon}}{-\frac{d\mu}{dt}}$$

Use formula 21 and rewrite further to obtain

$$\frac{t}{1-t} = \frac{\frac{\mu+\bar{\varepsilon}}{\bar{W}}}{e+1}$$

Again it can be seen that the revenue maximizing tax rate is similar to the optimal tax rate, but without redistributive preferences: in this case, the government attaches no welfare weights to any of the agents in the model. Contrary to the revenue maximizing tax rate in chapter 5, this expression does not necessarily provide a minimum for the optimal tax rate. It could be higher than the revenue maximizing tax rate if there is much inequality between the workers, or when the government attaches very different welfare weights to the workers. The optimal tax rate could also be lower if the government attaches enough value to the welfare of capital owners compared to the workers.

8 Optimal taxation with workers as capital owners

For the analysis of this situation, I assume that the workers share equally in the profits. The profit function is expressed in formula 4. When these profits become incorporated in their net income, they are of course taxed. There is no general transfer T for the firm, since the firm is owned by the two workers, who themselves already receive the transfer, contrary to the setting with capital owners. Using again $\frac{\varepsilon_1+\varepsilon_2}{2} \equiv \bar{\varepsilon}$ and $\frac{W_1+W_2}{2} \equiv \bar{W}$, the profit function can be formulated as follows:

$$\pi = 2\mu + 2\bar{\varepsilon} - 2\bar{W}$$

Hence,

$$U_1 = (1-t)W_1 + T - C(\mu) + \frac{1}{2}\pi(1-t) = (1-t)W_1 + T - C(\mu) + (1-t)(\mu + \bar{\varepsilon} - \bar{W})$$

and

$$U_2 = (1-t)W_2 + T - C(\mu) + (1-t)\frac{1}{2}\pi = (1-t)W_2 + T - C(\mu) + (1-t)(\mu + \bar{\varepsilon} - \bar{W})$$

The SWF the government wants to maximize now becomes

$$\phi = \alpha_1 U_1 + \alpha_2 U_2 = (1-t)(\alpha_1 W_1 + \alpha_2 W_2) + (\alpha_1 + \alpha_2)[T - C(\mu) + (1-t)(\mu + \bar{\varepsilon} - \bar{W})]$$

The GBC now looks like

$$t \cdot (W_1 + W_2 + 2\mu + 2\bar{\varepsilon} - 2\bar{W}) = t \cdot (2\mu + 2\bar{\varepsilon}) = 2T$$

This leads to the following Lagrangian:

$$\mathcal{L} = (1-t)(\alpha_1 W_1 + \alpha_2 W_2) + (\alpha_1 + \alpha_2)[T - C(\mu) + (1-t)(\mu + \bar{\varepsilon} - \bar{W})] + \gamma[t \cdot (2\mu + 2\bar{\varepsilon}) - 2T]$$

8.1 The optimal tax scheme

The corresponding First Order Derivatives with respect to T and t are calculated in order to find expressions for their optimal values. I make again use of the definition $\alpha_1 + \alpha_2 = 2$.

$$\begin{aligned}
\frac{d\,\mathcal{L}}{dT} = &\quad (\alpha_1 + \alpha_2) - 2\gamma = 0 \qquad \Longrightarrow \qquad \gamma = \frac{\alpha_1 + \alpha_2}{2} = 1 \\
\frac{d\,\mathcal{L}}{dt} = &\quad -(\alpha_1 W_1 + \alpha_2 W_2) - (\alpha_1 + \alpha_2)[\mu + \bar{\varepsilon} - \bar{W}] + \gamma(2\mu + 2\bar{\varepsilon}) + (1-t)\left(\alpha_1 \frac{dW_1}{dt} + \alpha_2 \frac{dW_2}{dt}\right) \\
&\quad -(\alpha_1 + \alpha_2)\left[C'(\mu)\frac{d\mu}{dt} - (1-t)\left(\frac{d\mu}{dt} - \frac{d\bar{W}}{dt}\right)\right] + \gamma t 2 \frac{d\mu}{dt} = 0
\end{aligned}$$

Some helpful definitions and equations can be used to rewrite this. It is still true that $\frac{dW_1}{dt} = \frac{dW_2}{dt} = \frac{\bar{W}}{dt}$, that $\frac{\alpha_1 + \alpha_2}{2} \equiv \bar{\alpha} = 1$, $C'(\mu) = 1 - t$ and $\frac{d\bar{W}}{dt} = \frac{d\mu}{dt} + \frac{\bar{W}}{1-t}$. Rewriting then leads to

$$\begin{aligned}
&-(\alpha_1 W_1 + \alpha_2 W_2) - (\alpha_1 + \alpha_2)(\mu + \bar{\varepsilon}) + (\alpha_1 + \alpha_2)\bar{W} + 2\mu + 2\bar{\varepsilon} + (1-t)(\alpha_1 + \alpha_2)\frac{d\bar{W}}{dt} \\
&-(\alpha_1 + \alpha_2)\left[(1-t)\frac{d\mu}{dt} - (1-t)\frac{d\mu}{dt} + (1-t)\frac{d\bar{W}}{dt}\right] + 2t\frac{d\mu}{dt} = 0
\end{aligned}$$

and subsequently to

$$-(\alpha_1 W_1 + \alpha_2 W_2) + 2\bar{W} + 2t\frac{d\mu}{dt} = 0$$

which, using formulas 17, 21 and 16, gives

$$\frac{t^*}{1-t^*} = \frac{-cov\left(\alpha_i, \frac{W_i}{\bar{W}}\right)}{e+1}$$

The optimal tax rate still depends on the degree of inequality aversion. The denominator is equal to what I calculated in the case of section 7.1. Something important can be noticed when comparing to that case. There, we had that $\frac{\alpha_1+\alpha_2+\alpha_\pi}{3} \equiv \bar{\alpha} = 1$. If the welfare weight of the capital owners in that case would be equal to the average of the welfare weights of the workers, such that $\alpha_\pi = \frac{\alpha_1+\alpha_2}{2} = 1$, then the $\frac{\mu+\bar{\varepsilon}}{\bar{W}}(1-\alpha_\pi)$-term disappears and the optimal tax rate is equal to the optimal tax rate in the current case where the workers are the capital owners themselves.

8.2 The revenue maximizing tax rate

Notice also that the top of the Laffer curve does not change compared to the one calculated in section 7.2 because there is only a shift in where the government redistributes to, not where the taxes are levied on. This can also be derived formally. Say again that τ denotes the total tax revenue and that it is defined with the formula

$$\tau = t \cdot (W_1 + W_2 + 2\mu + 2\bar{\varepsilon} - 2\bar{W}) = t \cdot (2\mu + 2\bar{\varepsilon})$$

The derivative with respect to the tax rate then yields

$$\frac{d\tau}{dt} = 2\mu + 2\bar{\varepsilon} + 2t\frac{d\mu}{dt} = 0$$

Rewriting leads to

$$-t\frac{d\mu}{dt} = \mu + \bar{\varepsilon}$$

and, making again use of formula 21, subsequently to

$$\frac{t}{1-t} = \frac{\frac{\mu+\bar{\varepsilon}}{\bar{W}}}{e+1}$$

which is indeed the same as in the previous case.

9 Comparison with other literature

In this section, the results are reviewed in the context of the related literature discussed in section 2.

9.1 Optimal taxation à la Mirrlees

First, I make a comparison of the results of this paper with the general results of the optimal taxation theory based on the Mirrlees paper (Mirrlees, 1971). Typically, for example in

Piketty and Saez (2013), the optimal linear tax rate in a Mirrleesian setting with a competitive labor market and wages that equal the marginal productivity is given by

$$\frac{t_M^*}{1-t_M^*} = \frac{-cov\Big(\alpha_i, \frac{W_i}{W}\Big)}{e} \tag{22}$$

where t_M^* is the optimal tax rate. This looks very much like the results I found in this paper. In all three optimal tax schemes that I have presented, the degree of inequality aversion is present, as well as the elasticity of the tax base with respect to the net-of-tax rate. However, there are two main differences in my findings compared to the Mirrleesian standard result. First, in the cases where the capital owners are not the workers, there is an additional positive term in the numerator, driving the optimal rate upwards. Second, in the cases where the capital owners are being taxed, there is a positive term in the denominator that neutralizes the mechanical effect that is incorporated in the elasticity. Let us look closer at these differences.

First, I look at the additional term in de numerator. In the case with foreign, untaxable capital owners, one can write

$$\frac{t^*}{1-t^*} = \frac{1}{e} + \frac{-cov\Big(\alpha_i, \frac{W_i}{W}\Big)}{e}$$

as the expression for the optimal tax rate. What happens here is that, although it is impossible to tax the capital owners directly, they still can be drained by the government via their relationship with the workers. The government freely raises taxes up to the revenue maximizing tax rate, and the firm responds by raising the wages to keep the workers participating in the job. That is what the $\frac{1}{e}$ part of the expression does. The other part is equal to the standard Mirrleesian result and tells us that the rate can be expanded to above the revenue maximizing rate, depending on the degree of inequality aversion. Theoretically, this can go on until the tax rate reaches 100%, since even with tax rates close to 100%, the firm still makes profits. This works as follows.

The revenue of the firm is linear with effort. The firm has no fixed costs, and the firms' variable costs are the wages, which in turn are dependent on the costs of effort and the tax rate. For the costs of effort, I have assumed that $C(\mu) > 0$, $C'(\mu) > 0$ and $C''(\mu) > 0$. So, for values of μ smaller than unity, the increase in $C(\mu)$ is smaller than the increase in μ, which implies a profit over that range[4].

[4]Because if $C'(\mu) = a \cdot \mu^b$, then $\lim_{\mu \to 0^+}(C'(\mu)) = 0$, $\forall\, b \in (0, \infty)$, while the increase in μ is linear and

Of course, when the tax rate reaches 100%, this story ends. At that point it would be senseless to exert any effort since the whole productivity is then taxed away.

In the case with local capital owners, the expression is somewhat more complicated:

$$\frac{t^*}{1-t^*} = \frac{\frac{\mu+\bar{\varepsilon}}{\bar{W}} - cov\left(\alpha_i, \frac{W_i}{\bar{W}}\right)}{e+1}$$

This is because the government can still drain the capital owners, this now comes at a cost since they are integrated in the tax system. The more it hurts their profits (symbolized by $\mu + \bar{\varepsilon}$), the lower the optimal tax. Furthermore, the government cares about capital owners when optimizing social welfare, which can be seen from the $1 - \alpha_\pi$ part. The more the government cares about them, the higher α_π and the lower the optimal tax rate will be.

Now I turn to the denominator. In this case, as well as in the case where workers are the capital owners, it is no longer only the elasticity that shows up there, but $e + 1$. Why is that? Well, as already explained, the elasticity is a constitution of both the direct mechanical effect and the indirect behavioral (substitution) effect. The mechanical effect is present since it is an automatic response from the firm to raise the wages when tax rate increases. What the workers lose in a tax regime change via the lower net-of-tax rate they get back from the firm in the form of higher gross wages. This holds for as long as I do not include the profits of capital owners in the optimization of welfare. But if the government decides otherwise, it takes into account that profits are hurt via this mechanism. The benefits from this "compensation reflex" for the workers exactly match the costs for the firm. Therefore, the mechanical effect is exactly cancelled out by the '1' in the denominator.

Concluding, it can be said that the central paradigm is similar in both the standard Mirrlees-type of optimal tax literature and in my model: it is optimal to redistribute as long as the welfare gains of this redistribution outweigh the costs of the distortion in labor supply. This conclusion holds even though my model differs in the setup of the way incomes are determined. Whereas in Mirrlees-like optax models, ability has a one-to-one relationship with earnings, in the case described in this paper other factors are taken into account via the noise term. Via this way, it is common in tournament-like models that two workers with

thus constant at 1. This also holds when $C(\mu)$ is scaled by $1-t$ for $\bar{W} = \frac{C(\mu)}{1-t}$ since the net-of-tax rate then acts as an additional constant like a.

the same earnings ability still end up with different net incomes. This is a phenomenon that is likely to occur in reality, but which is not captured in the Mirrlees model. It can now be seen that this absence of the one-to-one relationship does not lead to systematic differences in the optimal tax rates. Not only that, but also the formal expressions for the optimal linear rate are strikingly similar, employing the negative covariance between the welfare weights and the normalized wage (the degree of inequality aversion) and the elasticity of the tax base with respect to the net-of-tax rate. The differences stem from the different settings on the labor market with respect to capital owners. If the government includes them in the SWF, it must take into account that they are hurt by a tax increase through their compensation for the workers for the direct mechanical effect. This leads to a lower optimal tax rate than in the standard result in Mirrlees-like models and is explained in section 7 and 8. The "compensation reflex" can also be utilized by the government to tax the capital owners indirectly. This holds on the condition that the workers are not at the same time the capital owners of the firm. This leads to a higher tax rate compared to the standard Mirrleesian outcome and is explained in section 6 and 7.

9.2 Tournament model and similar researches

In this section, some comparative statics are calculated to see the differences in outcomes between this paper and that of Persson and Sandmo (2005). I discuss the differences and similarities and try to find an explanation for the differences. Remember that the main difference between their paper and this lies in the fact that the government transfer is universal in my model, rather than only granting it to people who actually work. This assumption forms the basis of the difference in the expression for the equilibrium wages. In their paper, equilibrium wages are formulated as follows:

$$W_{1,PS} = C(\mu) + \frac{1}{2g(0)}$$

$$W_{2,PS} = C(\mu) - \frac{1}{2g(0)}$$

The spread of the wages is determined by $\frac{1}{2g(0)}$, which is just the same as in the paper from PS. However, the average wage in my model is higher than in theirs. This is a result from the better bargaining position of the workers due to their improved outside option when not participating. Initially, their outside option was zero; now they receive T. If the firm wants the workers to participate, it must not only compensate them for the cost of effort they make, but also for this additional opportunity cost. Indeed, recalling equation 12, the difference between the average wage here compared to the average wage in Persson and Sandmo (2005) (which was simply $C(\mu)$) equals exactly the difference in the outside option:

$$\frac{1}{1-t}C(\mu) - C(\mu) = \frac{1-(1-t)}{1-t}C(\mu) = \frac{t}{1-t}C(\mu) = T$$

This difference in wages remains in terms of net incomes and utilities. Concerning the tax rate, they find that it is optimal to set the tax rate at the highest possible rate, which is 100%. They arrive at this result by taking the derivative of their SWF with respect to the tax rate. They then find that a tax rate increase always positively impacts welfare (Persson & Sandmo, 2005). Interestingly, welfare increases faster when the difference in wages is larger and when the difference in welfare weights is larger. This is very similar to the result in the setting that compares best to their model: with foreign capital owners. Also there, be it through the degree of inequality aversion, the government can have a stronger positive impact on welfare via tax increases if the difference in wages is larger and the difference in welfare weights is larger.

They also analyze the competitive case: where the firm is not a monopsonist in the labor market and workers consequently have market power. The optimal tax rate that they find there is characterized by

$$\frac{t^*_{PS}}{1-t^*_{PS}} = -\frac{1}{2g(0)} \cdot \frac{\phi_1 - \phi_2}{(\phi_1+\phi_2)\mu} \cdot \frac{1}{\eta_{PS}}$$

where $\phi_i,\ i = 1,2 \equiv \frac{d\phi}{dU_i}$ are the welfare weights they use and where η_{PS} is the elasticity of effort with respect to the net-of-tax rate (Persson & Sandmo, 2005). If I use their definition $\eta_{PS} \equiv \frac{d\mu}{d(1-t)}\frac{1-t}{\mu}$ and equation 15, their result can be rewritten to

$$\frac{t^*_{PS}}{1-t^*_{PS}} = \frac{-cov\Big(\phi_i, \frac{W_i}{W}\Big)}{e+1}$$

This appears to be very similar to my result from section 8 where the workers are the capital owners. The numerator is identical, except for the definition of the welfare weights. The denominator is completely identical, which can be explained intuitively as follows. In both situations, the worker's labor supply to the firm contributes to the firm's profits. Also, in both cases, workers benefit from the profits: in the model of Persson and Sandmo (2005), workers collect the firms' profits via higher wages, while in my model workers obtain the profits directly via their capital ownership. If a tax increase hits the labor supply of the workers via the substitution effect, profits go down accordingly, which in turn hurts the welfare of the workers. In both models, this goes through similar mechanisms as I explained above, and it therefore leads to the same expression in the denominator of the optimal tax rate.

9.3 Monopsony and the shift of surplus

This model brings forth one important implication concerning the battle for surplus between the different production factors. In a situation without the intervention of taxes, the firm can exploit the workers to the maximum since it has all the market power in the labor market. If then taxes are levied on the workers, the firm must choose between two options. Either it must accept that market forces make sure that all taxes levied on workers are passed on to the firm via the "compensation reflex", which can be seen as a shift on the intensive margin. Or, alternatively, it must get rid of the production factor labor altogether, which is a consequence on the extensive margin. This second option can be explored in an extension of this model where the firm can choose to replace labor by automated systems or technology. The workers will never be worse off since the firm offers in equilibrium a wage that makes workers just indifferent between working and retiring, which makes their utility level equal to the guaranteed income T. This implies that there is some redistribution between production factors possible: the return to entrepreneurship compared to the return to labor can be adjusted via taxes on labor.

10 Policy recommendations

Going back to my original research question, which was how the government can use the available instruments of redistribution to reach optimal welfare in a system with a monopsonistic firm and workers in a tournament setting, I can derive some recommendations for tax policy. I start with a general analysis of the tradeoffs that governments must make according to the model in this paper. Persson and Sandmo (2005) show that the government cannot via tax increases reduce inequality in incomes between the two workers. The firm will always make sure that the difference between the gross wages is large enough in order to provide a large enough incentive to win the tournament. As we have seen, a tax increase then only lowers the average wage level while keeping the wage difference constant, which surprisingly leads to more inequality in incomes. But, this effect is compensated by the reduction of the effort that workers put in, which subsequently lowers their costs of effort. Keeping this in mind, I arrive at the tradeoff between the welfare of the workers on the one hand, and that of the firm owners on the other hand. The economic reasoning for this is as follows. The firm has all the market power in the labor market. Therefore, the market itself ensures that the firm grabs all the surplus. This implies that any different tax scheme cannot make the workers worse off in expectation, since they are already at the lowest possible point on the intensive margin. The "compensation reflex" of the firm ensures that they will not change their behavior on the extensive margin and remain in the tournament. So, there are only two things the government must consider when setting the tax scheme. First, equality between the workers is not achievable due to the tournament setting. Second, the relative welfare of the capital owners to the workers can easily be influenced through taxes.

I now turn to the numerical values of the two most important values in the expressions for the optimal tax rates, namely the degree of inequality aversion and the elasticity of the tax base with respect to the net-of-tax rate. There is a lot of literature on the latter term, for example by Saez, Slemrod, and Giertz (2012), who argue that the elasticity of the tax base is, under certain assumptions, a sufficient statistic for the efficiency costs in tax policy analysis. Analyzing existing literature, they find that the most reliable estimates vary from 0.12 to 0.4, with a mean of 0.25. Another investigation is carried out by Chetty (2012), who focuses his analysis on frictions and finds a value of 0.33 for the compensated elasticity. Kleven and Schultz (2014) find an elasticity below 0.2 using a rich Danish dataset, but they also find that the elasticity is larger with larger tax increases (even up to values of 0.3 and 0.4 for some subgroups). This finding is supported by the theory that large tax rate changes can overcome factors like adjustment costs and frictions that blur the true value of the elasticity (Chetty, Friedman, Olsen, & Pistaferri, 2011).

Concerning the degree of inequality aversion, we know that it is bounded below one in my model, provided that both the wages and the welfare weights are nonnegative. I will briefly discuss four possible cases. The first case is when $W_2 = 0$ and the government has Rawlsian preferences, i.e. when it only cares about the lowest earner, implying $\alpha_1 = 0$ and $\alpha_2 = 2$. We then have that $-cov\Big(\alpha_i, \frac{W_i}{\bar{W}}\Big) = 1$. The fourth case is when the government is utilitarian, meaning that all agents have the same welfare weight. Then, $-cov\Big(\alpha_i, \frac{W_i}{\bar{W}}\Big) = 0$. The second and thirth case are somewhere in between: for the second case, I use $-cov\Big(\alpha_i, \frac{W_i}{\bar{W}}\Big) = \frac{2}{3}$, representing a leftist government, and for the third case $-cov\Big(\alpha_i, \frac{W_i}{W}\Big) = \frac{1}{3}$ which resembles a more right-wing government.

10.1 Optimal tax simulations

In Table 1, simulations for the optimal tax rate are carried out with different values for the elasticity and degree of inequality aversion for the three cases presented in chapters 6, 7 and 8. Notice that the formulas for the optimal tax rates are expressed in terms of $\frac{t^*}{1-t^*}$ and therefore should first be rewritten to be expressed in terms of just t^*. Since it would take up too much space to present all varieties of simulations, I limit them to the ones presented here, and assume for simplicity that $\mu + \bar{\varepsilon} = \bar{W}$ and that $\alpha_\pi = 0.5$ in all cases. Concerning the latter choice, we know that there is a long scientific debate about whether to tax capital gains at all. For example, Mankiw et al. (2009) say that capital should not be taxed, while Diamond and Saez (2011) argue that there should be a tax on capital. Empirics tell us that the share of capital income increases over the income distribution, which is clearly shown by Piketty and Saez (2003). Capital owners thus tend to be among the higher incomes, and are therefore expected to be assigned a relatively low welfare weight by the government. This justifies the choice for a below-average value for α_π. For comparison, I have also added simulations for the standard Mirrlees result, as presented in equation 22. Lastly, optimal tax rate percentages are rounded at one decimal for convenience.

Analyzing the simulations, it is clear that the optimal tax rate is very dependent on the setting with respect to capital ownership. When capital owners are not included in the SWF, optimal percentages are high. This sustains the theory that the government can freely tax the capital owners without the workers being hurt through the “compensation reflex” of the firm. What can also be seen is that the impact of different values of the elasticity is fairly low when capital owners are included in the SWF. The impact of the elasticity is then mitigated; both by the fact that there is capital that can be taxed as well as through the “compensation reflex” that ensures participation for the workers.

Notice that these tax rates are based on a model where redistribution is the only objective of the government. If the government also wants to raise funds for goals other than redistribution, the numerical simulations would of course be higher. Notice further that I have bounded the degree of inequality aversion by imposing restrictions on the welfare weights and the income levels. If these restrictions would be loosened, more extreme values of the

Table 1: Simulations for the optimal tax rate

	Foreign capital owners			Local capital owners		
$-cov(\cdot)$	$e = 0.12$	$e = 0.25$	$e = 0.4$	$e = 0.12$	$e = 0.25$	$e = 0.4$
1	94.3%	88.9%	83.3%	57.3%	54.5%	51.7%
$\frac{2}{3}$	93.3%	87.0%	80.6%	51.0%	48.3%	45.5%
$\frac{1}{3}$	91.7%	84.2%	76.9%	42.7%	40.0%	37.3%
0	89.3%	80.0%	71.4%	30.9%	28.6%	26.3%
	Workers as capital owners			Standard Mirrlees		
$-cov(\cdot)$	$e = 0.12$	$e = 0.25$	$e = 0.4$	$e = 0.12$	$e = 0.25$	$e = 0.4$
1	47.2%	44.4%	41.7%	89.3%	80.0%	71.4%
$\frac{2}{3}$	37.3%	34.8%	32.3%	84.7%	72.7%	62.5%
$\frac{1}{3}$	22.9%	21.1%	19.2%	73.5%	57.1%	45.5%
0	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%

$-cov\left(\alpha_i, \frac{W_i}{W}\right)$-term can occur, which makes more widespread values for the optimal tax rate possible. This term, representing the degree of inequality aversion, compiles both the wage difference and the welfare weights. Keep in mind that the latter is a purely political choice. The numbers presented here only provide insight in the impact of that choice.

11 Conclusions

This paper has tried to take a different point of view regarding optimal income taxation. The omission of factors other than earnings ability in the wage determination is addressed via the noise term in the tournament setting. Furthermore, the rising influence of superstar firms is taken into account by the monopsonistic labor market setting. It turns out that the optimal tax rate depends on the degree of inequality aversion of the government and the wage-elasticity with respect to the net-of-tax rate, which is similar to the standard Mirrlees result. In the comparison with Persson and Sandmo (2005), I find that, unlike in their paper, the optimal tax rate is finite and generally below 1. These are positive results. Nevertheless, this paper still has a very basic setup. To carry out a more elaborate investigation of this type of optimal taxation analysis, multiple expansions are possible, which are discussed below.

11.1 Suggestions for further research

The analysis carried out in this paper can be extended in several directions. This section mentions some of the possible ways to go in future research.

This model employs a linear income tax. It can be shown that a nonlinear tax rate always dominates a linear tax rate in terms of optimality. It would therefore be fruitful to allow for different tax rates at different income levels and explore the welfare gains that this generates. However, be careful since I've tried this and got stuck with exploding monsters of formulas.

One thing that is needed when we want different tax rates to be attached to different levels of income is an income distribution. This model lacks a any serious concept of an income distribution, albeit it is one of the crucial ingredients when a government wants to establish the optimal tax scheme. A possible avenue for further research is therefore to expand this model from two workers to a general case with an unspecified number of workers. This should be combined with a modified form of the model where it is possible to end up with any income, with different probability densities for all these incomes. This would be the way to create an income distribution.

It could also prove beneficial to explore the effects of attaching different ability levels to the agents. One could do that in such a way that, even with equal noise distributions, the ex ante probability of winning the tournament is not equal for everyone. However, it could be possible that a lower skilled agent takes the gamble to "steal" a promotion from a better counterpart if noise levels are sufficiently high. This would force the firm to design a more sophisticated tournament where noise can be reduced so that workers are being pushed to signal their true ability, for example by playing multiple rounds. For the government, it might then be possible to approach the Tinbergen rule more closely, since wages are then a more accurate reflection of the true earnings ability of the workers.

Furthermore, there is the issue of profits. Taking profit earners into account in the welfare evaluation can be done in different ways. In this paper it is done via capital owners who earn capital gains that are taxed the same way as wages. However, it is also possible to levy a separate tax on profits. That would be a useful extension, as it would improve the possibilities to redistribute between different production factors, something that might be of increasing importance given the developments related to the rising power of big firms in the labor market.

References

Autor, D., Dorn, D., Katz, L. F., Patterson, C., & Van Reenen, J. (2017). The fall of the labor share and the rise of superstar firms.

Berg, M. v. d., Boutorat, A., Garretsen, H., Mounir, A., & Stoker, J. (2019, 5). Managementkwaliteit en microdata: een dataverkenning (Tech. Rep. No. 1). 6412 EX Heerlen: CBS.

Boeri, T., & Van Ours, J. (2013). The economics of imperfect labor markets. Princeton University Press.

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

Chetty, R., Friedman, J. N., Olsen, T., & Pistaferri, L. (2011). Adjustment costs, firm responses, and micro vs. macro labor supply elasticities: Evidence from danish tax records. The quarterly journal of economics, 126(2), 749–804.

De Loecker, J., & Eeckhout, J. (2017). The rise of market power and the macroeconomic implications (Tech. Rep.). National Bureau of Economic Research.

Diamond, P., & Saez, E. (2011). The case for a progressive tax: from basic research to policy recommendations. Journal of Economic Perspectives, 25(4), 165–90.

Eggertsson, G. B., Robbins, J. A., & Wold, E. G. (2018). Kaldor and piketty's facts: The rise of monopoly power in the united states (Tech. Rep.). National Bureau of Economic Research.

Jacobs, B., Jongen, E. L., & Zoutman, F. T. (2017). Revealed social preferences of dutch political parties. Journal of Public Economics, 156, 81–100.

Karabarbounis, L., & Neiman, B. (2014). The global decline of the labor share. The Quarterly journal of economics, 129(1), 61–103.

Kleven, H. J., & Schultz, E. A. (2014). Estimating taxable income responses using danish tax reforms. American Economic Journal: Economic Policy, 6(4), 271–301.

Laffer, A. B. (2004). The laffer curve: Past, present, and future. Backgrounder, 1765, 1–16.

Lazear, E. P., & Rosen, S. (1981). Rank-order tournaments as optimum labor contracts. Journal of political Economy, 89(5), 841–864.

Lorenz, N., & Sachs, D. (2016). Identifying laffer bounds: A sufficient-statistics approach with an application to germany. The Scandinavian Journal of Economics, 118(4), 646–665.

Mankiw, N. G., Weinzierl, M., & Yagan, D. (2009). Optimal taxation in theory and practice. Journal of Economic Perspectives, 23(4), 147–74.

Mirrlees, J. A. (1971). An exploration in the theory of optimum income taxation. The review of economic studies, 38(2), 175–208.

Persson, M., & Sandmo, A. (2005). Taxation and tournaments. Journal of Public Economic Theory, 7(4), 543–559.

Piketty, T., & Saez, E. (2003). Income inequality in the united states, 1913–1998. The Quarterly journal of economics, 118(1), 1–41.

Piketty, T., & Saez, E. (2013). Optimal labor income taxation. In Handbook of public economics (Vol. 5, pp. 391–474). Elsevier.

Piketty, T., & Zucman, G. (2014). Capital is back: Wealth-income ratios in rich countries 1700–2010. The Quarterly Journal of Economics, 129(3), 1255–1310.

Rijksoverheid. (2020). Hoofdstuk 3. algemene bijstand. Retrieved 2020-03-27, from `https://wetten.overheid.nl/BWBR0015703/2020-01-01`

Saez, E. (2001). Using elasticities to derive optimal income tax rates. The review of economic studies, 68(1), 205–229.

Saez, E. (2002). Optimal income transfer programs: intensive versus extensive labor supply responses. The Quarterly Journal of Economics, 117(3), 1039–1073.

Saez, E., Slemrod, J., & Giertz, S. H. (2012). The elasticity of taxable income with respect to marginal tax rates: A critical review. Journal of economic literature, 50(1), 3–50.

Skåtun, J. D. (2017). Taxation, risk aversion, and the wage gaps in tournaments. Oxford Economic Papers, 69(3), 834–845.

Tinbergen, J. (1970). Belasting op bekwaamheid (tax on ability in dutch). Intermediair, 30(6), 1–3.

Waldman, M. (2007). Theory and evidence in internal labor markets.